UNIVERSITÀ DEGLI STUDI DI NAPOLI

FEDERICO II

PhD THESIS IN CHEMICAL SCIENCES

XXIX COURSE

INVESTIGATION OF MOLECULAR MECHANISMS INVOLVING PROTEIN TRAFFICKING

Tutor:

Prof. Maria MONTI

Coordinator:

Prof. Luigi PADUANO

Author:

Diana CANETTI

Amor ogni cosa vince

Leonardo da Vinci

Acknowledgments

UNIVERSITY OF NAPLES FEDERICO II

DEPARTMENT OF CHEMICAL SCIENCES

Prof. Maria MONTI

Prof. Piero PUCCI

Prof. Renata PICCOLI

Prof. Daria Maria MONTI

TELETHON INSTITUTE OF GENETICS AND MEDICINE

Dr. Roman POLISHCHUK

Prof. Giancarlo PARENTI

UNIVERSITY COLLEGE LONDON

Prof. Vittorio BELLOTTI

Dr. Patrizia MANGIONE

Dr. Graham TAYLOR

Table of Contents

Abbreviations

Acronym	Meaning
ACN	Acetonitrile
AL	Immunoglobulin light chain amyloidosis
AMBIC	Ammonium Bicarbonate
ApoAI	Apolipoprotein AI
APS	Ammonium Persulphate
BCS	Bathocuproine disulfonate
BSA	Bovine Serum Albumin
CFTR	Cystic fibrosis transmembrane conductance regulator
CHO	Chinese Hamster Ovary
CI-MPR	Cation independent-mannose-6-phosphate receptor
CID	Collision Induced Dissociation
Co-IP	Co-immunoprecipitation
CP	Ceruloplasmin
DDA	Data Dependent Acquisition
DTT	D,L- Dithiothreitol
EDTA	Acido EtilenDiamminTetraacetico
ER	Endoplasmic Reticulum
ERAD	ER-associated protein degradation
ERK	Extracellular signal-regulated kinase
ERT	Enzyme Replacement Therapy
ESI	Electrospray Ionization
FTICR	Fourier transform ion cyclotron resonance
GAA	Acid alpha-glucosidase
GlcNAc	N-Acetilglucosammine
GO	Gene ontology
GT	Gene Therapy
HDL	High density lipoproteins
HPLC	High Performance Liquid Chromatography
HSCT	Hematopoietic Stem Cell Transplantation
IAM	Iodoacetoamide
ICP-MS	Inductively coupled plasma mass spectrometry
IP	Immunoprecipitation
IT	Ion trap
JNK	c-Jun N-terminal kinase
LC	Liquid Chromatography
LSD	Lysosomal Storage Diseases
M6P	Mannose-6-phosphate
MALDI	Matrix Assisted Laser Desorption Ionization
MAPK	mitogen-activated protein kinase
MBD	Metal Binding Domains
MPR	Mannose-6-phosphate receptor
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MSD	Multiple Sulfatase Deficiency
MVB	Multi Vescicular endosome/Bodies
MW	Molecular weight
NB-DNJ	N-butyldeoxynojirimycin
nLC	nanoflow Cromatography Liquid
PAGE	PoliAcrylamide Gel Electrophoresis
PBS	Phosphate buffer saline
PC	Pre-cleaning
PCT	Pharmacological Chaperone Therapy
PD	Pompe diseasee
PM	Plasma membrane
Q	Quadrupole
rhGAA	Recombinant Enzyme GAA
ROS	reactive oxygen species
SDC	Sodium deoxycholate
SDS	Sodium dodecylsulfate
SRT	Substrate Reduction Therapy
TEMED	Tetramethylethylenediamine

Acronym	Meaning
TFA	Trifluoroacetic acid
TGN	Trans-Golgi Network
TGS	Tris-Glicine-SDS
TOF	Time Of Flight
TTR	Transthyretin
UPR	Unfolded Protein Response
WD	Wilson disease
WT	Wild type

Summary

The cellular protein traffic and the molecular mechanisms involved in the transport of each individual protein towards a specific destination are crucial events for cell physiology. The 2013 Nobel Prize in Physiology or Medicine honoured three scientists: Dr. Rothman, Dr. Schekman and Dr. Südhof, who have solved the mystery of how the cell organizes its complex transport system. They demonstrated that mechanisms employed by a cell to accomplish vesicle transport and fusion are independent by cell type, and are the same in all eukaryotic organisms. The description of the machinery regulating vesicle traffic, the major transport system in cells, represented a shift in our understanding of how the eukaryotic cell, characterised by a complex internal compartmentalization, organizes the routing of molecules and proteins to various intracellular destinations, as well as to the outside of the cell. The protein traffic plays a very critical role for a variety of physiological processes in which vesicle fusion and molecules delivery must be strictly controlled, such as during hormones and/ or cytokines release. In absence of this wonderful and precise organization, the cell would lapse into chaos. Defective vesicle transport leads to the development of a variety of neurological and immunological disorders, as well as in chronic pathologies, such as the diabetes. Many of these disorders occur in presence of mutation at level of different genes, that often give rise to proteins still catalytically active, but not able to reach the right cell compartment where operate, leading often to the pathological accumulation of metabolite over a toxic threshold. The comprehension at the molecular level how the protein trafficking process occurs in vivo and how it is impaired in pathological conditions might lead to the identification of new targets for therapeutic treatments.

In this perspective, my PhD project was focused on the investigation of molecular mechanisms involved in traffic processes of several proteins associated to genetic diseases when they are mutated. These investigations have been carried out basically by using functional proteomic approaches. The role that a specific protein plays in intra- and extra-cellular processes is clarified by the identification of its molecular partners. Indeed, the association of an individual protein, whose its function is unknown, with protein complexes involved in well definite cellular processes would be strongly suggestive of its biological function. A classical functional proteomics approach consists of isolation of protein complexes involving the target protein (bait) from a cell lysate by immunoprecipitation. Proteins so purified were then fractionated by SDS-PAGE, digested in situ with trypsin and identified by nano-LC-MS/MS methodologies integrated with protein database search.

In this PhD thesis, the above strategy was employed in the investigation of molecular mechanisms impaired in two rare genetic disorders: Wilson and Pompe disease.

Wilson disease, described in Chapter 2, is a disorder characterised by defective copper excretion, due to mutations in copper transporter ATP7B. In normal conditions, ATP7B binds Cu in the trans-Golgi network and moves to the plasma membrane where delivery the Cu in bile channels. Although the most frequent ATP7B mutants, such as ATP7B(H1069Q), potentially are able to bind Cu, they cannot reach the plasma membrane where the excess of Cu has to be removed. In order to evaluate which molecular pathways result altered by expression of ATP7B mutants, a comparison between the interactomes of the wild type protein and the mutant ATP7B(H1069Q) was performed by functional proteomics approach.

Pompe disease, discussed in Chapter 3, is a rare disorder of glycogen metabolism caused by mutations in gene encoding GAA, essential enzyme for the degradation of glycogen to glucose in lysosomes. Although the biochemistry GAA activity and genetic basis of the associated disorder are well characterized, the routes followed by wild type GAA to reach lysosomes and impaired in presence of mutants are still unclear. This work was focused on the investigation of the intracellular pathways controlling GAA traffic by the identification of protein partners associated to the enzyme on the route from ER to lysosomes. Similarly, the uptake pathway of recombinant GAA (rhGAA), used in Enzyme Replacement Therapy (ERT), was also studied in order to define the fate of recombinant protein once got into the cell.

In Chapter 4, the investigation of molecular mechanisms impaired in amyloidosis has been described. Amyloidosis diseases are associated to the formation of protein aggregates at level of different organs, often associated to the presence of genetic variants of secreted proteins. The study concerning amyloidosis has been carried out by using a double approach: the conformational characterization of some amyloidogenic ApoAI variants was investigated in vitro by a strategy based on complementary proteolysis coupled with mass spectrometry (LC-MS). This methodology allowed clarifying the effect of single point mutations on the protein folding and stability thus to make hypothesis on the structural basis of amyloidosis as protein misfolding/mistrafficking diseases. Moreover a new proteomic strategy for typing amyloid deposits was developed during the period spent abroad at Wolfson Drug Discovery Unit, Centre for Amyloidosis and Acute Phase Proteins, Division of Medicine, University College London (Royal Free Campus), London, UK. Diagnosis and treatment of systemic amyloidosis depends on the correct identification of the amyloid protein triggering the formation of aggregates. In fact, a precise identification of the amyloid fibril protein is essential for the definition of an appropriate therapy. In this field, the

proteomic analysis of amyloid deposits provides a chemical characterization of fibrillar constituents, which add important details to genetic sequencing and immunohistochemistry analysis. However, there are ambiguous cases, in which more than one potentially amyloidogenic protein is found within the patient's biopsy and an accurate diagnosis becomes challenging. The work was aimed at the development of a methodology based on decellularization of human fat biopsies, tryptic digestion and mass spectrometry analysis (LC-MS/MS) in order to eliminate background contamination and improve the specificity of amyloid typing. The use of deoxycholate detergent and the shaking of fat biopsy in tissue lyser resulted fundamental in decellularization procedure. The innovative strategy of decellularization proved to be a simple way to improve the accuracy and specificity of proteomic identification of amyloid fibril type removing most of cellular and plasma proteins background from tissue without altering properties of amyloid fibrils.

Chapter 1 – Introduction

1.1 Transport Machinery

Eukaryotic cells differ from prokaryotic cells by their more complex intracellular organization. In eukaryotes, specific cellular functions are compartmentalized into organelles surrounded by intracellular membranes. This compartmentalization is essential because improves the efficiency of many cellular functions and prevents potentially dangerous molecules from roaming freely within the cell [1].

In eukaryotes, an amount of signals, sorting constituents and energy are required in order to translocate efficiently the proteins into the endoplasmic reticulum (ER), the nucleus, mitochondria, peroxisomes, or to sort precisely vesicles between the membranes of the ER, Golgi, lysosomes, endosomes, autophagosomes, and plasma membrane (PM) (Figure 1.1).

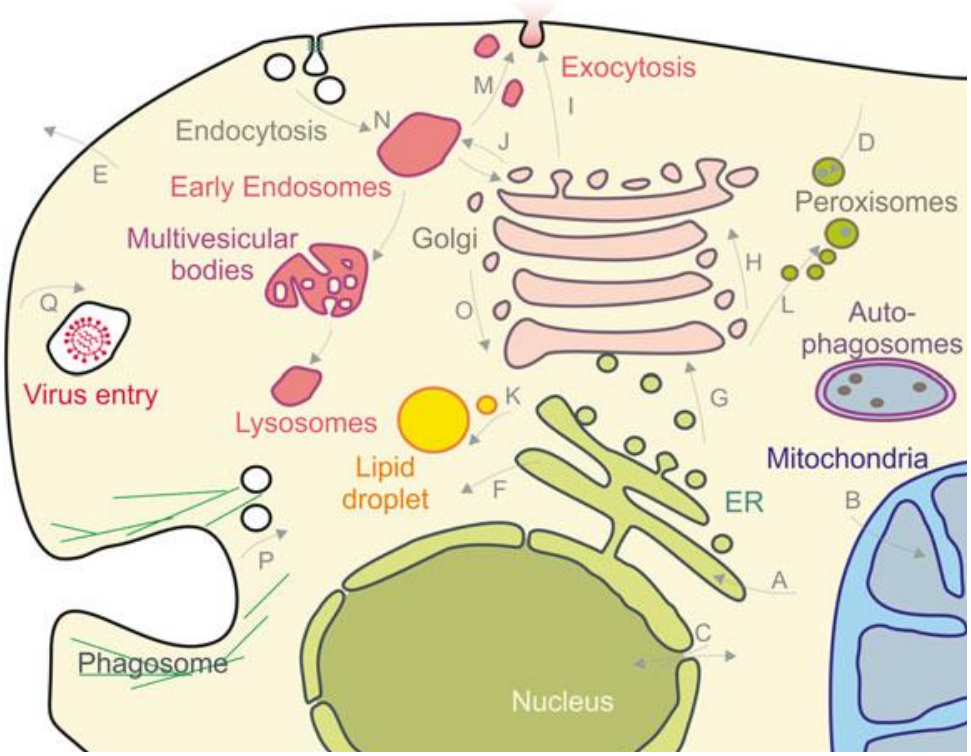


Figure 1.1: Schematic representation of intracellular protein trafficking. The distribution of proteins in eukaryotic cells relies on the translocation of proteins across membranes (A–E) and the dynamics of membrane structures, particularly the coordinated exchange of vesicles (F–P).

Along the secretory pathway, proteins are transported from the ER via the Golgi to lysosomes, endosomes, plasma membrane or to the extracellular space (Figure 1.1: G-M). On the other hand, the endocytic route mediates the uptake of peptides, proteins, fluids, and nutrients from the extracellular space (Figure 1.1: N-O). Moreover, the retrograde transport is critical for the recycling of components of the transport machinery [2].

Most of transport steps occur via *vesicles*, which work as specialized transport carriers that shield their content form the cytoplasm. In general, intracellular vesicle transport occurs through budding from a membrane, docking and fusing with a target membrane of another organelle releasing its contents (Figure 1.2).

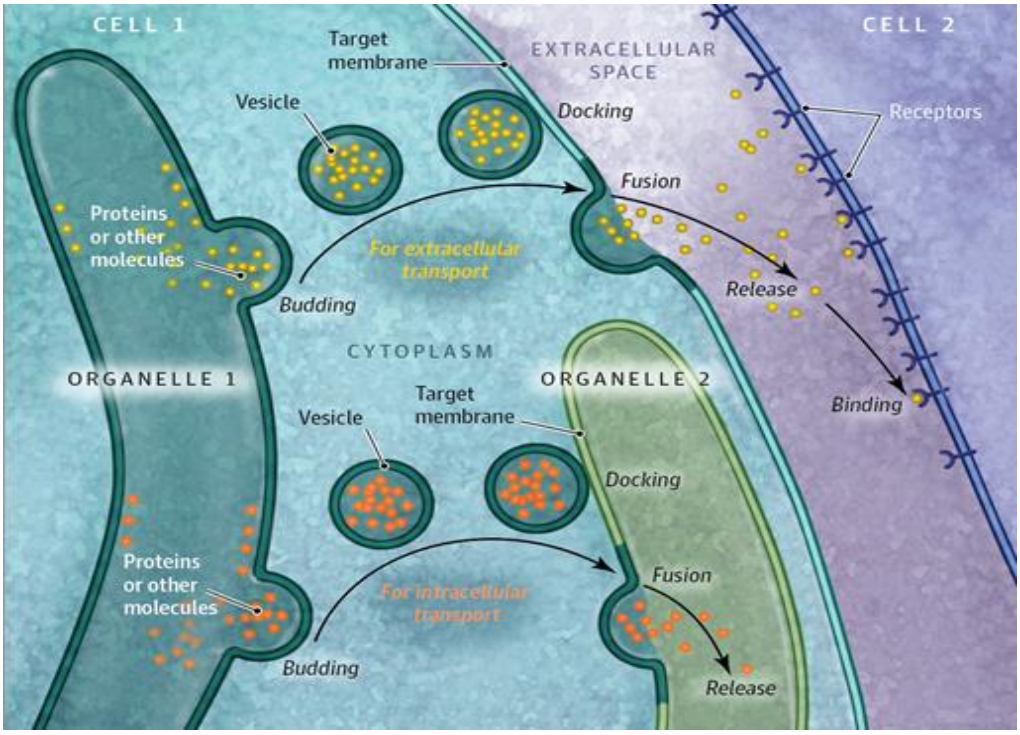


Figure 1.2: Schematic representation about vesicles transport and fusion in cell.

Following their release from the donor membrane, vesicles can interact with motor proteins, which hook them to the microtubules. Their transport on tracks of the cytoskeleton is critical for long-range transport such as from the ER to the Golgi and subsequently to the plasma membrane in mammalian cells like fibroblasts, and is further extremely critical along axons and dendrites of neurons during the release of neurotransmitters.

In addition to vesicular transport system, in eukaryotic cell several other transport system are active: for instance, the unfolded protein translocation occurs through a process that often depends on the *chaperones* activity. These proteins are involved in a plethora of processes including the folding of newly synthesized proteins, the refolding of misfolded and aggregated proteins, the transport of proteins across membranes, the transport of substrates to the proteasome, in particular of proteins that are misfolded in the ER and are retro-translocated to the cytosol for degradation [3].

In general, the specificity in the delivery of molecular cargo is essential for cell function and survival. It is required for several cellular processes, such as the release of neurotransmitters into the presynaptic region of a nerve cell to transmit a signal to a neighbouring nerve cell or for the

hormones export such as insulin to the cell surface. In order to ensure specific delivery of molecular cargo within and/or outside the cells, the molecular mechanisms must be strongly controlled, also in *the time* and *the space:* proteins or other molecules must be delivered to the specific place at the right time. [4]. How does the transport of proteins and other biomolecules occur with crucial specificity?

1.1.1 Traffic Control System Within/Outside the Cells

The 2013 Nobel Prize in Physiology or Medicine honoured three scientists: Dr. Rothman, Dr. Schekman and Dr. Südhof, who have solved the mystery of how the cell organizes its complex transport system. They demonstrated that mechanisms employed by a cell to accomplish vesicle transport and fusion are independent by cell type, and they are the same in human as well as in yeast organisms. Dr. Schekman used yeast genetics to identify a set of genes critical for vesicular trafficking, showing that these genes are essential for life and regulate different aspects of vesicle transport. Dr. Rothman identified proteins on the vesicle and target membranes that form a functional complex controlling cell fusion at specific destination.

Dr. Südhof was interested how vesicle fusion machinery was controlled. He unravelled the mechanism by which calcium ions trigger release of neurotransmitters, and identified key regulatory components in the vesicle fusion machinery (Figure 1.3).

The discovery of three scientists about the machinery regulating vesicle traffic, *the major transport system in cells*, represented a paradigm shift in our understanding of how the eukaryotic cell, through its complex internal compartmentalization, organizes the routing of molecules and proteins packaged in vesicles to various intracellular destinations, as well as to the outside of the cell [4][5].

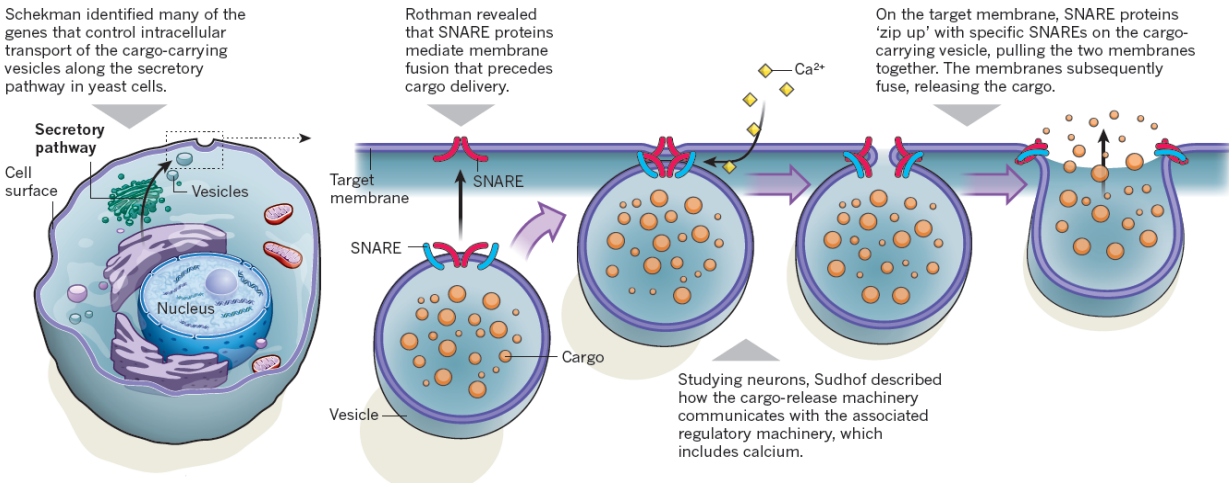


Figure 1.3: The discovery of three scientists: identification of the trafficking machinery regulating vesicle traffic, the major transport system in cells.

1.1.2 Protein Trafficking Defects in Human Disease

The protein traffic plays a critical role for a variety of physiological processes in which vesicle transport and fusion must be strictly controlled. In absence of this precise and well working transport system, every cell drops in chaotic condition leading often to the accumulation of metabolite over a toxic threshold and consequently leading to the development of several disorders. Many of these diseases occur in presence of genetic mutations that give rise to defective proteins whose functions in the traffic and / or transport are partially or completely impaired (Figure 1.4).

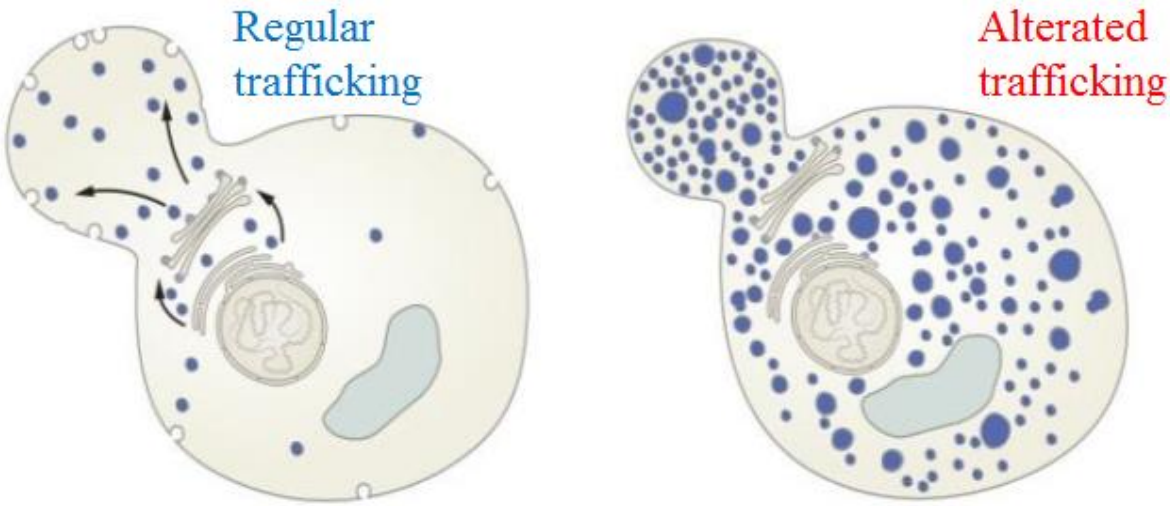


Figure 1.4: Schematic representation of normal and impaired transport system in cell.

Genetic mutations may often cause protein misfolding with consequently retention in the endoplasmic reticulum, followed by degradation; furthermore, aberrant folding behaviour or other alterations, such as misfolded proteins or proteins with defective post translational modification, like an abnormal glycosylation, might result in the failure of a protein to be correctly trafficked.

In presence of misfolded proteins, endoplasmic reticulum employs two distinct mechanisms responding to this phenomenon: the first is an ER-dedicated stress response termed the unfolded protein response (UPR), which acts to remodel the ER so as to increase its folding capacity [6]. The second, termed ER-associated degradation (ERAD), specifically recognizes terminally misfolded proteins and retrotranslocates them across the ER membrane into the cytosol, where they can be degraded by the ubiquitin-proteasome degradation machinery [7]. These two systems are intimately linked: UPR induction increases ERAD capacity, loss of ERAD leads to constitutive UPR induction, and simultaneous loss of ERAD and the UPR greatly decreases cell viability. Ubiquitin-proteasome degradation machinery is normally sufficient to prevent the accumulation of misfolded proteins, but under certain pathological conditions, the system of intracellular quality control appears to be incapable of recognizing and removing dangerous misfolded proteins. Thus,

proteins aggregate and/or interact inappropriately with other cellular components leading to impairment of cell viability and eventually to cell death [8].

Many disorders are known as misfolding or conformational diseases, in which peptides or proteins fail to fold and then aggregate giving arise to intracellular aggregates or extracellular amyloid deposits. Misfolded proteins which are not degradated by proteosome can be transported to extracellular space. This pathological protein traffic leads to the accumulation of toxic amyloidosis fibrils outside the cell.

Misfolding diseases include a range of sporadic, familial or transmissible degenerative diseases, some of which affect the brain and the central nervous system, such as Alzheimer's disease. Others pathologies, such as systemic amyloidoses and type II diabetes, involve peripheral tissues and organs like liver, heart and spleen, etc. [9][10].

Further examples of disorders caused by protein misfolding and/or mistrafficking are: *Wilson disease,* an inherited disorder of copper metabolism, associated to the presence of ATP7B pathological variants that lead to an accumulation of intracellular copper [11]; *Pompe disease,* a metabolic disorder caused by mutations in genes encoding GAA hydrolase, involved in degradation of glycogen in lysosomes; the malfunction of GAA causes glycogen accumulation in many tissues and organs [12]. *Cystic fibrosis*, a disease caused by mutations in CF transmembrane conductance regulator (CFTR). The most frequent mutation (F508del-CFTR) destabilises native protein conformation thus to trigger intracellular degradation of the protein. Therefore, mutated protein is not able to reach plasma membrane with a consequent deregulation of chloride homeostasis, and tissue damages, mostly at level of lungs, which become very vulnerable to bacterial infections which at long term, may eventually lead to death [13].

The comprehension at the molecular level how the protein trafficking process occurs in vivo and how it is impaired in pathological conditions might lead to the identification of new targets for therapeutic treatments. An accurate definition of the protein network involved in traffic processes can be pursued by functional proteomic approaches.

1.2 Proteomics

The term "*proteome*" was first coined at the end of the last century to denote the set of all proteins expressed by a genome of an organism in a precise biological context and time. Proteomics is the science focused on the systematic analysis of all the proteins of a proteome, from total cell extracts, tissue extracts or whole organisms, which are analysed at the same time and within their biological context.

Since it is now known the complete sequence of the genome of various organisms, including the human one, the number of proteins whose function is unknown has grown exponentially in recent times. This gave rise to the need to consider the knowledge of the human genome as a starting point, and not the end, for the understanding of the cellular processes [14].

The complexity of the proteome derives not only from the high number of components that constitute it, but also by its extreme variability, as it may differ from individual to individual and for a certain organism may change over time, in response to changes of the external environment or in signals arriving to the cell. Another characteristic of the proteome is its wide dynamic range of concentration: therefore few protein species are present in a large amount while a higher number of species in a very small quantity.

The investigation in proteomics field can essentially be divided into two main areas: the expression and functional proteomics. The first approach is focused on the qualitative and quantitative evaluation of the increase and / or decrease in protein expression levels following the change of experimental conditions, such as cellular stress, the presence of diseases, drugs treatments, etc. [15]. Functional proteomics aims to define the biological function of proteins whose role is not known yet by identifying protein-protein interactions in vivo, for the description at the molecular level of cellular mechanisms [16].

1.2.1 Functional Proteomics

The ability of any organism to survive depends on several complex cellular mechanisms that fulfill all of the functions needed for cell life. In addition, the cells are able to change the patterns of gene expression in response to signals extra- and intra-cellular in order to ensure that the correct amount of the appropriate subset of genes are expressed at the right time [17][18].

In the past, cellular mechanisms and signal transduction were considered to occur as a set of isolated processes where biological functions took place one after the other in a time-dependent fashion. The interactions involved in biological processes were first characterized individually, but

this approach suffered from a lack of information about time, space, and context in which the interactions occur in vivo.

Nowadays, it is well known that the biological mechanisms are pursued by a wide number of proteins gathering together to transiently form large functional complexes [19] or cell pathways [20]. Proteins inside the cell do not interact randomly, but protein associations need precise regulation. Protein complexes are involved in a continuous dynamic of association-dissociation events, in which two variables are very important: *the time* in which the proteins are associated to preside to a specific process, and *the space*, the place in the cell where this association occurs. Once their functions carried out, these complexes dissociate to form individual components free to associate with other complexes to perform different activities. In addition, several mechanisms and many signaling pathways are able to interact with each other thus to form an intricate network that integrates simultaneously extracellular and intracellular signals. These dynamic processes are the molecular networks through which signaling information flows within the cell [21][22].

A complete description of the intricate signal transcription and transduction pathway and the use of network to predict the full range of cellular behaviors are the main goals of systems biology [23][24]. A crucial role in the biological sciences contemporary and future can be played by functional proteomics, an area in proteomics research addressing the elucidation of protein functions and the definition cellular mechanisms at the molecular level [25][26].

A comprehensive description of cellular processes at the molecular level is strictly dependent on a clear definition of the individual protein components involved in these functional entities. The association of a protein, whose function is unknown, with partners belonging to a specific complex of proteins involved in particular cellular mechanisms may be strongly indicative of its biological function [27][28]. In addition, a detailed description of cellular signaling pathways might greatly benefit from the elucidation of protein-protein interactions in vivo.

1.2.2 Protein-Protein Interaction in Human Disease

“Life is a relationship between molecules, not a property of any one molecule. So is therefore disease which endangers the life” wrote Zuckerkandl and Pauling (1962) in their chapter on Molecular disease, evolution and genic heterogeneity. After more than 50 years we are still quite far from understanding the molecular mechanisms of most disease and the molecular interactions that are established in healthy and diseased organisms [29]. Indeed, many proteins carry out their biological function through molecular interactions, which are involved in supramolecular assemblies (collagens, elastic fibers, actin filaments), in the building of molecular machines (molecular motors, ribosomes, proteasome) and in main biological processes such as immunity

(antigen-antibody interaction), metabolism (enzyme-substrate interaction), signaling (interaction of messenger molecules, hormones, neurotransmitters with their cognate receptors), and gene expression (DNA-protein interactions). The investigation of protein-protein interaction maps (interactomes) is crucial to shed light on each protein networks role within the cell and is essential to reveal the molecular mechanism of processes that lead to disease. Genetic mutations or environmental factors interfere with protein-protein interaction leading to pathology. The kinetics of the processes and dynamics of the networks must be considered in order to understand how the perturbation of proteins and their interactions lead to the disease. In addition, the specific biological context (tissue, stage of disease and response) is also important to consider where the protein interaction occurs. At present there is still a gap between the identification of protein networks associated with the disease and the complete understanding of the mechanism of the disease. We are still far from understanding the etiology of most diseases, but further significant advances in technology and experimental field, such as those related to protein interactions, protein structure, gene expression, together with computational tools to organize and integrate these data will provide a step forward that direction. Since structural studies and the networks investigation are complementary approaches, their combination has the potential to be an excellent framework for the study of mechanisms impaired in diseases.

1.2.3 Proteomics Approach for the Study of Protein-Protein Interactions

For proteomics purposes, entire multiprotein complexes must be isolated and this can be accomplished by affinity based-approach or immunoprecipitation (IP) strategies. The latter consists in isolation of the protein of interest, also called *bait*, and its native protein complexes from a cellular/tissutal extract containing many thousands of different proteins. This is possible using affinity between the bait and an antibody, coupled on a solid substrate (beads), that recognizes its antigen. IP procedures essentially fall into two categories: direct immunoprecipitation of the endogenous protein bait from cellular extracts or in situ production of a tagged form of the bait followed by immunoprecipitation with antitag antibodies. This second category is used when there are no antibodies suitable for the protein of interest or when the endogenous bait is expressed at too low level. Therefore the protein is trasfected into appropriate cell line, in order to be expressed with a small peptide epitope (e.g., FLAG, V5, c-myc), located at either N- or C- terminus, thus to not effect protein tertiary structure and biological activity. After washing in order to remove unspecific proteins, bait and its partners are eluted in denaturing conditions or by competition with a relevant concentration of the free form of the tag peptide. Finally, once carried out the immunoprecipitation, protein components are fractionated by sodium dodecyl sulfate polyacrilamide gel electrophoresis SDS-PAGE. Protein bands are in situ digested

with a specific protease such as trypsin, and the resulting peptide mixtures analyzed by tandem mass spectrometry combined with reverse-phase capillary liquid chromatography (nano-LC-MS/MS), leading to sequence information required for the identification of proteins by database search (Figure 1.5).

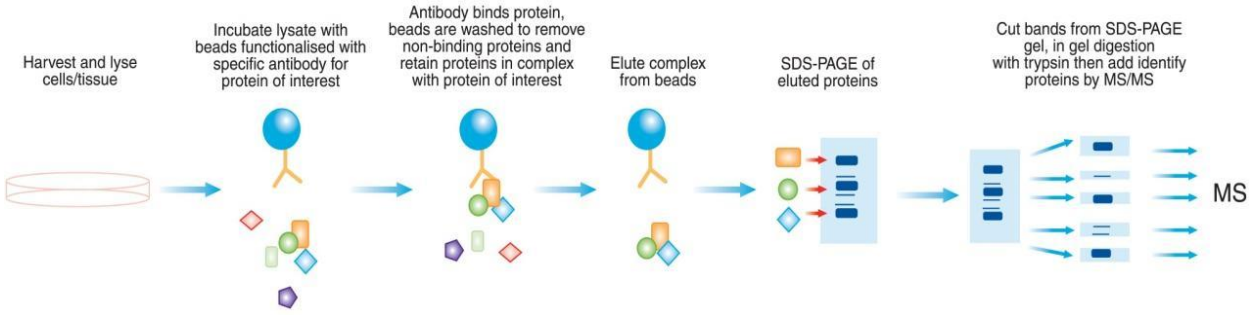


Figure 1.5: Description of IP strategy and identification of bait interacting proteins

These approaches usually results in the final identification of a very long list of interactors from a single functional proteomics experiment. Since a single protein might be involved in different complexes, each endowed with its specific biological function, the attempt to rationalize the list of interactors within a biologically significant picture constitutes a major challenge in a functional proteomics. Two approaches are normally used to manage these lists: the simplest strategy involves the choice of one or few interactors among the various protein candidates on the basis of previous data or literature information. The interaction of these proteins with the bait has to be confirmed by alternative approaches (such as co-immunoprecipitation and western blot assay) and functionally investigated. A second approach aims to assign each protein to a specific functional complex by using functional data provided by literature or protein-protein interaction database, keeping in mind that each protein might simultaneously belong to different entities since might play diverse biological functions. The final purpose of this second approach would be the achievement of a global functional overview of the protein bait interactome to be inserted into appropriate systems biology studies [16].

1.2.4 Proteomics in Clinical Diagnosis

Several metabolic diseases and genetic abnormalities are known to impair the cellular transcription and translation status and thereby creating a variation in proteome [30]. These alterations in proteome are an opportunity for development of diagnosis of molecular signatures for several diseases and lead to the development of a new area called *clinical proteomics* [31]. In this regard, for example, proteomics analysis is a powerful tool both for the diagnosis and for the investigation of molecular mechanisms impaired in amyloidoses, a large class of pathologies associated to protein misfolding and aggregation. Proteomics can be useful to elucidate different aspects of these

diseases, such as the characterization of amyloidogenic protein precursors in plasma or other compartments (with particular attention to post translational modifications); the identification of the nature of the fibrillar deposited proteins for disease typing; the investigation of the metabolic modifications in amyloid tissues, which lead to cellular toxicity and organ dysfunction; and the discovery of new biomarkers useful for the characterization of disease severity and response to treatment. In particular, amyloid typing is a key step in the management of amyloidoses, since the clinical course, treatment, and prognosis are critically dependent on the amyloid type (assigned on the basis of which protein forms the fibrils) [32].

1.3 Mass Spectrometry

Mass spectrometry (MS) is a powerful analytical methodology used to identify unknown compounds, to quantify known compounds, and to elucidate the structure and chemical properties of molecules, by generating ions in gas phase and measuring their mass to charge (m/z) ratio, also after fragmentation events.

Every mass spectrometer is composed of three main components: an ionization source, a mass analyzer and a detector (Figure 1.6). The ion source converts the sample into ions in gas phase, which then move to mass analyzer through an extraction and focusing system of electrostatic lens. The mass analyzer resolves the ions by their mass-to-charge ratio (m/z). Finally, the ions sorted by mass analyzer move to detector which provides abundances of each ion detected by measuring the associated current.

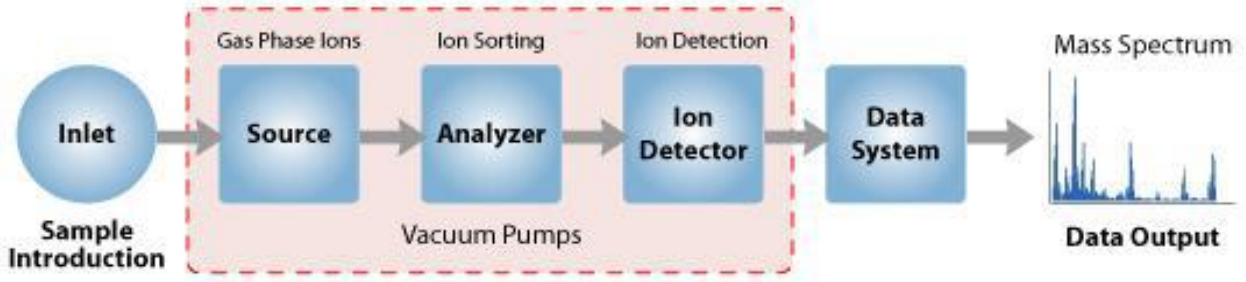


Figure 1.6: Schematic representation of a mass spectrometer.

Mass spectrometry was restricted for a long time to small, volatile and thermostable compounds untill the development, in the late 1980s, of two techniques, the electrospray ionization (ESI) and matrix assisted laser desorption/ionization (MALDI), able to generate molecular ions of intact biomolecules. These technological innovations dramatically changed the research scenarios made peptides and proteins accessible to mass spectrometric analysis. Starting from those years, the development of new mass analyzers and complex multistage instruments that would allow proteins and proteome analysis has been catalyzed [33].

In *MALDI* ion source, the sample intended for analysis is incorporated into a solid matrix containing compounds, such as 3,5-dimethoxy-4- hydroxycinnamic acid (sinapinic acid), which absorbs energy at the wavelength of the laser (typically 337 nm). Laser activation of the target by UV laser within the ion source leads to the release from the target of peptide/ protein $[M+H]^+$ ions into gas phase.

In contrast to the MALDI, *electrospray ionization (ESI)* involves the generation of peptide ions from aqueous solution. The solution containing the sample passed through a needle subjected to a high voltage. The solution stream is ejected from the needle orifice as a spray of droplets. The

solvent is eliminated from the droplets by the effect of a heated capillary and/or a desolvatation inert gas. Solutions with acidic pH favor protonation of the N-terminus and all basic aminoacid side chains, generating multiply-charged molecular-related ions, such as $(M+nH)^{n+}$. The number of charges depends on the chemical-physical characteristics of analytes; thus each analytes could assume more than one charge complicating the form assumed by the signal. For this reason this source is not useful to analyse complex mixtures except using pre-fractionation systems. In fact most uses of ESI-MS are made in combination with chromatographic separation, the so called LC-MS technique. A great enhancement of ESI ion sources has come from the reduction of the flow rate of the liquid used to create the spray to a nano-scale level. The advantages related to the use on μLC and nLC are the low consumption of sample and the higher sensitivity of the methods because of the increase in the concentration of the analyte as it elutes off the column.

Once generated gas phase ions, they move to the mass analyzers, which measure the mass-to-charge (m/z) ratio. Several types of mass analyzers have been developed including quadrupoles (Q), ion traps (IT), time-of-flight (TOF) and Fourier transform ion cyclotron resonance (FTICR), Orbitrap or combination of these in hybrid instruments, that are commonly use in proteomics (Figure 1.7); they differ for their resolution, sensitivity and accuracy and for the types of experiments that are able to perform.

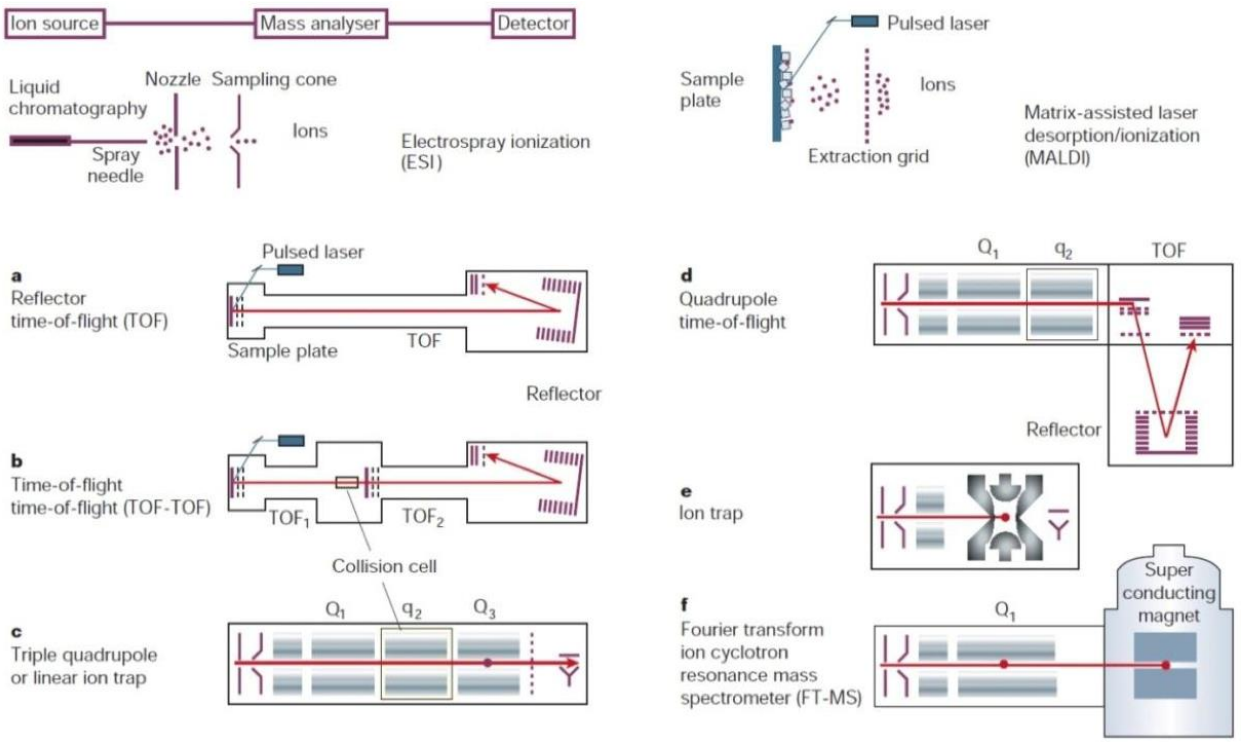


Figure 1.7: ESI and MALDI ionization source (upper row); types of mass analysers (lower row: a-f).

The instrumentations employed to carry out my PhD research projects were equipped with nano ESI sources and Ion Trap and/or Orbitrap as analysers.

Ion trap (IT) mass analysers are able to trap ions in a three-dimensional electrical field. The ion trap consists of a ring electrode and end cap electrodes, capable of analysing, isolating and retaining specific ions also for fragmentation upon collision with an inert gas in the same device. One end cap electrode usually has a hole in it to allow ions to enter the trap. Ions are injected axially and to overcome the exit of the ions, their high kinetic energy is dissipated by cooling them collisionally, filling the trap with gas molecules. Ions are dynamically stored in a three-dimensional quadrupole ion storage device by an electric quadrupole field generated by the applied RF voltage. The RF voltage can be scanned to eject successive mass-to-charge ratios from the trap to the detector. The discovery of the mass selective instability mode transformed the ion trap from a storage device into a powerful instrument capable of sequential ejection of selected or all ions out of the trapping cell to a detector.

Orbitrap mass analyzer was invented by Alexander Makarov. It consists of a spindle-shape central electrode surrounded by a pair of bell-shaped outer electrodes. The Orbitrap employs electric fields to capture and confine ions. Stable ion trajectories rotate around an axial central electrode with harmonic oscillations along it (Figure 1.8). The frequency of these harmonic oscillations along the z-axis depends only on the ion mass-to charge ratio m/z. Two split halves of the outer electrode of the orbitrap detect the image current produced by the oscillating ions. By Fast Fourier Transformation (FFT) of the image current, the instrument obtains the frequencies of these axial oscillations and therefore the m/z of the ions.

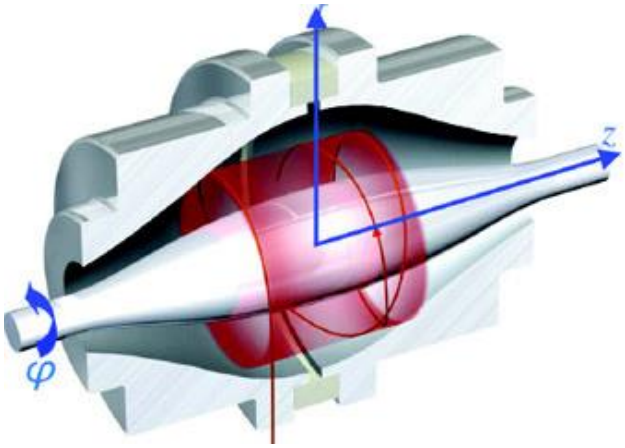


Figure 1.8: Orbitrap mass analyser

1.3.1 Tandem Mass Spectrometry

The power of mass spectrometry can be significantly increased by employing methods of tandem mass spectrometry (MS/MS) from which we obtain structural information about a compound by fragmenting specific sample ions inside the mass spectrometer and identifying the fragment ions. Instruments for tandem mass spectrometry are classified as *tandem in space* or *tandem in time.*

Tandem in space means that ion selection, ion fragmentation and fragments analysis, are events that occur in different regions of the spectrometer, while in tandem in time instruments, the three steps of analysis occur in the same region of the spectrometer, but in different times.

Tandem mass spectrometry uses two stages of mass analysis, the first one to pre-select a precursor or *parent* ion and the second to analyse fragments or *daughter* ions. These are separated by a collision cell into which an inert gas (e.g. nitrogen, argon, and helium) is admitted to collide with the selected ions causing the fragmentation. This process is known as *Collision Induced Dissociation* (CID). In elastic collisions between the precursor ion with a high translational energy and a neutral target gas cause the conversion of part of the translational energy into internal energy of the ion, leading to subsequent decomposition.

The collisions can break several bonds along the backbone (NH-CαH, CαH-CO and CO-NH) originating different fragment ions. Each bond breakage gives rise to two species, one neutral and the other one charged, and only the charged species are monitored by the mass spectrometer. Hence there are six possible fragment ions for each amino acid residue and these are labelled with the *a*, *b*, and *c* ions having the charge retained on the N-terminal fragment, and the *x*, *y* and *z* ions having the charge on the C-terminal fragment (Figure 1.9). The most probable fragmentation event occur at level of peptide CO-NH bonds, giving rise to the *b* and the *y* ion series, characterised by the charge retention on the N or C terminus, respectively.

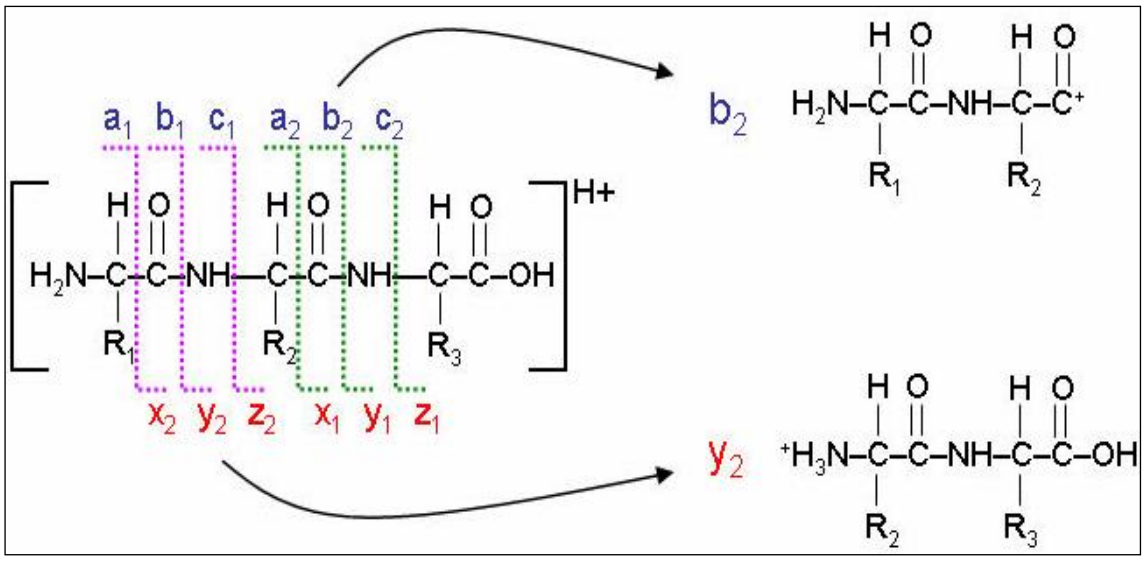


Figure 1.9: Peptides fragmentation scheme.

1.4 Protein Identification

Data obtained by MS/MS spectra provide information about amino acid sequence of peptides, and are crucial for univocal protein identification.

Data provided from an LC-MS/MS analysis, i.e. molecular weight and fragmentation spectra of the main abundant ions occurring in the starting mixture, are included in a peak list, entered into Mascot software, using the option MS/MS Ion Search and are employed to consult protein databases. This software, indeed, by miming the hydrolysis experiment, calculates the expected molecular weight of all peptides theoretically generated by each protein present in the database, according to protease specificity and possible chemical modifications occurring. An algorithm compares the experimental mass values obtained as m/z by MS and MS/MS analyses, with the theoretical mass values generated by Mascot on the basis of protein sequences, each match is associated to a probability score of identification. The score s is defined as: $s= 10*\log(P)$, where P is the probability that the coincidence of observed mass values with the theoretical ones is a random event; then, lower P value correspond to a higher possibility of identification. Only the identifications with a score higher of a score threshold chosen by the software, according to the type of mass spectrometer used, its accuracy and resolution and other experimental conditions, such as database consulted for the identification, are considerate significant.

1.5 Aim of the PhD Thesis

The protein traffic and the molecular mechanisms involved in the transport of each individual protein towards a specific destination are of crucial importance for normal cellular physiology. The knowledge of how the protein trafficking processes occur in vivo and how it is impaired in pathological conditions might lead to the identification of new therapeutic targets.

In this perspective, my PhD project is focused on the investigation of molecular mechanisms involved in traffic processes both in physiological and in psthological conditions by using proteomic approaches. The role that a specific protein plays in intra- and extra-cellular processes is clarified by the identification of its molecular partners.

In addition to proteomics approach, a study of protein conformational stability of amyloidogenic protein variants by employing complementary proteolysis approach coupled with mass spectrometry (LC-MS), has been performed to clarify the effect of single point mutations in the development of protein misfolding/mistrafficking diseases.

The description of the work has been structured as follows:

- ✓ **Chapter 2**: The investigation of molecular mechanisms impaired in Wilson disease.

The work was addressed to the study of molecular pathways impaired in Wilson disease comparing wild type ATP7B and the mutant ATP7B interactomes. The differences occurred between the two set of ATP7B interacting proteins was allowed the identification of new molecular targets for a possible therapeutic treatment of Wilson.

- ✓ **Chapter 3**: The investigation of molecular mechanisms impaired in Pompe disease.

This study was focused on the investigation of the intracellular pathways controlling GAA protein traffic in physiological condition and on the internalization process of rhGAA protein in order to elucidate the fate of recombinant protein once got into the cell by proteomics approaches. The identification of intracellular factors regulating the traffic of GAA protein, involved in Pompe disease, might unravel the cell processes impaired in presence of GAA mutant causative of this disorder. Moreover, the description of uptake pathway followed by recombinant GAA (rhGAA), used in enzyme replacement therapy (ERT) of Pompe disease, has been useful to shed light on why only low amount of protein correctly reaching the right destination leading inefficacy ERT in numerous cases.

✓ **Chapter 4**: The investigation of molecular mechanisms impaired in Amyloidosis.

This study was addressed to add knowledge to the molecular mechanisms by which amyloidogenic mutations impair Apolipoprotein AI (AApoAI) protein stability and native conformation.

The last part of research project was focused on the development of new proteomic strategy based on decellularization of human fat biopsies coupled with mass spectrometry analysis (LC-MS/MS) for a correct typing of amyloid deposits extracted from adipose tissue of patients. Proper amyloid typing represents a crucial step in the diagnosis and then in the choice of a specific and effectiveness treatment of systemic amyloidosis.

1.6 References

[1] Rothman JE, Sheckman RW, Südhof TC, Machinery regulating vesicle traffic, a major transport system in our cells. *Nature Vol. 502: 149–150, 2013.*

[2] Tang BL, Membrane trafficking, *Methods Molecular Biology 1270, 2015.*

[3] Saibil H, Chaperone machines for protein folding, unfolding and disaggregation. *Nat Nature Reviews Molecular Cell Biology, Vol. 14(10): 630-42, 2013.*

[4] Friedrich MJ, 2013 Nobel Prize recognizes work of scientists who illuminated molecular transport system of cells. *JAMA Vol. 310(19): 2027-9, 2013.*

[5] Bonifacino JS, Vesicular Transport Earns a Nobel. *Trends in Cell Biology Vol. 24(1): 3-5, 2014.*

[6] Schroder M, Kaufman, RJ The mammalian unfolded protein response. *Annual Review of Biochemistry Vol. 74: 739–789, 2005.*

[7] Romisch K. Endoplasmic reticulum-associated degradation. *Annual Review of Cell and Developmental Biology Vol. 21: 435–456, 2005.*

[8] Stefani M, Dobson CM, Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *Journal of Molecular Medicine Vol. 81(11): 678-99, 2003.*

[9] Kelly J, Alternative conformation of amyloidogenic proteins and their multi-step assembly pathways. *Current Opinion in Structural Biology Vol. 8: 101–6, 1998.*

[10] Dobson CM, The structural basis of protein folding and its links with human disease. *Philosophical Transactions of the Royal Society Lond B Vol. 356: 133–45, 2001.*

[11] Gitlin JD, Wilson Disease. *Gastroenterology, Vol. 125(6): 1868-77, 2003.*

[12] Fiore Manganelli and Lucia Ruggiero, Clinical features of Pompe disease. *Acta Myologica, Vol. XXXII: 82-4, 2013.*

[13] Hegde RN, Parashuraman S, Iorio F, Ciciriello F, Capuani F, Carissimo A, Carrella D, Belcastro V, Subramanian A, Bounti L, Persico M, Carlile G, Galietta L, Thomas DY, Di Bernardo D, Luini A. Unravelling druggable signalling networks that control F508del-CFTR proteostasis. *eLIFE 4: e10365, 2015.*

[14] O'Donovan C, Apweiler R, Bairoch A, The human proteomics initiative (HPI). *Trends in Biotechnology, Vol. 19 (5): 178-81, 2001.*

[15] Corthals GL, Wasinger VC, Hochstrasser DF, Sanchez JC, The dynamic range of protein expression: a challenge for proteomic. *Electrophoresis, Vol. 21 (6): 1104-15, 2000.*

[16] Monti M, Cozzolino M, Cozzolino F, Vitiello G, Tedesco R, Flagiello A, Pucci P, Puzzle of proteomic complexes in vivo: a present and future challenge for functional proteomics. *Expert Review of Proteomics, Vol. 6 (2): 159-69, 2009.*

[17] Kobayashi Y, Weigel D, Move on up, it's time for change-mobile signals controlling photoperiod-dependent flowering. *Genes & Development, Vol 21: 2371-84, 2007.*

[18] Qiu Z, Ghosh A, A calcium-dependent switch in a CREST-BRG1 complex regulates activity-dependent gene expression. *Neuron, Vol. 60: 775-87, 2008.*

[19] Michaud GA, Snyder M, Proteomic approaches for the global analysis of proteins. *Biotechniques, Vol. 33: 1308–16, 2002.*

[20] Souchelnytskyi S, Proteomics in studies of signal transduction in epithelial cells. *Journal of Mammary Gland Biology and Neoplasia, Vol. 7: 359–71, 2002.*

[21] Perkins JR, Diboun I, Dessailly BH, Lees JG, Orengo C, Transient protein-protein interactions: structural, functional, and network properties. *Structure, Vol. 18 (10): 1233-43, 2010.*

[22] van Bentem S, Mentzen WI, de la Fuente A, Hirt H, Towards functional phosphoproteomics by mapping differential phosphorylation events in signalling networks. *Proteomics, Vol. 8: 4453-65, 2008.*

[23] Lecaudey V, Cakan-Akdogan G, Norton WH, Gilmour D, Dynamic Fgf signaling couples morphogenesis and migration in the zebrafish lateral line primordium. *Development, Vol. 135, 2695-705, 2008.*

[24] Pfeifer AC, Timmer J, Klingmüller U, Systems biology of JAK/STAT signalling. *Essays in Biochemistry, Vol. 45: 109-20, 2008.*

[25] Zak DE, Aderem A, Systems biology of innate immunity. *Immunological Reviews, Vol. 227: 264-82, 2009.*

[26] Schlieker CD, Van der Veen AG, Damon JR, Spooner E, Ploegh HL, A functional proteomics approach links the ubiquitin-related modifier Urm1 to a tRNA modification pathway. *Proceedings of the National Academy of Sciences, Vol. 105: 18255-60, 2008.*

[27] Souza TA, Chen X, Guo Y, Sava P, Zhang J, Hill JJ, Yaworsky PJ, Qiu Y, Proteomic identification and functional validation of activins and bone morphogenetic protein 11 as candidate novel muscle mass regulators. *Molecular Endocrinoogyl, Vol. 22: 2689-702, 2008.*

[28] Brown KA, Ham AJ, Clark CN, Meller N, Law BK, Chytil A, Cheng N, Pietenpol JA, Moses HL, Identification of novel Smad2 and Smad3 associated proteins in response to TGF-beta1. *Journal of Cellular Biochemistry, Vol. 105 (2): 596-611, 2008.*

[29] Kann MG, Protein interactions and disease: computational approaches to uncover the etiology of diseases. *Briefings in Bioinformatics, Vol. (5): 333-46, 2007.*

[30] Di Girolamo F, Del Chierico F, Caenaro G, Lante I, Muraca M, Putignani L, Human serum proteome analysis: new source of markers in metabolic disorders. *Biomarkers in Medicine, Vol. 6 (6): 759-73, 2012.*

[31] Ahmad Y, Arya A, Gangwar A, Paul S and Bhargava K, *Proteomics in Diagnosis: Past, Present and Future*. *Journal of Proteomics and Genomics, Vol. 1 (1): 103, 2014.*

[32] Brambilla F, Lavatelli F, Merlini G, Mauri P, Clinical proteomics for diagnosis and typing of systemic amyloidoses. *Proteomics Clinical Applications, Vol. 7 (1-2): 136-43, 2013.*

[33] Domon B, Aebersold R, Mass spectrometry and protein analysis. *Science, Vol. 312 (5771): 212-7, 2006.*

Chapter 2 – Investigation of Molecular Mechanisms Impaired in Wilson Disease

2.1 Introduction

2.1.1 The Role of Cu-Transporting ATPases in Copper Homeostasis

Copper is a micronutrient that plays an essential role in human metabolism. It is a cofactor of key metabolic enzymes, which are involved in cellular respiration, neurotransmitter biosynthesis, radical detoxification, iron metabolism, and many other physiological processes [1]. Numerous dietary foods are rich in copper (shellfish, chocolate, nuts, raw kale, seeds, mushrooms and legumes) that is absorbed primarily through the intestinal mucosa and this process is regulated by copper intake. In the serum, newly absorbed copper is transported to the liver bound to two proteins: albumin and transcuprein. Most of the copper is taken up by the liver, central organ of copper homeostasis and is primarily responsible for the export of excess copper out of the body (Figure 2.1). In hepatocytes, most of the copper is incorporated into ceruloplasmin in the trans-Golgi network (TGN), which is then secreted into the blood [2][3].

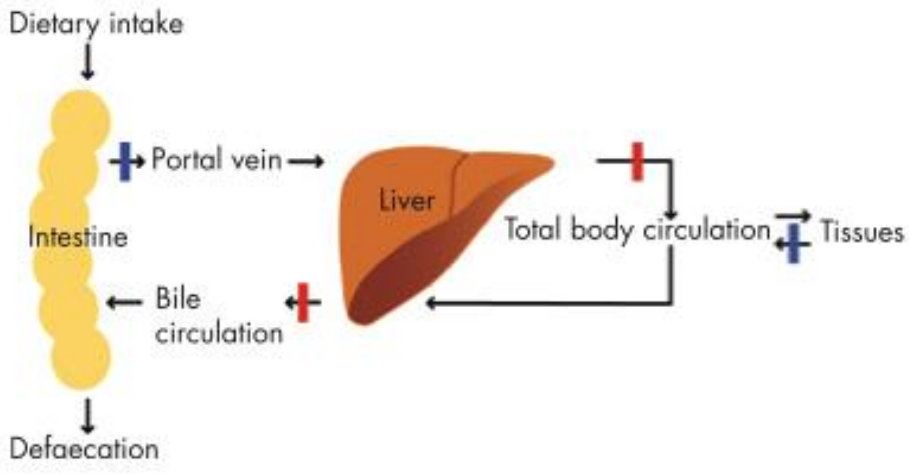


Figure 2.1: Scheme of physiology of copper homeostatis.

In human, copper is in the oxidized Cu(II) form, then it is reduced to Cu(I) by one or more metalloreductases upon entry into the cell and reoxidized when exiting the cell [1][4].

Although copper is essential for human health, in amounts that exceed cellular needs is highly toxic leading to cellular oxidative damage, production of reactive oxygen species (ROS) that can have devastating effects, including DNA and RNA damage and oxidation of proteins and lipids [2][5]. Therefore, the level of copper in the body must be kept under tight control. The organisms have evolved an elaborated network of molecular mechanisms in order to maintain homeostatic balance of copper. These mechanisms involve an intricate balance between uptake, distribution

and utilization, storage and detoxification, and efflux pathways for copper. The breakdown of these mechanisms leads to Cu imbalance that is a crucial factor in the etiology and pathology of several neurodegenerative diseases [2].

Several proteins, called metallochaperones, are involved in intracellular copper uptake and translocation. Copper uptake into cells is mediated by the high-affinity copper transporter Ctr1; once internalized, the copper may be delivered by CCS to Cu/Zn-SOD1, or alternatively transferred by COX17 within the mitochondrial intermembrane space for metallation and assembly of cytochrome C oxidase, or, chaperoned by ATOX1 to the copper-ATPases (Figure 2.2) [2].

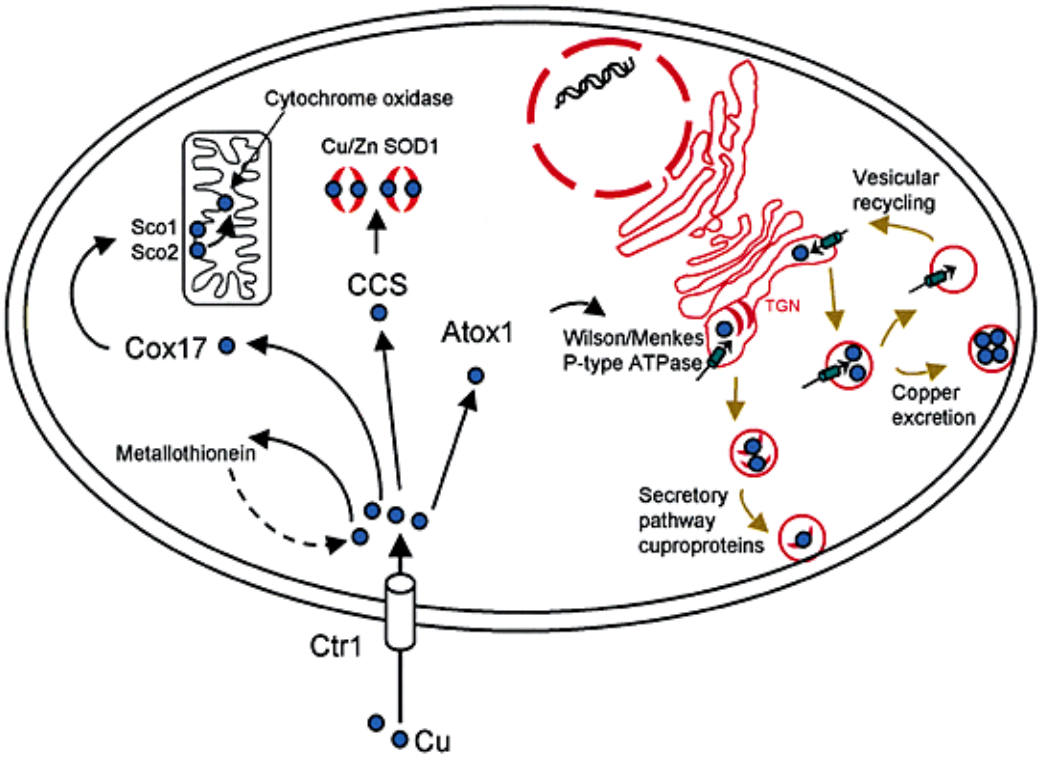


Figure 2.2: Schematic representation of intracellular pathways of copper traffic.

Several physiological processes depend on the proper and timely transport of copper by a family of proteins, known as "*Copper-transporting P-type ATPases*" (Cu-ATPases). *ATP7A* and *ATP7B* enzymes belong to Cu-ATPases family, and have a key role for the regulation and maintenance of homeostatic balance of intracellular copper.

Cu-ATPases are polytopic membrane proteins that translocate copper from the cytosol across cellular membranes using the energy of ATP hydrolysis. This transport process is essential to reduce the intracellular copper concentration and to contribute to homeostatic control of copper in the body. In addition to their homeostatic function, Cu-ATPases perform important physiological functions delivering copper into the secretory pathway where the Cu ion is incorporated into

copper-dependent enzymes, such as dopamine-β-hydroxylase, tyrosinase, lysyl oxidase, peptidylglycine-α- amidating monooxygenase, ceruloplasmin, and others [1][5].

The two Cu-ATPases share 50-60% sequence identity and their expression patterns is somewhat complementary: ATP7A is expressed in the majority of tissues except for the liver, and ATP7B is primarily present in the liver, but also in the kidney and placenta, and at lower levels in brain, heart, lungs, retinal pigment epithelium and in the ciliary body during retinal development. Under normal condition, both ATPases reside at the trans-Golgi network (TGN) into the cell [1][2].

The absence or malfunction of ATP7A and ATP7B proteins leads to disorders of copper metabolism, known as Menkes (MD) and Wilson (WD) diseases, related to a lack or an excess of intracellular copper, respectively [1].

2.1.2 ATP7B Protein

ATP7B gene encodes a large membrane protein of 1465 amino acids (~160 kDa), constituted by eight transmembrane regions and several functional domains (Figure 2.3). Within the N-terminal tail, six Metal Binding Domains (MBDs) are present, each containing the core sequence MxCxxC. Essential motifs for ATP catalysis include the nucleotide binding domain (N-domain), the phosphorylation domain (P-domain) and the actuator domain (A-domain). C-terminal cytoplasmatic tail contains structural motifs involved in endocytosis [1][3].

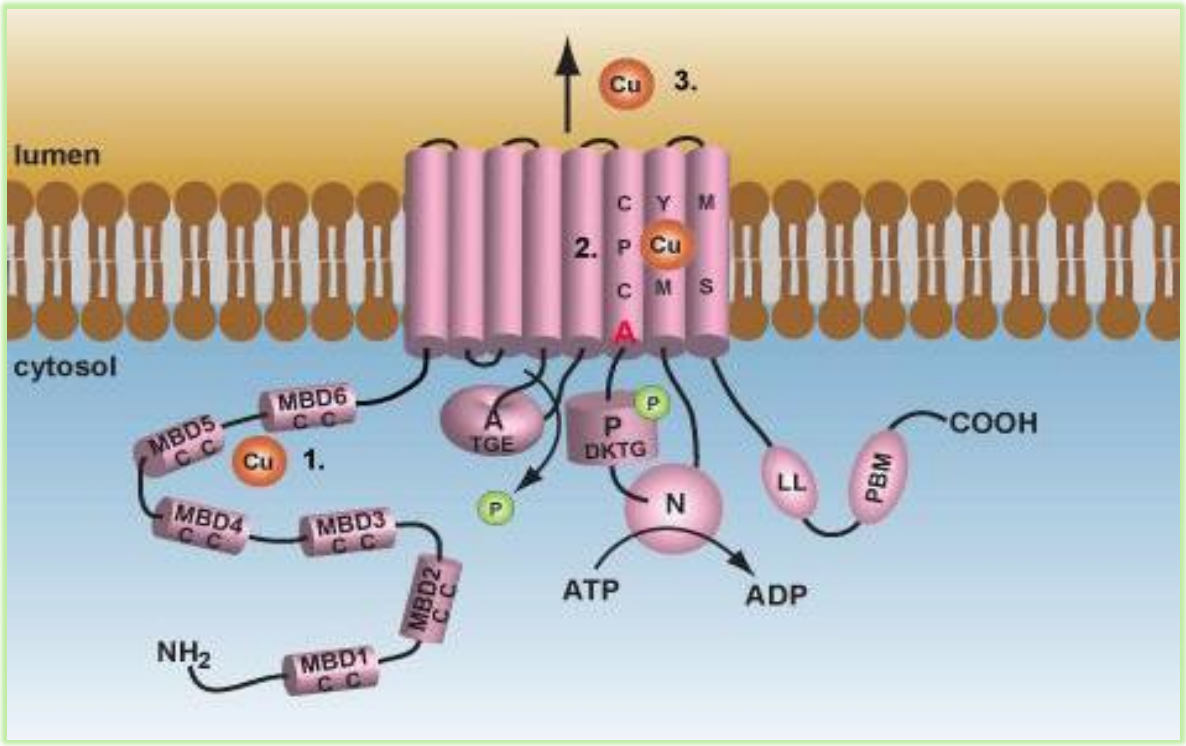


Figure 2.3: ATP7B protein: The six metal binding domains at the N-terminal are indicated by CxxC. Sequence motifs TGE, DKTG, SEHPL, GDGVND, and CPC depict conserved elements of Cu-ATPases.

ATP7B protein is a *Cu-transporting P-type ATPase,* which translocates copper from the cytosol across cellular membranes using the energy of ATP hydrolysis.

The N-terminal domain regulates Cu-ATPase activity and contains six repetitive sequences, each harboring the sequence motif GMT/HCxxCxxxIE. Each of these repeats forms a subdomain with a single metal binding site CxxC, where the two Cys residues, coordinate copper in the reduced Cu(I) form. The total stoichiometry is six copper ions per N-terminal domain.

By analogy with other P-type ATPases, the catalytic activity of ATP7B is mediated through coordinated action of the A-domain, containing conserved TGE motif, essential for enzymatic function of P-type ATPases and the ATP-binding domain. The latter consists of two portions: N-domain and P-domain that includes the site of catalytic phosphorylation and highly conserved DKTG motif. D1027 residue in DKTG motif is presumed to be the target of phosphorylation by the gamma-phosphate of ATP. Indeed, mutation of this aspartic acid residue in ATP7B completely prevents formation of an acylphosphate intermediate during hydrolysis process [1][3].

Copper translocation process by P-type ATPases occurs through a general cycling model involving several stages in which ATP hydrolysis drives translocation of the Cu(I) ion (Figure 2.4).

The stages involved in copper translocation are:

1. Binding of the Cu(I) ion to the MBD
2. Binding of ATP to the N-domain
3. After ATP binding to the N-domain, conformational changes take place that bring the ATP binding site within the N-domain in close proximity to the P-domain This could potentially promote ATP binding, phosphorylation of the P-domain and ADP release.
4. Translocation of the Cu(I) ion
5. The translocation stimulates the dephosphorylation of the P-domain mediated by intrinsic phosphatase activity, in which the A-domain plays a key role.

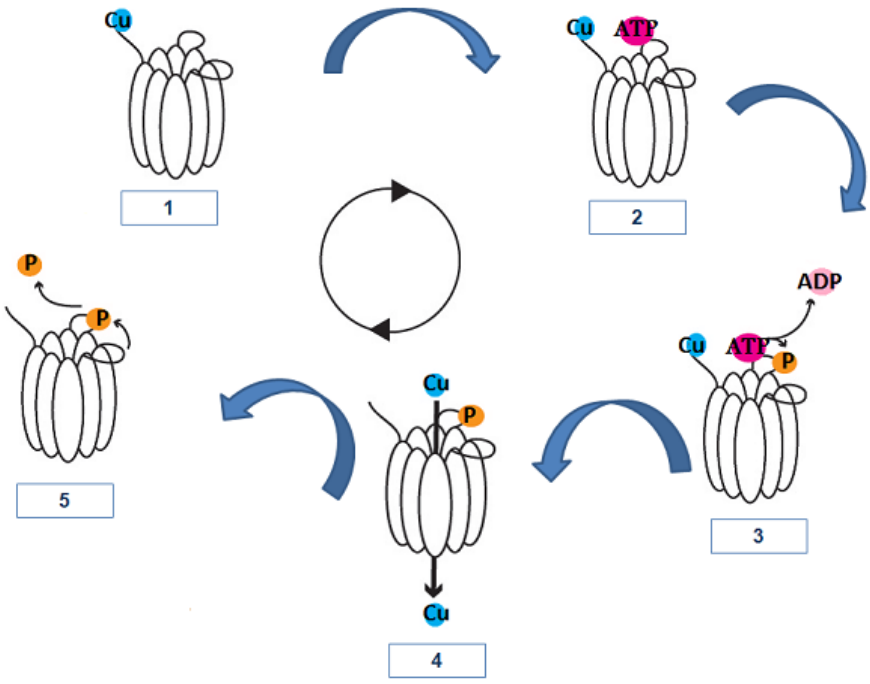


Figure 2.4: Schematic representation of Cu-ATPase catalytic cycle.

ATP7B is abundantly expressed in hepatocytes and localises within the trans-Golgi compartment. It is involved in cuproenzyme biosynthesis and is essential for homeostatic regulation of cellular copper. Under basal copper levels, ATP7B resides at the TGN, where copper is incorporated into cuproenzymes such as ceruloplasmin (CP), a ferroxidase that is subsequently secreted into the blood.

In response to elevated copper level, ATP7B binds Cu(I) ions and moves within cytoplasmic vesicles from TGN towards canalicular area of hepatocytes. Recent study demonstrated that ATP7B reaches lysosomes and imports Cu into their lumen. ATP7B enables lysosomes to undergo exocytosis through the interaction with p62 subunit of dynactin that allows lysosome translocation toward the canalicular surface of hepatocytes (Figure 2.5). Moreover, the activation of lysosomal exocytosis stimulates both the delivery of ATP7B to the canalicular membrane domains and the release of excess Cu into the bile indicating that ATP7B-containing lysosomes and lysosomal exocytosis have a key role in copper homeostasis [6]. When copper levels decrease, ATP7B is endocytosed and cycles back to the TGN [2].

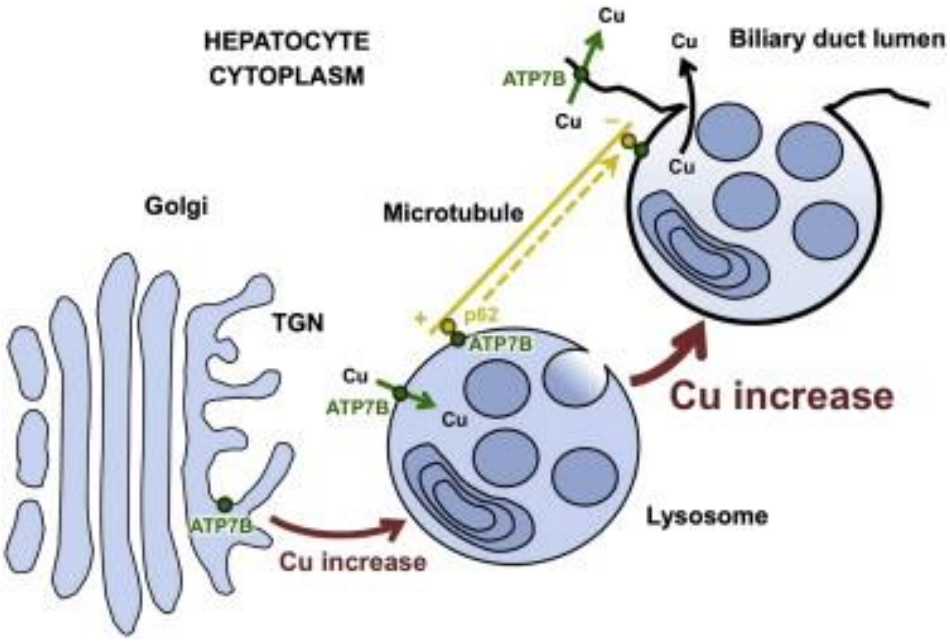


Figure 2.5: ATP7B protein trafficking in hepatocyte.

The malfunction of the ATP7B protein, correlated to the presence of mutations in the corresponding gene, which are often associated with genetic disorders, such as Wilson's disease, involves both a homeostatic balance disorder, resulting in accumulation of intracellular copper, and dysfunctions at the level of different cuproenzymes.

2.1.3 Wilson Disease

Wilson disease (WD) is an autosomal recessive disorder characterised by defective copper excretion, with an estimated incidence between 1:30 000 and 1:100 000; worldwide the heterozygous carrier rate is approximately 1:100. Almost 300 mutations in ATP7B that are associated with WD development have been described. In contrast to Menkes disease, relatively few small insertions/deletions and splice site mutations have been identified in patients affected by WD; in fact, almost 60% of all identified WD-causing mutations are missense mutations. The most prevalent mutations are H1069Q in Europe and North America, and R778L in southeast Asia [3]. Several ATP7B mutations, including H1069Q and R778L, result in protein product, which, although still functional, remain in the endoplasmic reticulum, presumably because misfolded [7][8], or is trapped in cytoplasmic clusters not reaching Cu excretion sites. This mislocalization leads to toxic accumulation of copper primarily in the liver of WD patients, but also in other organs, such as brain, kidney, cornea.

Clinical features of WD include hepatic abnormalities (cirrhosis and chronic hepatitis), neurological defects (parkinsonian features), and psychiatric symptoms (personality changes, depression, psychosis). In some severe cases, patients present fulminant liver failures. A

characteristic feature of patients with WD is the Kayser–Fleischer ring, a deposition of copper in the Descemet membrane visible as a gold-brown ring around the periphery of the cornea [9].

In addition, serum levels of the cuproenzyme ceruloplasmin are often found greatly reduced in WD patients as a result of rapid degradation of the copper-free form of ceruloplasmin [3].

2.1.4 Treatment of Wilson Disease

The therapeutic goal in WD patients is to maintain normal copper homeostasis. In most circumstances, this can be accomplished with a number of different oral chelating agents, such as D-Penicillamine, Trientin, Tetrathiomolybdate. Penicillamine has been shown to effectively ameliorate hepatic and neuropsychiatric signs and symptoms in most patients and to prevent the onset of disease in asymptomatic individuals. Patients show a rapid and dramatic response within the first few months of penicillamine treatment, but poor compliance may result in an acute neurological deterioration, presumably secondary to rapid changes in copper distribution with the central nervous system. The measurement of urinary copper excretion can be a useful sign of initial response to the treatment. Although it is effective, penicillamine is toxic, gives hypersensitivity reactions, might induce bone marrow suppression and the development of autoimmune disease, and patients must be monitored during treatment.

Trientine and tetrathiomolybdate are alternative drugs with distinct mechanisms of action that may be also used with similar effectiveness.

Dietary restriction is reasonable suggested to patients that they avoid foods with an abundance of copper (shellfish, chocolate, nuts, and legumes).

In general, these treatment options are very effective; however, if treatment is ineffective, or in circumstances of fulminant liver failure, liver transplantation remains the only cure [9].

2.1.5 Aim of the Project

Although the most frequent ATP7B mutants potentially are able to transport Cu, they cannot reach the Cu excretion sites to remove excess Cu from hepatocytes. ER retention of ATP7B mutants, such as ATP7B(H1069Q), occurs because the mutation destabilizes the native structure which is recognized as defective and directed toward the ER-associated protein degradation (ERAD) pathway.

This project, carried out in collaboration with the research group of Dr. Roman Polishchuk of TIGEM (Telethon Institute of Genetics and Medicine), was focused on the investigation of molecular mechanisms impaired in Wilson disease (WD). In order to evaluate which molecular

pathways result altered by expression of ATP7B mutants, a comparison between the interactomes of the wild type protein and the mutant ATP7B(H1069Q) was performed by functional proteomic approach. This latter lead to the identification of molecular pathways crucial for the recovery of partially or fully active ATP7B(H1069Q) mutant from the ER to the appropriate functional compartment(s). These pathways would be also targets for drugs in new therapies applicable for the majority of Wilson disease patients.

2.2 Materials and Methods

2.2.1 Materials

All chemicals for electrophoresis, 30% Acrylamide Bis-acrylamide solution, Tetramethylethylenediamine (TEMED), sodium dodecylsulfate (SDS), Ammonium Persulphate (APS), Tris-Glicine-SDS (TGS), Tris-HCl and the mixture of protein standards of known molecular weight (total blue) were from BIO-RAD.

In addition, Triton X-100, "BioRad Protein Assay" kit and bovine serum albumine (BSA) used in Bradford assay were purchased from BIO-RAD.

Protease inhibitors (Complete mini EDTA-free) were ROCHE.

The Colloidal Coomassie Blue Brilliant was PIERCE.

Beads for pre-cleaning and immunoprecipitation (Sepharose Protein-A), dithiothreitol (DTT), trypsin, sodium chloride (NaCl), ethylenediaminetetraacetic acid (EDTA), glycerol, NP-40 detergent, trifluoroacetic acid (TFA) and acetonitrile (ACN) LC-MS Grade were SIGMA-ALDRICH products.

The iodoacetoamide (IAM) and ammonium bicarbonate (AMBIC) were from FLUKA. Formic acid (HCOOH) and acetonitrile HPLC Grade were Baker products.

The antibodies anti-GFP and the HepG2 transfected cells were provided by our collaborators of TIGEM.

2.2.2 Methods

2.2.2.1 Cell Lysis and Protein Quantification

Protein extracts were prepared starting from HepG2 cells grown in presence of 100μM $CuSO_4$ and infected with GFP-ATP7B(WT), GFP-ATP7B(H1069Q) and GFP. This latter constitutes the control.

Lysis buffer composition was:

- ✓ 20mM Tris HCl pH=8
- ✓ 150mM NaCl
- ✓ 1mM EDTA pH=8
- ✓ 10% glycerol
- ✓ 0.5% NP-40
- ✓ 0.5% Triton X-100
- ✓ 1 mini pill of protease inhibitor for 10 mL of lysis buffer

A ratio of 1:5 (v/v) pellet/buffer was used for lysis of each cell pellet. After buffer addition, pellets were left for 10 minutes in ice, and successively shacked and agitated on the wheel at 4°C for 30 minutes. Finally, the suspensions were centrifuged for 15 minutes at 13200 rpm in order to separate protein extracts from debris.

Bradford assays were carried out for a quantitative evaluation of protein extracts. Bovine serum albumin (BSA) was used as standard for the calibration curve construction, obtained by reporting the absorbance measurements at 595 nm in function of concentrations of BSA standard solutions.

The total protein concentrations defined by Bradford protein assay are reported in Table 2.1.

Table 2.1: Total protein concentration from Bradford assay.

	Total protein concentration
HepG2 expressing GFP-ATP7B(WT)	10.6 µg/µL, (11.8mg total protein)
HepG2 expressing GFP -ATP7B(H1069Q)	8 µg/µL, (8.8mg total protein)
HepG2 expressing empty-GFP (control)	9.7 µg/µL, (10.7mg total protein)

2.2.2.2 Isolation of Protein Complexes by Immunoprecipitation

7 mg of protein extract from each HepG2 protein pellet were subjected to the immunoprecipitation procedure. As first step, every cell extract was pre-cleared in order to remove unspecific proteins adsorbed on beads by incubating protein extracts with Sepharose beads Protein-A (20µL of slurry resin per mg of protein extract), previously conditioned in lysis buffer, for 3h at 4°C on the wheel.

Pre-cleaned extracts were then incubated with specific anti-GFP antibody (policlonal mouse) (2.5 µg of antibody per mg of protein extract), over-night at 4°C on the wheel.

Finally, the samples were transferred in new tubes and incubated with beads Protein-A (40µL of slurry resin per mg of protein extract) for 4h at 4°C under gently agitation on the wheel in order to capture antigen-antibody complexes. The unbound proteins were then removed, and the beads were repetitively washed with lysis buffer containing either low (150mM) or high (300mM) NaCl concentrations, in order to remove unspecific proteins. Proteins retained were eluted by using Laemli Buffer (100mM Tris HCl pH=6.8, 4% SDS, 20% glycerol, bromophenol blue) for 10 min at 99°C in thermomixer.

2.2.2.3 Preparative SDS-PAGE

100mM DTT (dithiothreitol) was added to each samples following 10 minutes of boiling at 99°C.

Sample and control eluates were loaded on a 16x16 cm, polyacrylamide gel gradient 6%-15% SDS-PAGE fractionation. The electrophoretic run was performed at initial voltage of 200 V.

The gel was stained by Colloidal Blue Comassie over-night and the excess of dye was removed by extensive washing with deionized water.

2.2.2.4 Hydrolysis in situ

After electrophoresis and staining, the whole lanes were cute in several bands by scalpel, chopped and placed in tubes. The bands were bleached by ACN, removed following agitation with vortex and fast centrifugation; then, 50µl of NH_4HCO_3 (AMBIC) 50mM, pH 8 was added and left few minutes for the gel pieces rehydration. These alternative washings were repeated until the complete destaining of gel bands.

Protein bands were treated with 60µL of 10mM DTT in 50mM NH_4HCO_3, for 45 min at 56°C to reduce cysteine residues. At the end of the reduction, the bands were first dehydrated with ACN and then rehydrated with 60µL 50mM NH_4HCO_3, containing 55mM iodoacetamide for the alkylation reaction. This latter was carried out in the dark at room temperature for 30 minutes. The excess of reagent was finally removed by washing with ACN 50mM/NH_4HCO_3 as previously described.

The dehydrated gel bands were then treated with 10µL of 10ng/µL trypsin solution in 50mM NH_4HCO_3 pH=8 at 4°C, for 1h. Finally, a supplemental volume of 50mM NH_4HCO_3 was added to cover the gel bands and samples were placed over-night at 37°C.

At the end of hydrolysis, samples were centrifuged at 13200 rpm for few seconds and the supernatants were collected and placed in tube. Each sample was acidified by 20% trifluoroacetic acid. Then the remaining gel pieces were shrunk with 50µL of acetonitrile, in order to extract any peptides still present in the gel. The extracts were joined to the respective supernatants, previously removed. A second extraction was performed adding 20µL of HCOOH 0,2% followed by ACN to gel bands and the surnatants recovered and joined to respective peptide mixtures. The obtained mixtures were dried by a Speed-Vac system.

2.2.2.5 LC-MS/MS Analysis

Each peptide mixture was re-suspended in 10μL of 0.2% HCOOH, and analyzed by nano-chromatography tandem mass spectrometry (MS/MS) on a CHIP MS Ion Trap XCT Ultra (Agilent Technologies, Palo Alto, CA).

After loading, the peptide mixture was first concentrated and desalinated in the reverse-phase pre-column working at a flow rate of 4μL/min with 0.1% formic acid as eluent. Each peptide sample was then fractionated on a C18 reverse-phase capillary column (C18 EasyColumn ID=75 μm, L=43 mm, *Agilent Technologies*) working at a flow rate of 200nL/min, with a gradient of eluent B (0,2% formic acid, 95% acetonitrile LC-MS Grade) and eluent A (0,2% formic acid, 2% acetonitrile LC-MS Grade) from 5 to 60% in 50 mins.

Peptide analysis was performed using data-dependent acquisition (DDA) of one MS scan (mass range from 400 to 2000 m/z) followed by MS/MS scans of the three most abundant ions in each MS scan, obtained by CID (*Collision Induced Fragmentation*) fragmentation in the ion trap.

2.2.2.6 Proteins Identification

Raw data from nano-LC-MS/MS analyses were processed and converted in .mgf files to be introduced into the MASCOT software (Matrix Science Boston, USA) to search a non-redundant protein database. Both peptide mass data and the data obtained by fragmentation spectra were included in the *peak list* for protein identification.

The research was done by setting the following parameters:

- ✓ Database: NCBI
- ✓ Taxonomy: Homo sapiens
- ✓ Enzyme: Trypsin
- ✓ Fixed modification: carbamidomethyl (C)
- ✓ Variable modification: oxidation (M), pyro-Glu (N-term Q), pyrocarbamidomethyl (N-term C)
- ✓ Peptide tolerance: 600 ppm
- ✓ MS/MS tolerance: ± 0,6 Da
- ✓ Ions charge: +2, +3
- ✓ Instrument: ESI-TRAP

It has also been used the software ExPASy Blast Form to assign to each protein a code that uniquely identify each sequence in the annotated database UniProt KB.

2.3 Results

The investigation of molecular mechanism involving ATP7B(WT) and ATP7B(H1069Q) protein trafficking was carried out by the identification of protein partners interacting with both GFP tagged proteins expressed in HepG2 cells.

HepG2 cells expressing GFP-ATP7B(WT), GFP-ATP7B(H1069Q) and GFP (control) were lysated as reported in Methods paragraph 2.2.2.1 and 7 mg of each protein extract immunoprecipited as reported in 2.2.2.2 section.

Proteins specifically retained by ATP7B-GFP proteins or by GFP (control) were eluted in denaturing conditions and individual proteins were fractionated by SDS-PAGE. 30 Protein bands were excised from the gel (Figure 2.6), digested in situ with trypsin and the resulting peptide mixtures were directly analyzed by nano-chromatography tandem mass spectrometry (MS/MS) on a CHIP MS Ion Trap XCT Ultra (Agilent Technologies, Palo Alto, CA).

Since the samples were fractionated by one-dimensional electrophoresis, a low-resolution technique, each band might contain more than one protein. Therefore, the identification of all proteins required the employment of tandem mass spectrometry associated with reverse phase liquid chromatography (nanoLC-MS/MS) methodologies for the analyses of complex peptide mixtures as those obtained by in situ hydrolysis of each band.

The data obtained by mass spectrometry analyses, which included both the m/z ratios of intact peptides and the correspondent ions (peak lists), were inserted in the MS/MS Ion Search option of Mascot software, leading to proteins identification.

Unambiguous protein identification is achieved by a large number of peptides, with scores exceeding the minimum confidence threshold suggested by Mascot, that allows a high percentage of sequence coverage. The peptide scores are determined by the quality of the fragmentation spectra in terms of signal/ noise ratio, the accuracy of the m/z values and consequently by the number of signals that could be assigned to the expected amino acid sequences present in the databases, and, finally, the size of the database in which the search is performed.

The databases used for the ATP7B interacting protein identification was NCBI, which has the advantage to be continuously updated, although suffers from redundancy problems. Once identified, all proteins have been associated to UniProt database code.

The lists of ATP7B(WT) and ATP7B(H1069Q) putative interactors (Table 2.2) were obtained by subtracting the proteins identified in the control (IP GFP) to those identified in IP ATP7B(WT) and ATP7B(H1069Q) samples. For each protein, the molecular weight, the total number of

identified peptides (in parentheses those with a score higher than the reliability threshold defined by Mascot), the UniProt code, the gene name, keywords for GO Biological process are also reported.

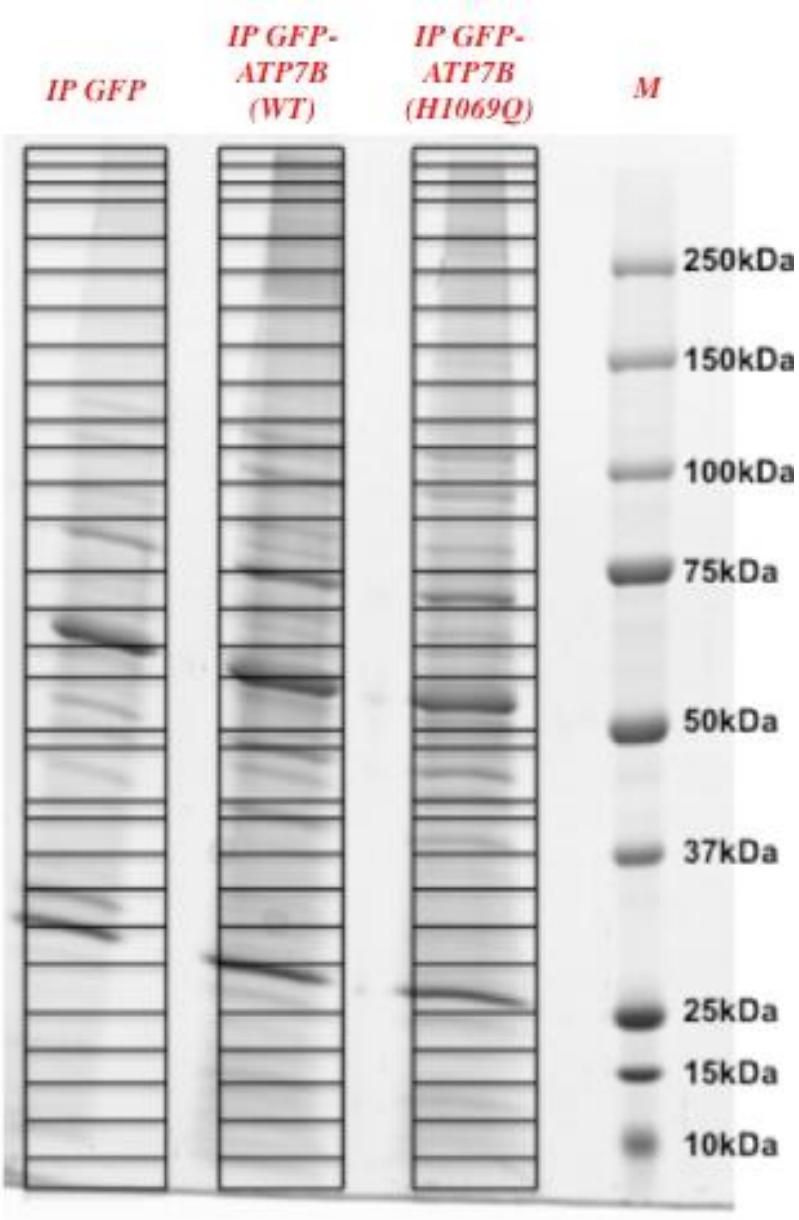


Figure 2.6: SDS-PAGE of the immunoprecipitated samples from HepG2 cells expressing 1. GFP-ATP7B(WT), 2. GFP-ATP7B(H1069Q) and 3. GFP (control). M represents a mixture of known molecular weight proteins used as standards.

Table 2.2: Putative ATP7B(WT) and ATP7B(H1069Q) protein partners. Protein name, gene name, number of peptides in both experiments, in parenthesis those with scores above the threshold UniProt code and GO Biological Process are reported.

MW (kDa)	Protein name	Gene	Peptides in WT	Peptides in H1069Q	UniProt code	GO Biological Process
>250	Cation-independent mannose-6-phosphate receptor	IGF2R	2(1)	-	P11717	Signal transduction
	Myosin-10	MYH10	27(11)	37(11)	P35580	Actin filament-based movement
	Myosin-14	MYH14	5(1)	8(1)	Q7Z406	Actin filament-based movement
	Copper-transporting ATPase 2	ATP7B	35(21)	27(14)	P35670	Cellular copper ion homeostasis
	Eukaryotic translation initiation factor 3 subunit A	EIF3A	8(2)	6(2)	Q14152	Translation
150-100	Tight junction protein ZO-2	TJP2	8(5)	12(4)	Q9UDY2	Hippo signaling
	Vigilin	HDLBP	7(2)	9(5)	Q00341	Lipid transport
	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-1	PLCG1	-	3(1)	P19174	Activation of MAPKK activity
	Large proline-rich protein BAG6	BAG6	3(1)	-	P46379	Protein stabilization
	Protein transport protein Sec31A	SEC31A	4(2)	4(3)	O94979	ER to Golgi vesicle-mediated transport
	Hypoxia up-regulated protein 1	HYOU1	9(2)	15(6)	Q9Y4L1	ER to Golgi vesicle-mediated transport
	LIM domain only protein 7	LMO7	4(2)	-	Q8WWI1	Regulation of signaling
	Ubiquitin-associated protein 2-like	UBAP2L	4(2)	5(2)	Q14157	Regulation of gene expression
	Ubiquitin-associated protein 2	UBAP2	3(1)	-	Q5T6F2	Regulation of gene expression
	Cullin-associated NEDD8-dissociated protein 1	CAND1	8(2)	9(4)	Q86VP6	Protein ubiquitination

MW (kDa)	Protein name	Gene	Peptides in WT	Peptides in H1069Q	UniProt code	GO Biological Process
150-100	Ubiquitin conjugation factor E4 A	UBE4A	-	2(1)	Q14139	Ubl conjugation pathway
	Cell cycle and apoptosis regulator protein 2	CCAR2	6(2)	-	Q8N163	Transcription
	Kinesin-1 heavy chain	KIF5B	12(7)	9(4)	P33176	Vesicle transport along microtubule
	Importin-5	IPO5	11(6)	11(6)	O00410	Protein transport
	Poly [ADP-ribose] polymerase 1	PARP1	5(3)	3(2)	P09874	Transcription
	Protein transport protein Sec24C	SEC24C	3(1)	-	P53992	ER to Golgi vesicle-mediated transport
100	General vesicular transport factor p115	USO1	9(3)	6(1)	O60763	ER to Golgi vesicle-mediated transport
	Heat shock protein 105 kDa	HSPH1	6(3)	6(4)	Q92598	Chaperone mediated protein folding
	5'-3' exoribonuclease 2	XRN2	3(2)	-	Q9H0D6	Regulation of transcription
	Transferrin receptor protein 1	TFRC	4(1)	-	P02786	Cellular iron ion homeostasis
	Alpha-actinin-1	ACTN1	7(5)	7(5)	P12814	Actin filament organization
	Eukaryotic translation initiation factor 4 gamma 2	EIF4G2	2(2)	-	P78344	Translation regulation
	Pyridoxal-dep decarboxylase domain-containing p. 1	PDXDC1	8(2)	4(2)	Q6P996	Carboxylic acid metabolic process
	Conserved oligomeric Golgi complex subunit 5	COG5	2(1)	-	Q9UP83	ER to Golgi vesicle-mediated transport
	DNA replication licensing factor MCM3	MCM3	-	6(3)	P25205	DNA replication
	Heat shock 70 kDa protein 4L	HSPA4L	-	2(2)	O95757	Unfolded protein binding

MW (kDa)	Protein name	Gene	Peptides in WT	Peptides in H1069Q	UniProt code	GO Biological Process
100	Dynamin-2	DNM2	-	4(2)	P50570	Endocytosis
	26S proteasome non-ATPase regulatory subunit 2	PSMD2	-	3(1)	Q13200	MAPK cascade
	Rho GTPase-activating protein 18	ARHGAP18	-	4(2)	Q8N392	Signal transduction
	E3 UFM1-protein ligase 1	UFL1	-	5(1)	O94874	Ubl conjugation pathway
	Interleukin enhancer-binding factor 3	ILF3	-	7(3)	Q12906	Transcription
	Calnexin	CANX	-	7(5)	P27824	Unfolded protein binding
	DNA replication licensing factor MCM6	MCM6	-	4(1)	Q14566	DNA replication
	Methionine--tRNA ligase, cytoplasmic	MARS	-	7(4)	P56192	Protein biosynthesis
100-75	Mitochondrial inner membrane protein	IMMT	4(2)	7(4)	Q16891	Mitochondrial calcium ion homeostasis
	DNA replication licensing factor MCM7	MCM7	6(2)	9(3)	P33993	DNA replication
	6-phosphofructokinase, liver type	PFKL	-	3(3)	P17858	Glycolysis
	Polyadenylate-binding protein 4	PABPC4	-	8(3)	Q13310	Translation
	ORF protein	ORF	-	8(6)	Q7KYM9	mRNA splicing
	Thimet oligopeptidase	THOP1	-	2(1)	P52888	Peptide metabolic process
	Long-chain-fatty-acid--CoA ligase 4	ACSL4	-	10(4)	O60488	Lipid metabolism
	Lysine--tRNA ligase	KARS	-	8(3)	Q15046	Protein biosynthesis
	Eukaryotic translation initiation factor 4B	EIF4B	3(2)	3(2)	P23588	Protein biosynthesis
	NADH-ubiquinone oxidoreductase 75 kDa subunit	NDUFS1	7(2)	6(2)	P28331	Mitochondrial electron transport
	Sec1 family domain-containing protein 1	SCFD1	9(6)	2(1)	Q8WVM8	ER-Golgi transport
	Sorting nexin-1	SNX1	-	3(1)	Q13596	Protein transport

MW (kDa)	Protein name	Gene	Peptides in WT	Peptides in H1069Q	UniProt code	GO Biological Process
100-75	Annexin A6	ANXA6	10(3)	2(1)	P08133	Ion transmembrane transport
	Transmembrane protein 214	TMEM214	4(2)	3(1)	Q6NUQ4	Apoptosis
75	Arginine--tRNA ligase, cytoplasmic	RARS	4(2)	4(2)	P54136	Protein biosynthesis
	ATP-dependent Clp protease ATP-binding subunit clpX-like	CLPX	5(2)	-	O76031	Chaperone mediated protein folding
	EH domain-containing protein 1	EHD1	3(1)	-	Q9H4M9	Intracellular protein transport
	Pyruvate kinase isozymes M1/M2	PKM	-	4(1)	P14618	Glycolysis
	Asparagine--tRNA ligase, cytoplasmic	NARS	-	3(1)	O43776	Protein biosynthesis
	EH domain-containing protein 4	EHD4	-	3(2)	Q9H223	Regulation of endocytosis
	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	DLAT	-	3(1)	P10515	Glucose metabolism
	Phenylalanine--tRNA ligase beta subunit	FARSB	-	6(4)	Q9NSD9	Translation
	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	RPN1	-	8(5)	P04843	Cellular protein modification process
	Inositol-3-phosphate synthase 1	ISYNA1	3(1)	-	Q9NPH2	Lipid metabolism
75-50	Glucose-6-phosphate isomerase	GPI	10(6)	9(4)	P06744	Glycolysis
	T-complex protein 1 subunit delta	CCT4	8(5)	12(4)	P50991	Chaperone mediated protein folding
	Importin subunit alpha-2	KPNA2	6(2)	5(1)	P52292	Protein transport
	Serine--tRNA ligase, cytoplasmic	SARS	4(2)	3(2)	P49591	Protein biosynthesis
	Hydroxymethylglutaryl-CoA synthase, cytoplasmic	HMGCS1	8(4)	7(6)	Q01581	Lipid metabolism

MW (kDa)	Protein name	Gene	Peptides in WT	Peptides in H1069Q	UniProt code	GO Biological Process
75-50	Inosine-5'-monophosphate dehydrogenase 2	IMPDH2	9(2)	6(4)	P12268	Purine biosynthesis
	T-complex protein 1 subunit epsilon	CCT5	-	6(3)	P48643	Chaperone mediated protein folding
	T-complex protein 1 subunit beta	CCT2	-	8(4)	P78371	Chaperone mediated protein folding
	tRNA-splicing ligase RtcB homolog	RTCB	-	2(1)	Q9Y3I0	tRNA processing
	DnaJ homolog subfamily C member 7	DNAJC7	-	2(1)	Q99615	Protein folding
	Coiled-coil domain-containing protein 47	CCDC47	-	3(1)	Q96A33	ER overload response
	Beta-1-syntrophin	SNTB1	-	3(1)	Q13884	Muscle contraction
	Aldehyde dehydrogenase, mitochondrial	ALDH2	2(2)	4(2)	P05091	carbohydrate metabolic process
50-37	Serpin H1	SERPINH1	2(2)	3(3)	P50454	Stress response
	Rab GDP dissociation inhibitor beta	GDI2	5(2)	6(1)	P50395	Vesicle-mediated transport
	Synaptic vesicle membrane protein VAT-1 homolog	VAT1	4(2)	7(3)	Q99536	Regulation of mitochondrial fusion
	Hydroxysteroid dehydrogenase-like protein 2	HSDL2	3(2)	2(2)	Q6YN16	Oxidation-reduction process
	Ribonucleoside-diphosphate reductase subunit M2	RRM2	2(2)	-	P31350	DNA replication
	Erlin-1	ERLIN1	2(2)	3(2)	O75477	Lipid metabolism
	Annexin A11	ANXA11	-	2(1)	P50995	Cell cycle
	26S proteasome non-ATPase regulatory subunit 13	PSMD13	-	5(2)	Q9UNM6	MAPK cascade
	Protein DDI1 homolog 2	DDI2	-	2(1)	Q5TDH0	Proteolysis
	Hsp90 co-chaperone Cdc37	CDC37	-	3(1)	Q16543	Protein folding

MW (kDa)	Protein name	Gene	Peptides in WT	Peptides in H1069Q	UniProt code	GO Biological Process
50-37	Cytochrome b-c1 complex subunit 1, mitochondrial	UQCRC1	-	4(1)	P31930	Oxidation-reduction process
	Thioredoxin domain-containing protein 5	TXNDC5	-	5(2)	Q8NBS9	Response to endoplasmic reticulum stress
	Lymphokine-activated killer T-cell-originated protein kinase	PBK	-	2(2)	Q96KB5	Regulation of stress-activated MAPK cascade
	6-phosphogluconate dehydrogenase, decarboxylating	PGD	-	6(3)	P52209	Oxidation-reduction process
	Alpha-aminoadipic semialdehyde dehydrogenase	ALDH7A1	-	4(2)	P49419	Oxidation-reduction process
	Aldehyde dehydrogenase X, mitochondrial	ALDH1B1	-	4(2)	P30837	Oxidation-reduction process
	DnaJ homolog subfamily B member 1	DNAJB1	3(2)	3(1)	P25685	Chaperone mediated protein folding
37	Ubiquitin fusion degradation protein 1 homolog	UFD1L	3(2)	5(3)	Q92890	Ubl conjugation pathway
	26S proteasome non-ATPase regulatory subunit 7	PSMD7	-	2(2)	P51665	MAPK cascade
	Nucleophosmin	NPM1	-	2(1)	P06748	Response to stress
	Proteasome activator complex subunit 1	PSME1	-	8(3)	Q06323	MAPK cascade
	Ubiquitin thioesterase OTUB1	OTUB1	-	2(1)	Q96FW1	Ubl conjugation pathway
	Proteasome activator complex subunit 2	PSME2	-	5(2)	Q9UL46	MAPK cascade
	Ubiquitin carboxyl-terminal hydrolase isozyme L3	UCHL3	-	2(2)	P15374	Ubl conjugation pathway
	PDZ and LIM domain protein 1	PDLIM1	5(3)	4(2)	O00151	Regulation of transcription
37-25	Voltage-dependent anion-selective channel protein 1	VDAC1	6(5)	6(5)	P21796	Ion transport

MW (kDa)	Protein name	Gene	Peptides in WT	Peptides in H1069Q	UniProt code	GO Biological Process
37-25	Receptor of activated protein C kinase 1	RACK1	11(5)	-	P63244	Regulation of protein localization
	ATP synthase subunit gamma, mitochondrial	ATP5C1	3(2)	-	P36542	Ion transport
	60S ribosomal protein L7	RPL7	8(2)	-	P18124	Translation
	ADP/ATP translocase 3	SLC25A6	7(2)	-	P12236	Transport
	60S ribosomal protein L7a	RPL7A	8(4)	-	P62424	Translation
	Signal recognition particle receptor subunit beta	SRPRB	5(3)	-	Q9Y5M8	Signal transduction
	Estradiol 17-beta-dehydrogenase 8	HSD17B8	2(2)	-	Q92506	Lipid metabolism
	Proteasome subunit alpha type-1	PSMA1	4(2)	6(3)	P25786	Protein polyubiquitination
	Vesicle-associated membrane protein-associated protein B/C	VAPB	3(2)	-	O95292	ER to Golgi vesicle-mediated transport
	60S ribosomal protein L8	RPL8	3(2)	4(2)	P62917	Translation
	Glutathione S-transferase omega-1	GSTO1	6(2)	-	P78417	Oxidation-reduction process
	Proteasome subunit alpha type-7	PSMA7	6(3)	4(3)	O14818	Protein polyubiquitination
	Tumor protein D54	TPD52L2	4(3)	3(2)	O43399	Regulation of cell proliferation
	Proteasome subunit alpha type-6	PSMA6	4(3)	4(3)	P60900	Protein polyubiquitination
	14-3-3 protein theta	YWHAQ	2(2)	-	P27348	Signal transduction
	14-3-3 protein eta	YWHAH	2(2)	-	Q04917	Intracellular protein transport
	Proteasome subunit alpha type-3	PSMA3	4(3)	3(2)	P25788	Protein polyubiquitination
	Phosphoglycerate mutase 1	PGAM1	4(2)	5(2)	P18669	Glycolysis
	B-cell receptor-associated protein 31	BCAP31	2(2)	-	P51572	ER-Golgi transport

MW (kDa)	Protein name	Gene	Peptides in WT	Peptides in H1069Q	UniProt code	GO Biological Process
37-25	Ran-specific GTPase-activating protein	RANBP1	2(2)	-	P43487	Signal transduction
	Ras-related protein Rab-5C	RAB5C	4(3)	2(2)	P51148	Protein transport
	Vesicle-trafficking protein SEC22b	SEC22B	3(3)	3(2)	O75396	ER-Golgi transport
	Eukaryotic translation initiation factor 6	EIF6	3(2)	2(2)	P56537	Protein biosynthesis
	40S ribosomal protein S9	RPS9	4(1)	-	P46781	Protein biosynthesis
	Superoxide dismutase [Mn], mitochondrial	SOD2	2(1)	-	P04179	Oxidation-reduction process
	Apoptosis regulator BAX	BAX	4(2)	4(2)	Q07812	Apoptosis
	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 4	NDUFB4	3(2)	-	O95168	Oxidation-reduction process
	NADH dehydrogenase[ubiquinone]1 alpha subcomplex subunit 2	NDUFA2	2(2)	-	O43678	Oxidation-reduction process
15-10	Peptidyl-prolyl-cis-trans isomerase FKBP1A	FKBP1A	-	3(3)	P62942	Protein folding
	COX5B protein	COX5B	-	3(2)	Q6FHJ9	Unfolded protein binding
	D-dopachrome decarboxylase	DDT	-	2(1)	P30046	Melanin biosynthesis
	Protein transport protein Sec61 subunit beta	SEC61B	-	2(1)	P60468	Ubl conjugation pathway
	Glutaredoxin-1	GLRX	-	2(1)	P35754	Cell redox homeostasis

2.4 Discussion

The molecular mechanisms involving ATP7B protein traffic and impaired in WD are still unclear. Several ATP7B variants, including one of the most common mutants H1069Q, although still functional, are mislocated in the endoplasmic reticulum, resulting in the toxic buildup of Cu in the liver of WD patients. Therefore, the identification of molecular targets for correction of endoplasmic reticulum-retained ATP7B mutants, might lead to the protein to proper functional sites in the cell and restore Cu excretion. This would benefit a large number of WD patients. In order to understand the molecular pathways resulting impaired by expression of ATP7B mutants, a comparative interactomes investigation of ATP7B(WT) and ATP7B(H1069Q) was performed by employing a functional proteomic approach.

The investigation of ATP7B(WT) and ATP7B(H1069Q) interactomes were carried out in HepG2 cells, and allowed the identification of 83 and 103 putative interactors respectively, classified according to the main biological function using information reported in literature (Figure 2.7 and 2.8).

In addition, a bioinformatic analysis of identified proteins was performed with Cytoscape tool allowing us to obtain a graphical representation of the molecular interactions networks both for ATP7B(WT) and ATP7B(H1069Q) interactors (Figure 2.9).

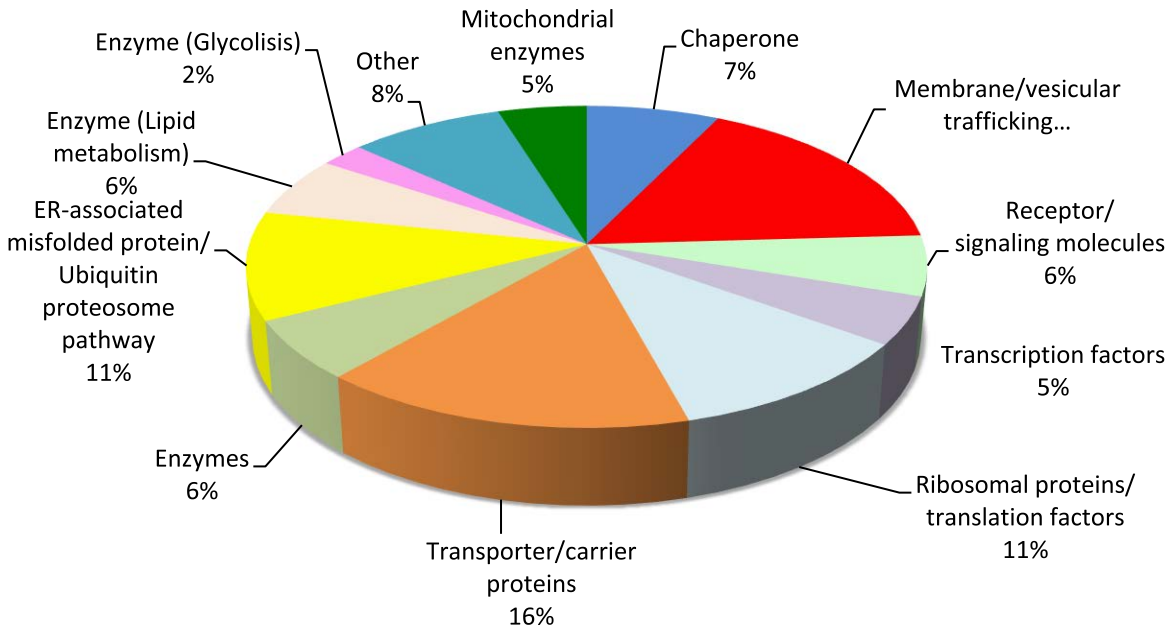


Figure 2.7: Functional classification of ATP7B(WT) putative interactors.

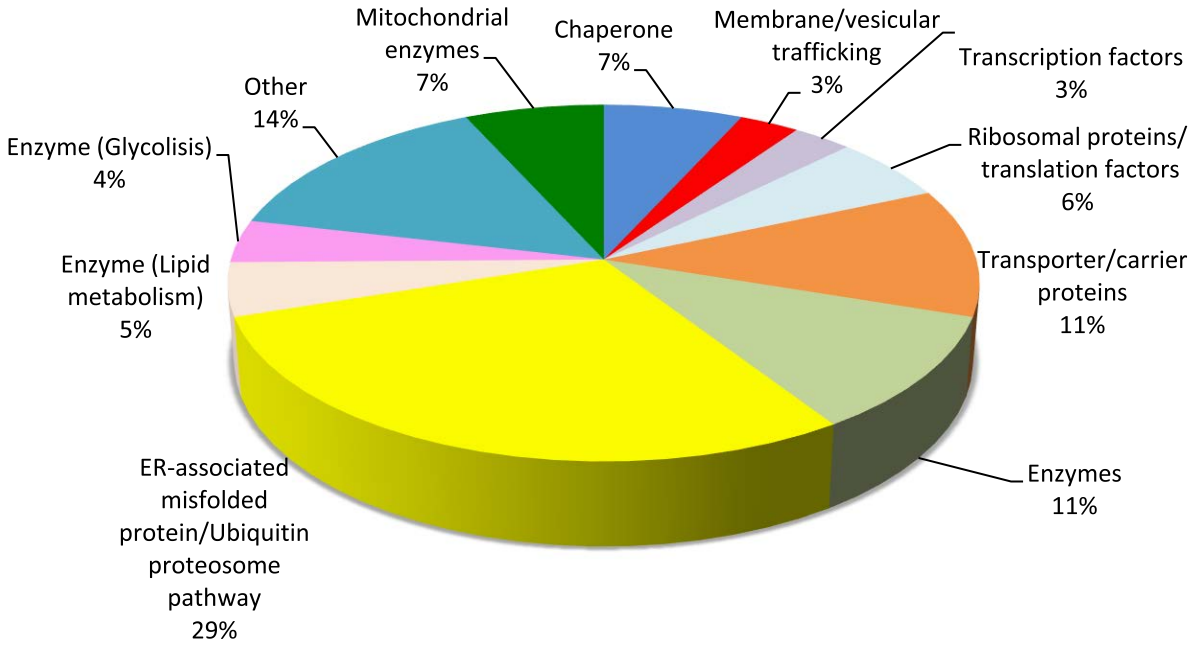


Figure 2.8: Functional classification of ATP7B(H1069Q) putative interactors.

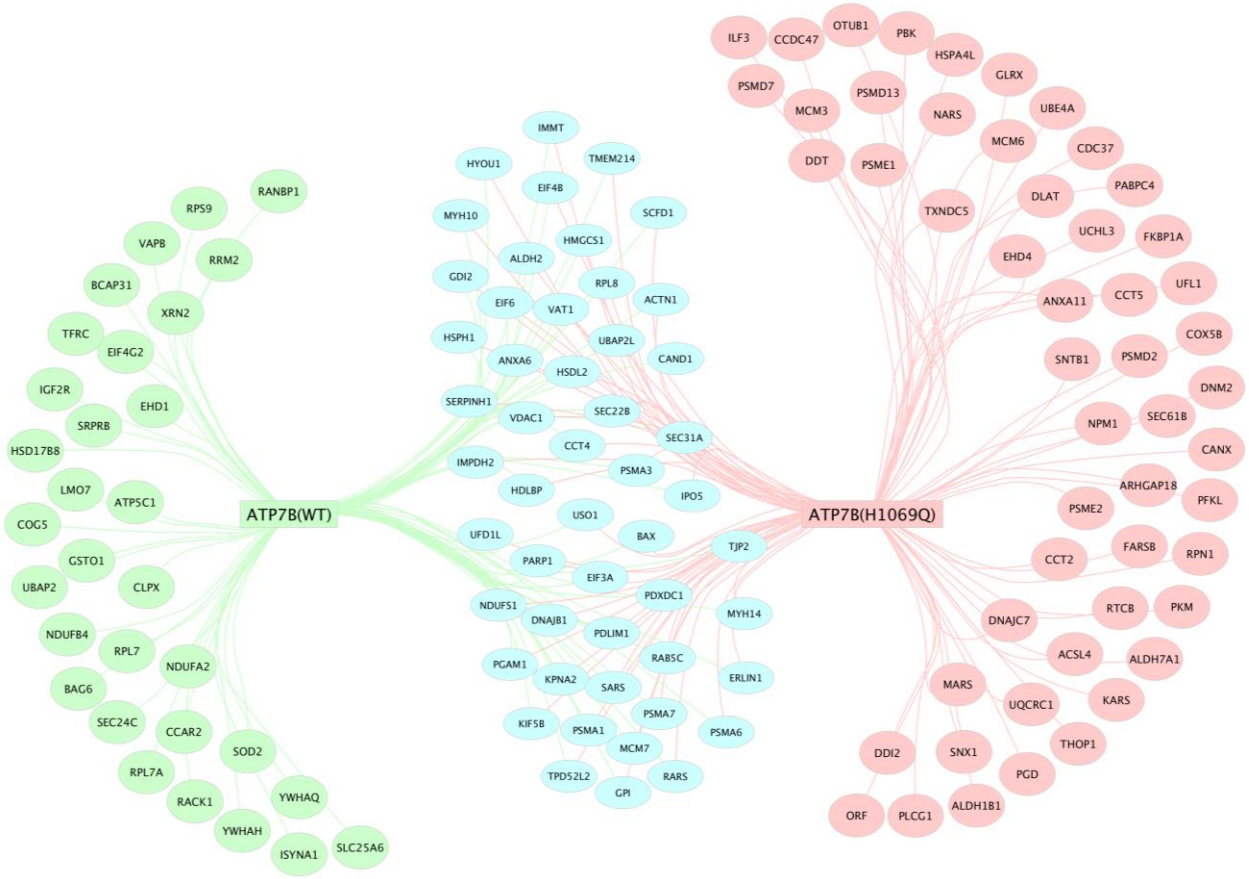


Figure 2.9: Graphical representation of protein interaction networks specific for ATP7B(WT) in green, ATP7B(H1069Q) in pink and those in common in light blue, obtained by Cytoscape software.

As shown in Figure 2.9, 31 and 51 protein partners have been identified as specific for ATP7B(WT) and ATP7B(H1069Q) respectively. As expected, a large number of interactors (52

proteins) have been identified as common putative partners of both ATP7B(WT) and ATP7B(H1069Q) proteins suggesting that the mutant protein retains roughly native threedimensional structure and then the ability to bind many interactors of wild type protein. Common interactors include essentially enzymes involved in lipid metabolism, glycolysis pathway and mitochondrial processes. The presence of the latter might be attributed to the existence of a smaller isoform of ATP7B (140kDa), localised in mitochondria, where it participates or assists to Cu- dependent processes taking place in this organelle [10].

Differential proteomic experiments carried out in Long-Evans Cinnamon (LEC) rats, an animal model of WD, demonstrated that the excess of Cu induces mitochondrial injury and carbohydrate metabolism disorders affecting the expression levels of several enzymes involved in these processes [11].

GO analysis of Biological processes indicated that specific binding partners of ATP7B(WT) were enriched in "*intracellular trafficking*" category (Figure 2.10). Among this latter, the *Vesicle-associated membrane protein-associated protein B/C (VAPB*), the *B-cell receptor-associated protein 31 (BCAP31)* and the *Protein transport protein Sec24C (SEC24C),* component of the COPII coat, that covers ER-derived vesicles, are involved in transport of new synthesized proteins from ER to the Golgi complex [12][13][14].

Also other putative ATP7B(WT) interactors, the *14-3-3 proteins,* such as *YWHAH and YWHAQ,* play a key role in intracellular protein transport. Recent studies demonstrated that 14-3-3 facilitates SAC1 protein transport from the endoplasmic reticulum by stimulating the packaging of SAC1 into COPII vesicles. This process involves the interaction of *SEC24C* with 14-3-3 [15]. All these data support the hypothesis that ATP7B translocates from ER to Golgi stacks through COPII vesicles traffic. The absence of these proteins among the variant interactors suggests that the traffic of mutant is impaired at level of the exit from ER and the entrance to Golgi.

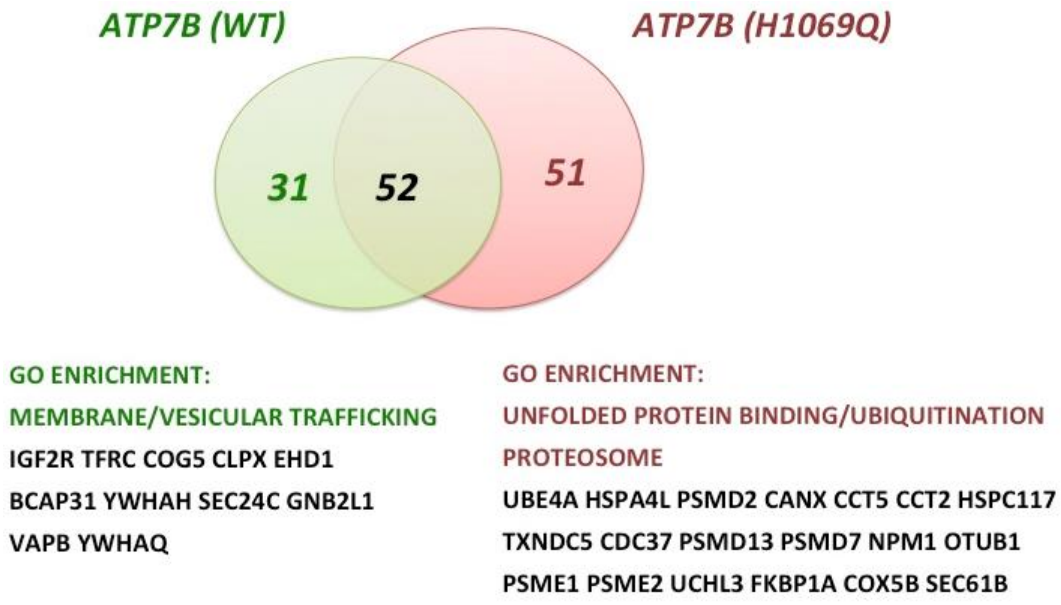


Figure 2.10: GO classification of ATP7B(WT) and ATP7B(H1069Q) specific protein partners.

Other proteins involved in traffic to and from endosome were identified exclusively in ATP7B(WT) interactome: the *Conserved oligomeric Golgi complex subunit 5 (COG5)*, an essential protein for vesicle transport within the Golgi apparatus and from endosomal compartments to the Golgi [16], the *Cation-independent mannose-6-phosphate receptor (IGF2R),* involved in sorting of lysosomal enzymes from Golgi compartment towards endosomes [17] and the *EH domain-containing protein 1 (EHD1),* a protein that acts in early endocytic membrane fusion and membrane trafficking of recycling endosomes [18]. These proteins might be associated to the ATP7B recycling from plasma membrane to Golgi passing through endosomes (Figure 2.11).

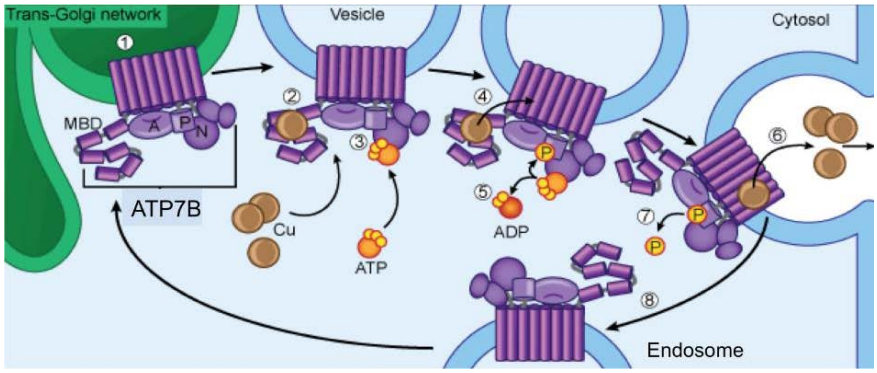


Figure 2.11: Representation of the endocytic trafficking of ATP7B. When cytosolic copper levels decrease, ATP7B is internalized and delivered to endosomes.

Among ATP7B(WT) specific interactors we also found *Receptor of activated protein C kinase 1,* called *RACK1 (GNB2L1)* that regulates protein traffic and transport from ER/Golgi compartment to cell surface and which is involved in the recruitment, assembly and/or regulation of a variety of signaling molecules [19][20][21][22]. RACK1 regulates various biological and physiological

functions as a scaffold/anchoring protein. Within this context, RACK1 is thought to promote efficient intracellular signaling through accurate and highly reproducible spatial and temporal compartmentalization of binding partners. RACK1 function is critical for many proteins physiological activities, as the correct localization of other plasma membrane proteins, such as ABCB4 [23] and the chloride channel CFTR [24]. RACK1 regulates the apical localization of the CFTR protein through the interaction with protein kinase C (PKCε) and Na^+/H^+ exchanger regulatory factor (NHERF1), two proteins essential for the activity of the ion channel. Figure 2.12 shows a hypothetical model of the RACK1 role for CFTR apical localization. RACK1 interacts with activated PKCε and NHERF1. The PKCε-RACK1-NHERF1 complex binds tubulin through other NHERF1 molecules. In this scenario, RACK1 works as an anchor to localize thin molecular machinery at apical region of epithelial cell. Down regulation of RACK1 reduces the apical localization of CFTR, indicating a role for a RACK1 in regulating the residency of CFTR at the cell surface [24].

Interestingly RACK1 is absent among ATP7B(H1069Q), suggesting that the interaction with ATP7B might play crucial for the correct localisation of the protein. It is possible that the absence of this interaction might depend either by the retention of ATP7B variant in the ER or by the conformation changes occurring in the ATP7B mutant. In order to deeply investigate the role of RACK1 in ATP7B trafficking, RACK1 siRNA experiment has been carried out. Silencing RACK1, ATP7B(WT) is retained in the ER thus confirming the importance of the interaction for the correct localization of the protein.

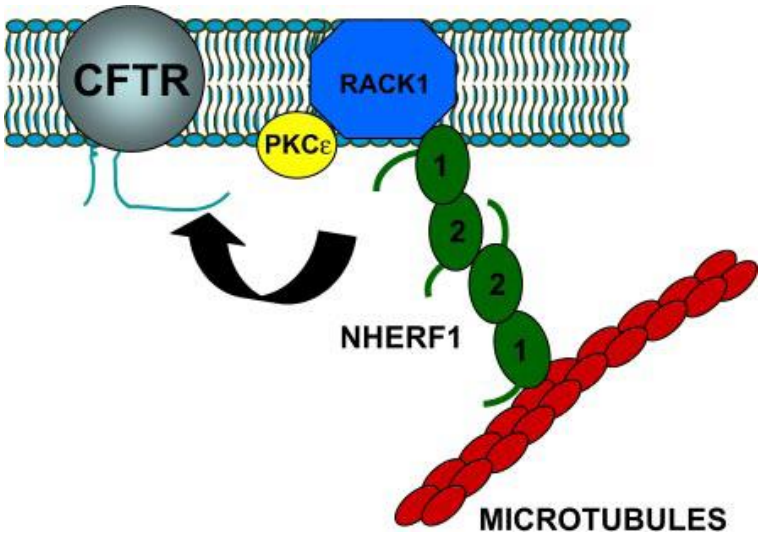


Figure 2.12: Hypothetical model of an epithelial RACK1 proteome for the apical localization of CFTR.

All these ATP7B(WT) specific interactors could assist the biological role of ATP7B in copper transport along the path from the ER/Golgi, where the protein is synthesized and subsequently uploaded copper, towards the plasma membrane where the excess copper is ejected from the cell.

In contrast, the ATP7B(H1069Q) interactome lacks almost at all of "*protein trafficking category*" while numerous proteins belonging to the "*unfolded protein binding*", "*proteasome*", and "*ubiquitination*" GO categories have been specifically recognized (Figure 2.9). Indeed a large number of specific mutant interacting proteins belong to molecular chaperones (*Heat shock 70 kDa protein 4L,* several *T-complex protein 1 subunits, DnaJ homolog subfamily C member 7, Hsp90 co-chaperone Cdc37*) indicating that the ATP7B(H1069Q) mutant might be partially misfolded thus to require the assistance of this specific protein family. The ER accumulation of misfolded protein activates a cell stress response as demonstrated by the presence of specific proteins (*Coiled-coil domain-containing protein 47, Serpin H1, Lymphokine-activated killer T-cell-originated protein kinase, Thioredoxin domain-containing protein 5 and Nucleophosmin*) that recognised and binds the ATP7B variant.

According to proteomics data, immuno electron microscopy (EM) showed that in presence of $CuSO_4$, ATP7B(WT) is localized in vesicles and at cell membrane. In contrast the mutant is retained in the ER (Figure 2.13).

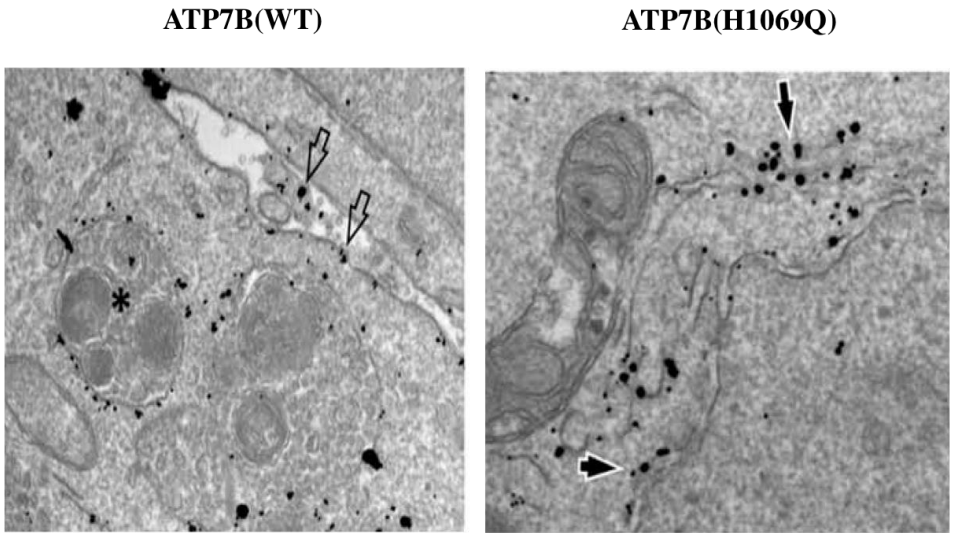


Figure 2.13: HepG2 cells were infected with Ad-ATP7B(WT)-GFP or Ad-ATP7B(H1069Q)-GFP, incubated for an additional 2 hours with 100 μM $CuSO_4$ and prepared for immuno-EM. Empty arrows indicate ATP7B(WT) signal at the PM, while asterisk label ATP7B-positive multivesicular body-like vesicles. Arrows indicate ATP7B(H1069Q) in the ER. Scale bars 260 nm.

Moreover the presence of several proteins involved in the process of ubiquitin conjugation to protein substrates (Ubiquitin-associated protein 2-like, Cullin-associated NEDD8-dissociated protein 1, Ubiquitin conjugation factor E4 A, Ubiquitin fusion degradation protein 1 homolog, E3 UFM1-protein ligase 1, Ubiquitin thioesterase OTUB1) and of a large number of proteasome subunits (alpha subunits and non-ATPase regulatory subunits), suggest that the ATP7B mutant recognized as defective by the ER quality control, is strongly addressed to proteasomal degradation.

Notably, many non-ATPase regulatory subunits interacting with ATP7B mutant are associated to MAPK activating pathways. These latter have recently been shown to affect proteostasis of proteins such as cystic fibrosis transmembrane conductance regulator (CFTR) [25] and to regulate the expression of ER quality control/degradation genes [26].

Among mutant protein partners, p38 (MAPK14) and its activators PLCG1 and PBK were identified, indicating a specific activation of stress kinase pathways.

Proteomic data are coherent with a comparative transcriptomic analysis carried out by microarrays on mRNA extracted from cell expressing ATP7B(WT) and mutant that revealed significative differences in expressed genes. Gene ontology analysis of these data indicated enrichment in genes that regulate cell death and response to stress (Figure 2.14). Additional bioinformatics analysis with String tool revealed that more than 100 gene altered in presence of ATP7B(H1069Q) mutation form a network enriched in “response to stress” (Figure 2.14). Several genes in this network (SQSTM1-TXNIP-IL6R-FOSL1-MAP2K6-NET1) operate in signaling pathways mediated by two pro-apoptotic stress kinases: p38 MAPK and JNK. These data are in agreement with the results obtained from functional proteomic analyses.

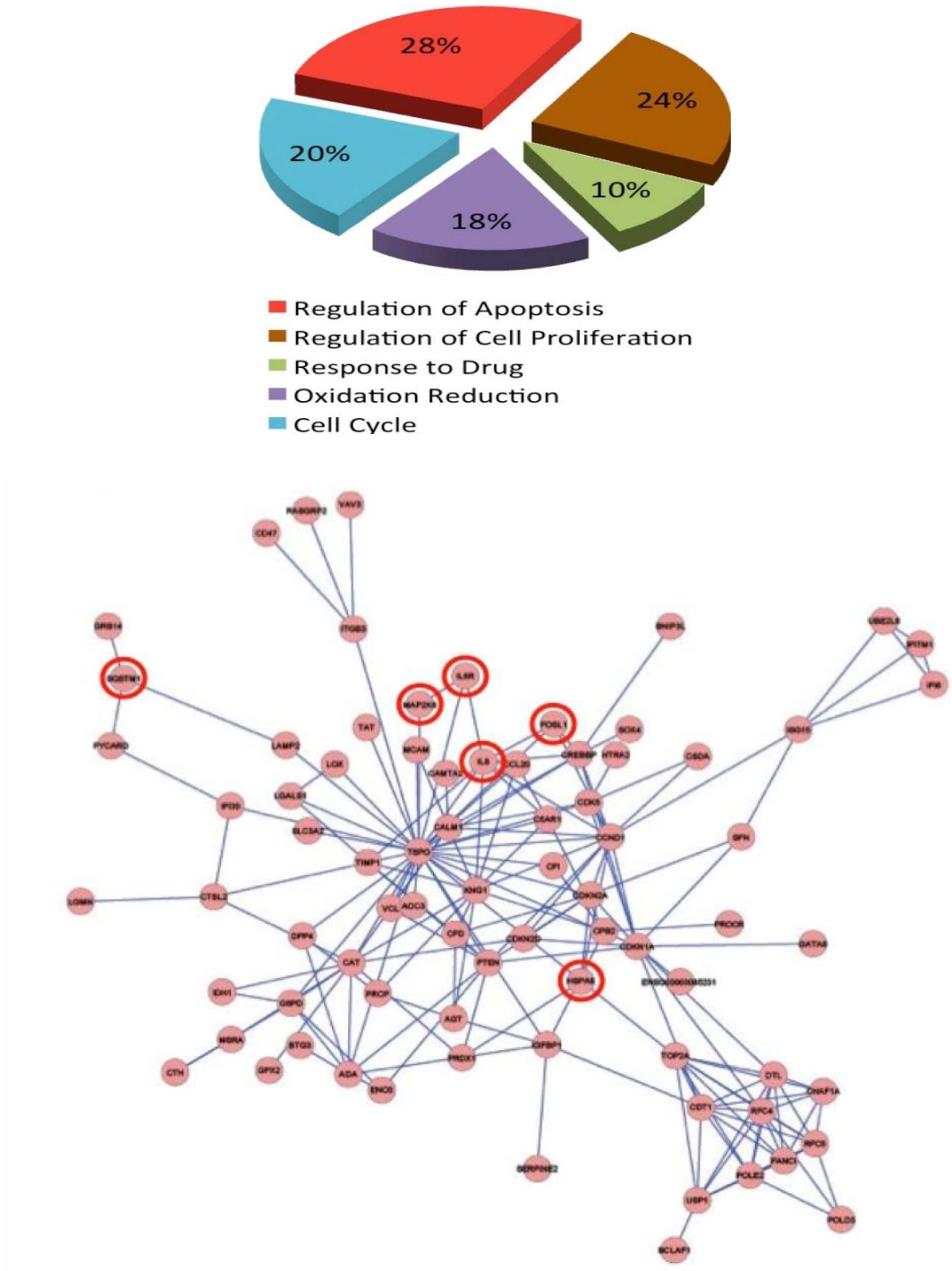


Figure 2.14: HepG2 cells were infected with ATP7B(WT)-GFP or ATP7B(H1069Q)-GFP and prepared for microarray analysis. Genes that were differently expressed in cells expressing ATP7B(H1069Q) were analyzed for GO enrichment. The pie diagram (upper row) shows the GO categories that were enriched among the altered genes in ATP7B(H1069Q) expressing cells, as opposed to cells expressing ATP7B(WT). String tool showed that genes altered in presence of ATP7B-H1069Q mutation form a network enriched in "response to stress" (lower row). Circled genes (SQSTM1-TXNIP-IL6R-FOSL1-MAP2K6-NET1) operate in signaling pathways mediated by two pro-apoptotic stress kinases: p38 MAPK and JNK.

Moreover, a significant increase in p38 MAPK and JNK phosphorylation in cells expressing the mutant versus cell expressing ATP7B(WT) confirmed the occurrence of a specific activation of p38 MAPK and JNK pathways in pathological conditions, as shown by Western blot analysis (Figure 2.15).

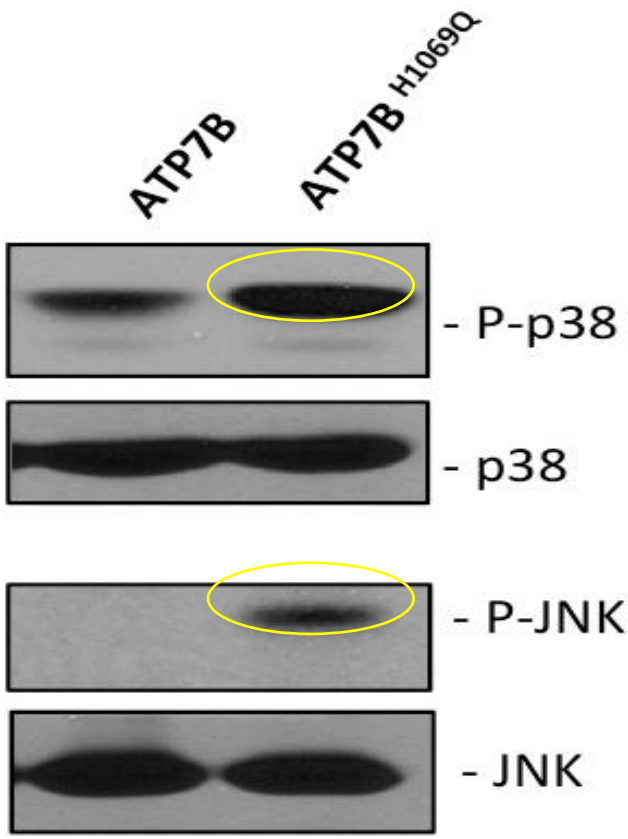


Figure 2.15: HepG2 cells were infected with ATP7BWT-GFP or ATP7BH1069Q-GFP and analyzed with Western blot. Phosphorylated forms of p38 or JNK increased in cells expressing the ATP7BH1069Q mutant, while overall amounts of p38 or JNK remained similar in wild type-expressing and mutant-expressing cells.

To test whether deactivating the p38 and JNK pathways facilitates ATP7B(H1069Q) export from the ER, the activity of p38 and JNK was suppressed first by using specific chemical inhibitors, SB202190 (SB90) and SP600125 (SP125), and then by RNA interference. Neither p38 and JNK inhibitor (added for 24 hours) affected ATP7B(WT), but both reduced ATP7B(H1069Q) within the ER and improved the delivery of the mutant to the Golgi (Figure 2.16 panel B). The RNA interference gave same results. In control experiments the ERK inhibitor failed to correct the mutant from the ER to the Golgi indicating that only p38 and JNK branches of the MAPK signaling cascade are involved in mutant retention in the ER.

Following inhibition of p38 and JNK, the mutant moves from ER to vesicles and PM in response to copper overload, showing a localization compatible with ATP7B(WT) (Figure 2.16). These data suggested that SB90 and SP125, respectively p38 and JNK inhibitors, might act as molecular chaperones promoting ATP7B mutant to escape from ER quality control, and leading to a right cell collocation of protein.

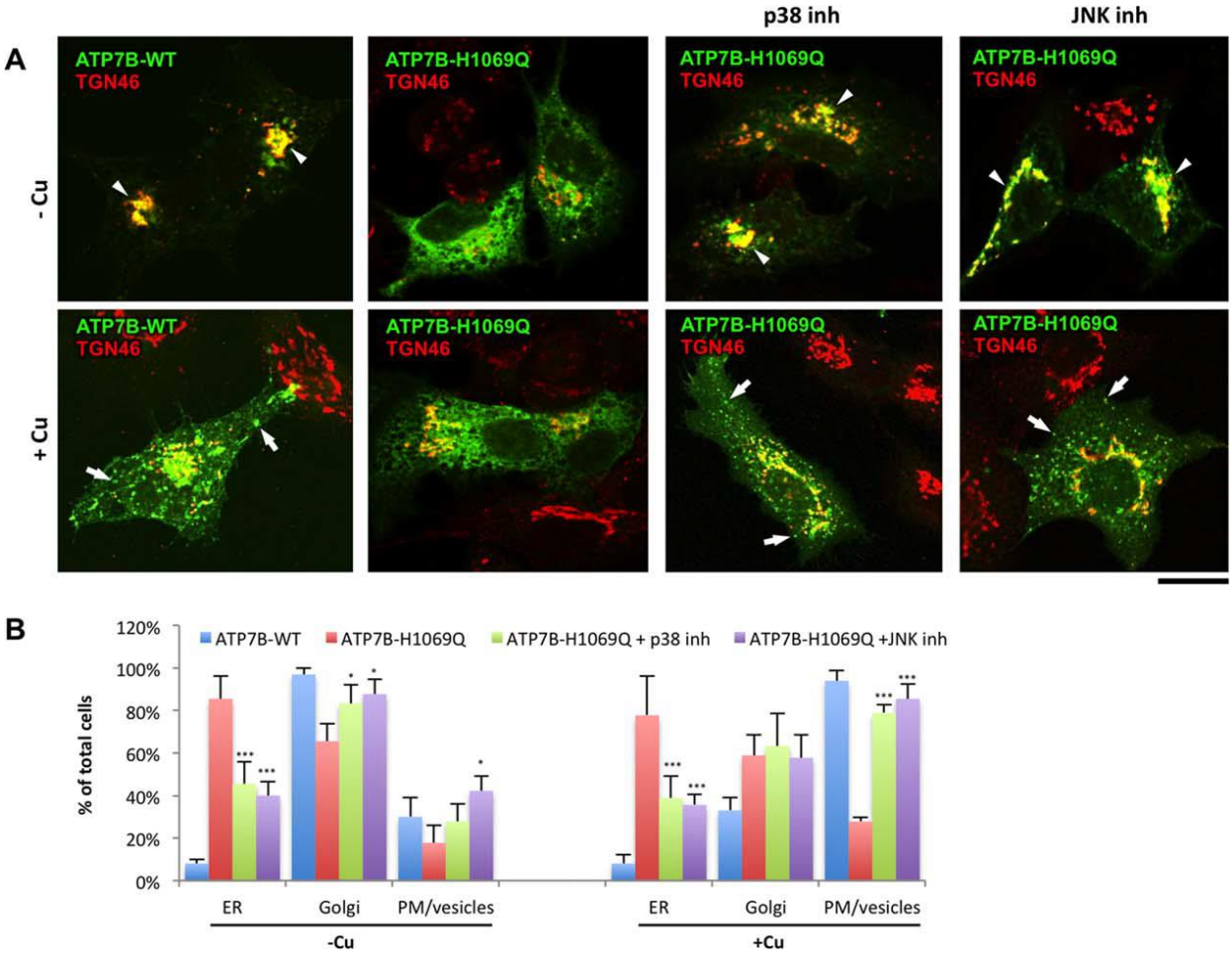


Figure 2.16: Suppression of p38 and JNK corrects localization and trafficking of ATP7B(H1069Q) mutant in hepatic cells. (A) HepG2 cells were infected with Ad-ATP7B(WT)-GFP or Ad-ATP7B(H1069Q)-GFP, incubated overnight with 200 μM BCS, and fixed or incubated for an additional 2 hours with 100 μM CuSO4. In response to Cu, ATP7B(WT) (left column) traffics from the Golgi (arrowheads in upper panels) to the PM and vesicles (arrows in lower panels), while ATP7B(H1069Q) (second column) is retained in the ER under both low-Cu and high-Cu conditions. p38 or JNK inhibitors were added to the cells 24 hours before fixation (as indicated in the corresponding panels). Fixed cells were then labelled for TGN46 and visualized under a confocal microscope. Both p38 and JNK inhibitors corrected ATP7B(H1069Q) from the ER to the Golgi (arrowheads in upper panels) under low-Cu conditions and to the PM and vesicles (arrows in lower panels) upon Cu exposure. (B) Cells were treated as in (A). The percentage of cells (average ± standard deviation, n =10 fields) with ATP7B signal in the ER, Golgi, or PM/vesicles was calculated. Both p38 and JNK inhibitors reduced the percentage of cells exhibiting ATP7B(H1069Q) in the ER independently from the Cu levels and increased the number of cells in which ATP7B was corrected to the Golgi under low-Cu conditions and to the PM and vesicles upon Cu stimulation.

Once the mutant has reached the membrane, its ability to excrete copper from the cell were also investigated by ICP-MS analysis. The quantization of intracellular Cu revealed that the treatment with p38 or JNK inhibitors caused a significant decay of the intracellular copper levels, confirming that the mutant following inhibition of p38 and JNK not only reaches the correct location to plasma membrane, but there it works properly eliminating the excess of copper from the cell (Figure 2.17) [27].

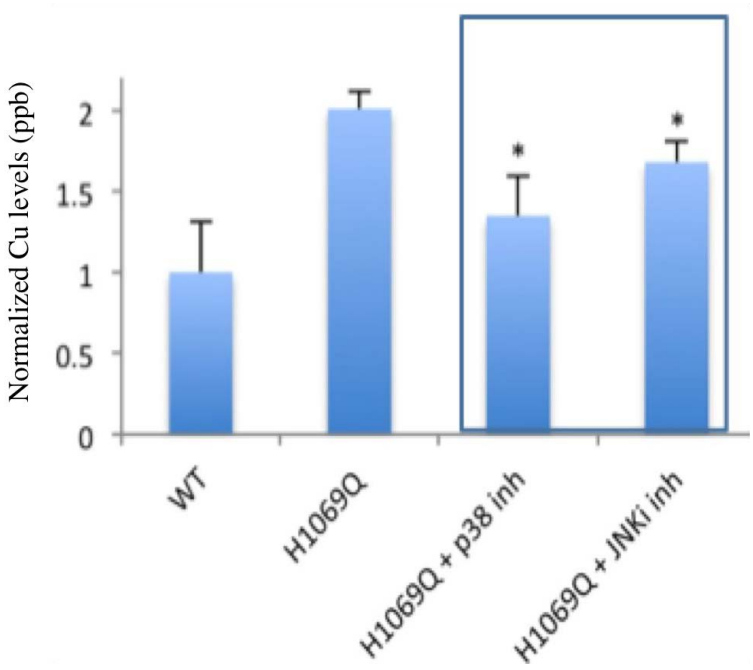


Figure 2.17: Inhibitors of p38 and JNK reduced Cu levels in cells expressing ATP7BH1069Q mutant. HepG2 cells were infected with Ad-ATP7BWT-GFP or Ad-ATP7BH1069Q-GFP and incubated with p38 or JNK inhibitor. In presence of CuSO4. The cells were examined by ICP-MS which revealed an increase in normalized intracellular Cu levels (average ± standard deviation, n=3 experiments) in ATP7B(H1069Q) expressing cells. Both p38 and JNK inhibitors reduced Cu levels in ATP7B(H1069Q) expressing cells.

2.5 Conclusions

In conclusion, proteomic and transcriptomic analyses revealed that the molecular mechanisms causative of Wilson disease triggered by the presence of ATP7B(H1069Q) is the retention of mutant protein in the ER compartment, due by its partial misfolding. This conformational change affects interactions occurring between ATP7B and several proteins controlling protein trafficking and translocation from ER to PM, which is impaired in the presence of mutant. On the other hand, the conformational instability of ATP7B(H1069Q) induces a strong stress response in the cell, with the activation of both protein degradation and apoptotic pathways.

In particular, the activation of p38 and JNK pathways has been associated to the disease onset and development [28]. Surprisingly, by inhibiting these two pathways, the mutant is redirect out from the ER to the secretory pathway, subtracting the protein to degradation and leading to the restoring the right copper level in the cell.

All together, the results of present research suggest that p38 and JNK might represent two new attractive targets for novel therapeutic strategies in the cure of Wilson disease's patients.

2.6 References

[1] Lutsenko S, Barnes NL, Bartee MY, Dmitriev OY, Function and Regulation of Human Copper-Transporting ATPases. *Physiol Reviews, Vol. 87 (3): 1011-46, 2007.*

[2] La Fontaine S, Mercer JF, Trafficking of the copper-ATPases, ATP7A and ATP7B: Role in copper homeostasis. *Archives of Biochemistry and Biophysics, Vol. 463: 149–67, 2007.*

[3] de Bie P, Muller P, Wijmenga C, Klomp LW, Molecular pathogenesis of Wilson and Menkes disease: correlation of mutations with molecular defects and disease phenotypes. *Journal of Medical Genetics, Vol. 44: 673-88, 2007.*

[4] Lee J, Peña MM, Nose Y, Thiele DJ, Biochemical Characterization of the Human Copper Transporter Ctr1. *The Journal of Biological Chemistry, Vol. 277 (6): 4380-7, 2001.*

[5] Veldhuis NA, Gaeth AP, Pearson RB, Gabriel K, Camakaris J, The multi-layered regulation of copper translocating P-type ATPases. *Biometals, Vol. 22: 177–90, 2009.*

[6] Polishchuk EV, Concilli M, Iacobacci S, Chesi G, Pastore N, Piccolo P, Paladino S, Baldantoni D, van IJzendoorn SC, Chan J, Chang CJ, Amoresano A, Pane F, Pucci P, Tarallo A, Parenti G, Brunetti-Pierri N, Settembre C, Ballabio A, Polishchuk RS, Wilson Disease Protein ATP7B Utilizes Lysosomal Exocytosis to Maintain Copper Homeostasis. *Developmental Cell, Vol. 29: 686–700, 2014.*

[7] Payne AS, Kelly EJ, Gitlin JD, Functional expression of the Wilson disease protein reveals mislocalization and impaired copper-dependent trafficking of the common H1069Q mutation. *Proceedings of the National Academy of Sciences of the United States of America Vol. 95: 10854–10859, 1998*

[8] Huster D1, Kühne A, Bhattacharjee A, Raines L, Jantsch V, Noe J, Schirrmeister W, Sommerer I, Sabri O, Berr F, Mössner J, Stieger B, Caca K, Lutsenko S, Diverse Functional Properties of Wilson Disease ATP7B Variants. *Gastroenterology, Vol. 142 (4): 947-56, 2012.*

[9] Gitlin JD, Wilson Disease. *Gastroenterology, Vol. 125 (6): 1868-77, 2003.*

[10] Lutsenko S, Cooper MJ, Localization of the Wilson's disease protein product to mitochondria. *Proceedings of the National Academy of Sciences of the United States of America, Vol. 95 (11): 6004-9, 1998.*

[11] Beom HL, Jae-Min K, Sun HH, Joo HM, Jihun K, Joo HK, Hye YJ, Gu-Hwan K, Jin-Ho C, Han-Wook Y, Proteomic analysis of the hepatic tissue of Long–Evans Cinnamon (LEC) rats according to the natural course of Wilson disease. *Proteomics, 11(18), 3698–3705, 2011.*

[12] Prosser DC, Tran D, Gougeon PY, Verly C, Ngsee JK, FFAT rescues VAPA-mediated inhibition of ER-to- Golgi transport and VAPB-mediated ER aggregation. *Journal of Cell Science, Vol. 121: 3052-61, 2008.*

[13] Annaert WG1, Becker B, Kistner U, Reth M, Jahn R, Export of Cellubrevin from the Endoplasmic Reticulum Is Controlled by BAP31 *The Journal of Cell Biology,Vol. 139 (6): 1397–1410, 1997*

[14] Pagano A, Letourneur F, Garcia-Estefania D, Carpentier JL, Orci L, Paccaud JP, *Sec24* Proteins and Sorting at the Endoplasmic Reticulum. *The Journal Of Biological Chemistry, Vol. 274 (12): 7833–40, 1999.*

[15] Bajaj Pahuja K, Wang J, Blagoveshchenskaya A, Lim L, Madhusudhan MS, Mayinger P, Schekman R, Phosphoregulatory protein 14-3-3 facilitates SAC1 transport from the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America, Vol. 112 (25): E3199-206, 2015.*

[16] Ha JY, Pokrovskaya ID, Climer LK, Shimamura GR, Kudlyk T, Jeffrey PD, Lupashin VV, Hughson FM, Cog5–Cog7 crystal structure reveals interactions essential for the function of a multisubunit tethering complex. *Proceedings of the National Academy of Sciences of the United States of America, 111 (44): 15762-7, 2014.*

[17] Braulke T, Bonifacino JS, Sorting of lysosomal proteins. *Biochimica et Biophysica Acta, Vol. 1793: 605–14, 2009.*

[18] Naslavsky N, Boehm M, Backlund PS Jr, Caplan S, Rabenosyn-5 and EHD1 Interact and Sequentially Regulate Protein Recycling to the Plasma Membrane. *Molecular biology of the cell, Vol. 15 (5): 2410-22, 2004.*

[19] Parent A, Laroche G, Hamelin E, Parent JL, RACK1 regulates the cell surface expression of the G protein-coupled receptor for thromboxane A(2). *Traffic, Vol. 9 (3): 394-407, 2008.*

[20] Reiner CL, McCullar JS, Kow RL, Le JH, Goodlett DR, Nathanson NM, RACK1 associates with muscarinic receptors and regulates M(2) receptor trafficking. *PLoS One, Vol. 5 (10): e13517, 2010.*

[21] Ikebuchi Y, Ito K, Takada T, Anzai N, Kanai Y, Suzuki H, Receptor for activated C-kinase 1 regulates the cell surface expression and function of ATP binding cassette G2. *Drug Metabolism and Disposition, Vol. 38 (12): 2320-8, 2010.*

[22] Li S, Esterberg R, Lachance V, Ren D, Radde-Gallwitz K , Chi F, Parent JL, Fritz A, Chen P, Rack1 is required for Vangl2 membrane localization and planar cell polarity signalling while attenuating canonical Wnt activity. *Proceedings of the National Academy of Sciences of the United States of America, Vol. 108 (6): 2264-9, 2011.*

[23] Ikebuchi Y, Takada T, Ito K, Yoshikado T, Anzai N, Kanai Y, Suzuki H, Receptor for activated C-kinase 1 regulates the cellular localization and function of ABCB4. *Hepatology Research, Vol. 39: 1091–1107, 2009*

[24] Auerbach M, Liedtke CM, Role of the scaffold protein RACK1 in apical expression of CFTR. *American Journal of Physiology - Cell Physiology, Vol. 293(1): C294-304, 2007.*

[25] Hegde RN, Parashuraman S, Iorio F, Ciciriello F, Capuani F, Carissimo A, Carrella D, Belcastro V, Subramanian A, Bounti L, Persico M, Carlile G, Galietta L, Thomas DY, Di Bernardo D, Luini A, Unravelling druggable signalling networks that control F508del-CFTR proteostasis. *eLIFE, Vol. 4: e10365, 2015.*

[26] Darling NJ, Cook SJ, The role of MAPK signalling pathways in the response to endoplasmic reticulum stress, *Biochimica et Biophysica Acta, Vol. 1843 (10): 2150-63, 2014.*

[27] Chesi G, Hegde RN, Iacobacci S, Concilli M, Parashuraman S, Festa BP, Polishchuk EV, Di Tullio G, Carissimo A, Montefusco S, Canetti D, Monti M, Amoresano A, Pucci P, van de Sluis B, Lutsenko S, Luini A, Polishchuk RS, Identification of p38 MAPK and JNK as new targets for correction of Wilson disease-causing ATP7B mutants. *Hepatology, Vol. 63 (6): 1842-59, 2016.*

[28] Kadowaki S, Meguro S, Imaizumi Y,1 Sakai H, Endoh D, Hayashi M, Role of p38 Mapk in Development of Acute Hepatic Injury in Long-Evans Cinnamon (LEC) Rats, an Animal Model of Human Wilson's Disease. *The Journal of Veterinary Medical Science, 75(12): 1551–1556, 2013.*

Chapter 3 – Investigation of Molecular Mechanisms Impaired in Pompe Disease

3.1 Introduction

3.1.1 Lysosomes and Lysosomal Enzymes

Lysosomes are subcellular organelles that break down and recycle a wide range of complex substrates, including glycosaminoglycans, sphingolipids, glycogen, and several proteins. Their catabolic function is performed through the concerted action of approximately 60 different acidic hydrolases belonging to different protein families, such as glucosidases, sulfatases, peptidases, phosphatases, lipases, and nucleases. Lysosomal catabolic function depends on direct interaction between enzymes and substrates in the lysosomal lumen [1]. Lysosomal enzymes are synthesized in the endoplasmic reticulum (ER), where they adopt their native conformations and reach the lysosomes via specialized pathways. The most of these hydrolases are targeted to lysosomes through the addition of mannose-6-phosphate (M6P) residues onto the oligosaccharide moieties of the enzymes, a modification that occurs in the late Golgi compartment, and through the M6P receptor (MPR) pathway [2]. Some specific enzymes do not depend on MPR for lysosomal delivery, such as β-glucocerebrosidase, which is transported to lysosomes by lysosomal integral membrane protein 2 (LIMP-2) [3].

Substrates are translocated to lysosomes through different paths: specific endocytic mechanisms, such as phagocytosis, macropinocytosis, clathrin-mediated endocytosis, caveolin-mediated endocytosis, and clathrin and caveolin-independent endocytosis, are preferentially exploited according to the nature of the cargo. Intracellular materials are transported to the lysosomes mainly through autophagy, a process that delivers cytoplasmic components and organelles to lysosomal degradation and recycling (Figure 3.1).

The biological role of lysosomal hydrolases is crucial for the cell survival: the total or partial lack of their enzymatic activity causes intra-lysosomal accumulation of undegraded substrates and abnormalities of several cellular pathways. Mutations in the genes that encode lysosomal enzymes are responsible for genetic diseases, called *Lysosomal Storage Diseases* [1].

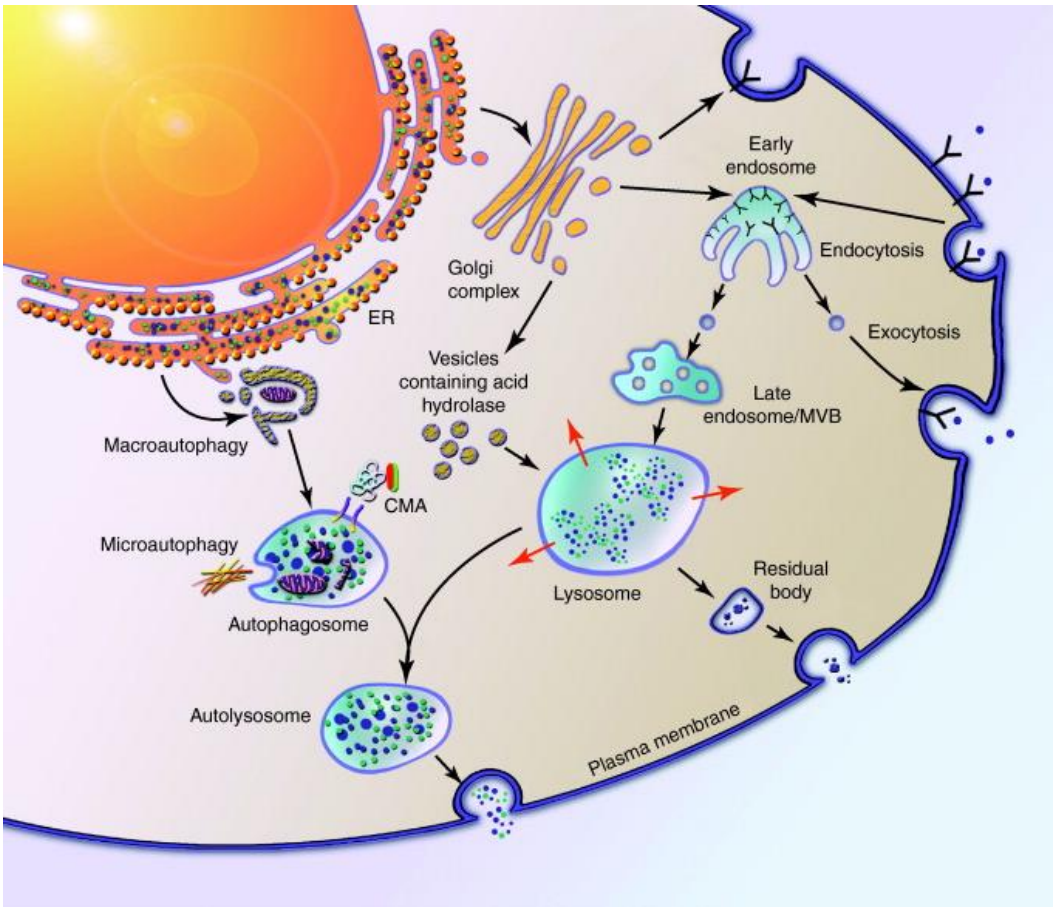


Figure 3.1: Overview of cellular trafficking to the lysosome and autolysosome. Lysosomal proteins reach the lysosome directly, via the trans-Golgi network (TGN), or indirectly by way of the plasma membrane. In the direct pathway, upon trafficking to the TGN, some oligosaccharide side chains acquire the mannose-6-phosphate (M6P) recognition marker. Enzymes with this modification are recognized by cation-dependent and cation-independent M6P receptors (M6PRs) in the TGN, mediating their sorting into the endosome–lysosome pathway. M6P-tagged molecules that escape M6PR recognition in the TGN are transported to the cell surface and secreted into the extracellular fluid. Secreted M6P-labeled proteins are recognized by M6PR on the extracellular plasma membrane and are delivered to the lysosome by receptor-mediated endocytosis.

3.1.2 Lysosomal Storage Diseases

Lysosomal storage diseases (LSDs) are a group of more than 50 rare inherited metabolic disorders caused by mutations in genes encoding soluble acidic hydrolases, integral membrane proteins, activator and transporter proteins that are essential for lysosomal functions. These genetic mutations result in the impairment of one or more functions of lysosomes. Indeed, the missense mutations may cause protein retention in the endoplasmic reticulum, often followed by degradation; analogously, other alterations in post-translational modification, such as an abnormal glycosylation might induce mistrafficking of the mutant protein through the Golgi apparatus, trans-Golgi network (TGN), and to the lysosomes [1][4]. Functional deficiencies in lysosomal proteins trigger pathogenetic cascades that lead to intra-lysosomal accumulation of undegraded substrates in multiple tissues and organs (Figure 3.2). The accumulation of specific substrates in lysosomal compartment affects several cellular functions with mechanisms that include the activation of

receptors by a non-physiological over accumulation of ligands, the modulation of receptor responses and signal transduction cascades, the activation of an inflammatory response, the altered intracellular traffic of vesicles, of membranes and membrane-bound protein, the alteration of related to the autophagy [5].

The clinical consequences affect different organs with visceral, ocular, haematological, skeletal, or neurological manifestations, and with phenotypes overlapping among different disorders. Symptoms may emerge at variable ages, in some cases starting in utero or during the new-born period, or becoming evident in late adulthood. In general, the diseases progress and evolve over the time.

The lysosomal storage diseases are generally classified by the nature of the accumulated substrates. Some examples of LSDs are listed in Table 3.1 [1][5].

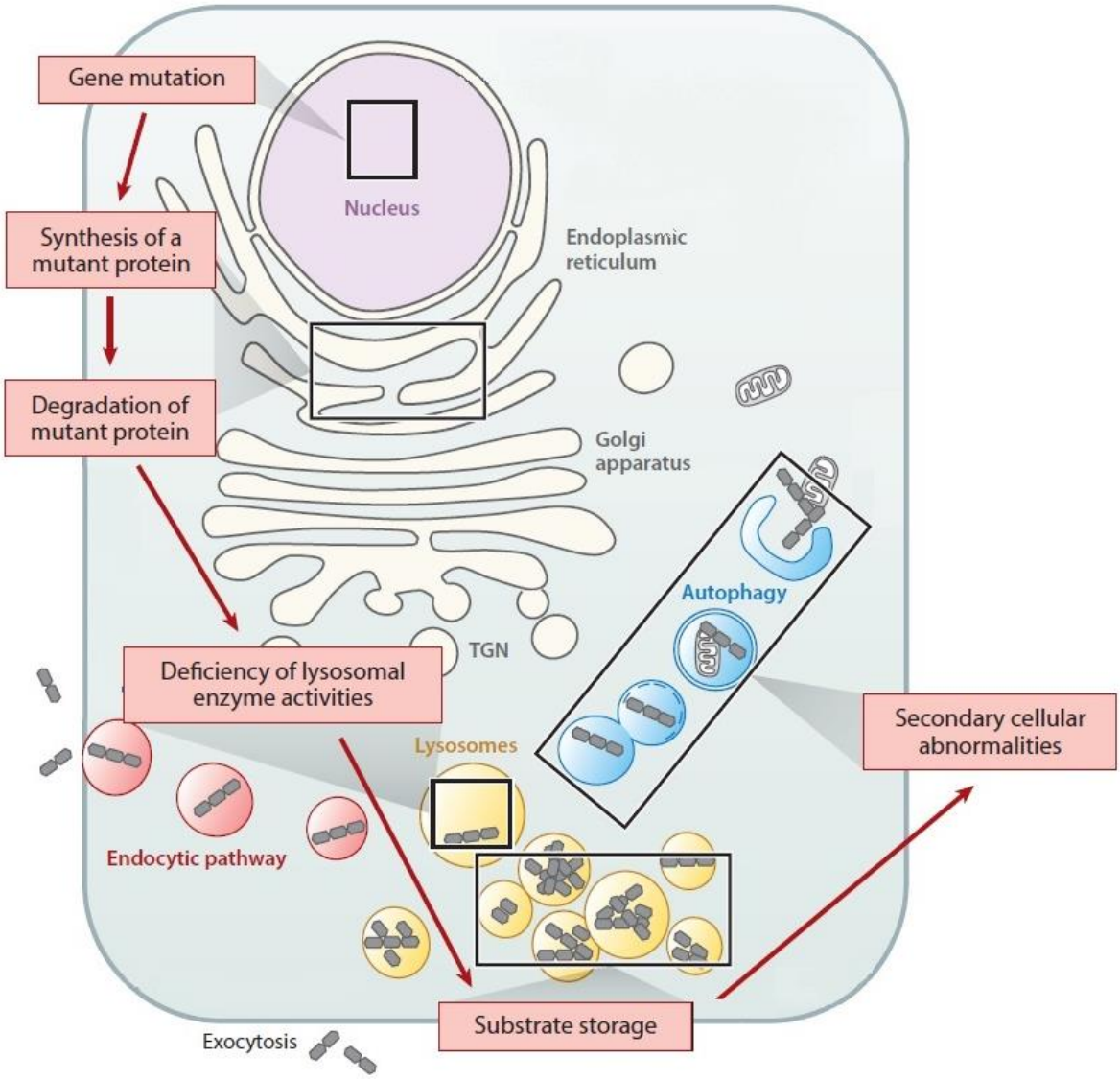


Figure 3.2: Cascade of events underlying the pathogenesis of LSDs.

Table 3.1: Some examples of Lysosomal Storage Diseases.

Disease	Defective protein	Storage materials	Major symptoms
Pompe	Acid α-glucosidase	Glycogen	Cardiomyopathy, cardiomegaly, severe hypotonia, limb-girdle weakness, muscle weakness, respiratory dysfunction
Fabry	α-Galactosidase A	Globotriaosylceramide, Galabiosylceramide, Globotriaosylsphingosine	Painful paresthesias, angiokeratoma, renal disease, cardiomyopathy and stroke. There is no neuronal involvement.
Multiple Sulfatase Deficiency	Formylglycine-generating enzyme	Heparan sulphate, Dermatan sulphate, Chondroitin-4- and -6-sulphates, Sulpholipids	Neurologic symptoms prevail. In addition, patients have attenuated signs of mucopolysaccharidosis like hepatomegaly, dysostosis multiplex and coarse facial features.
Gaucher	Acid β-Glucosidase	Glucosylceramide, Glucosylsphingosine	Type 1: Dysfunction of liver, spleen and bone marrow. Symptoms involve hepatosplenomegaly, thrombocytopenia, skeletal deformations and bone fractures. There is no nervous system involvement. Type 2: In addition to visceral symptoms, patients have severe neurologic involvement. Brain biopsies show extensive neuronal death. Type 3: intermediate symptoms between type 1 and 2.
Niemann–Pick A and B	Sphingomyelinase	Sphingomyelin, Cholesterol, Bismonoacylglycerophosphate, Glucosylceramide, Lactosylceramide, Globotriaosylceramide, Globotetraosylceramide	Hepatosplenomegaly and rapid progressive neurodegeneration.

3.1.3 Pompe Disease

Pompe Disease (PD), also known as glycogen type II storage disease (GSDII) or acid maltase deficiency, is a rare disorder of glycogen metabolism with an estimated frequency of 1 in 40,000 live births. PD is an autosomal recessive disease, occurring when an individual inherits two alleles, one from each parent, mutated in the gene encoding acid α-1,4-glucosidase (GAA), which maps on the long arm of chromosome 17 (location q25.2-q25.3) [6].

Lysosomal acid α-glucosidase (GAA, acid maltase) is an exo- 1,4- and -1,6-α-glucosidase that hydrolyzes glycogen to glucose in lysosome (Figure 3.3).

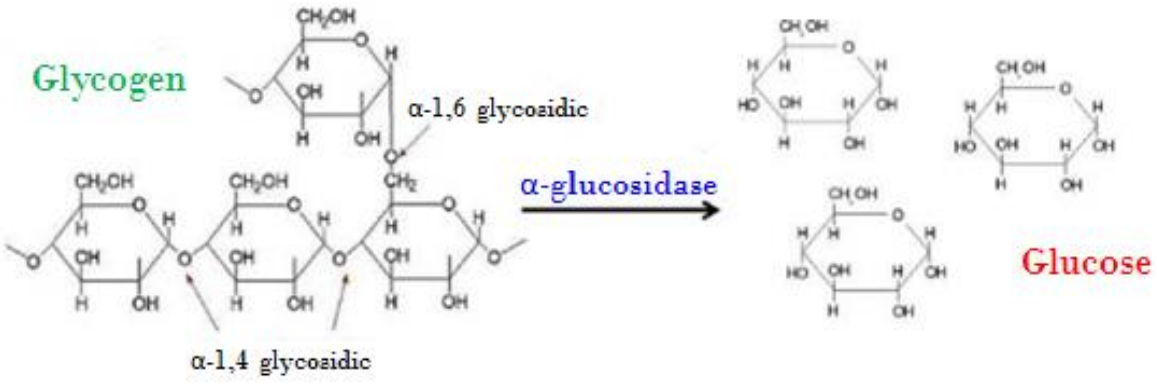


Figure 3.3: Enzymatic degradation of glycosidic bonds by acid α- glucosidase.

More than 450 mutations in *GAA* have been reported in individuals affected by PD, leading to both a partial deficiency of the enzymatic function or to the complete absence of the active protein [7]. Furthermore, the GAA protein undergoes complex post-translational modifications trafficking from ER towards lysosomes, including glycosylation and proteolytic maturation (Figure 3.5) [8]. In several PD forms, many abnormalities might occur along this pathway, thus affecting the availability of an opportune amount of active enzyme at level of lysosome [1][7].

GAA is synthesized as precursor with an amino-terminal signal peptide for co-translational transport into the endoplasmic reticulum lumen, where it is *N*-glycosylated at seven glycosylation sites, resulting in a glycosylated precursor of 110-kDa. The 110-kDa precursor, containing N-linked carbohydrate groups modified with mannose-6-phosphate, after the transport through the Golgi complex and targeting to the endosome/lysosome, is proteolytically processed at the amino terminus, resulting in a 95-kDa intermediate with a sequence beginning at amino acid 122. Than 95-kDa intermediate is proteolytically processed to a 76-kDa form, at the carboxyterminal. The 76-kDa form is then proteolytically processed at the amino terminus at amino acid 204 to give the 70-kDa mature form (Figure 3.4).

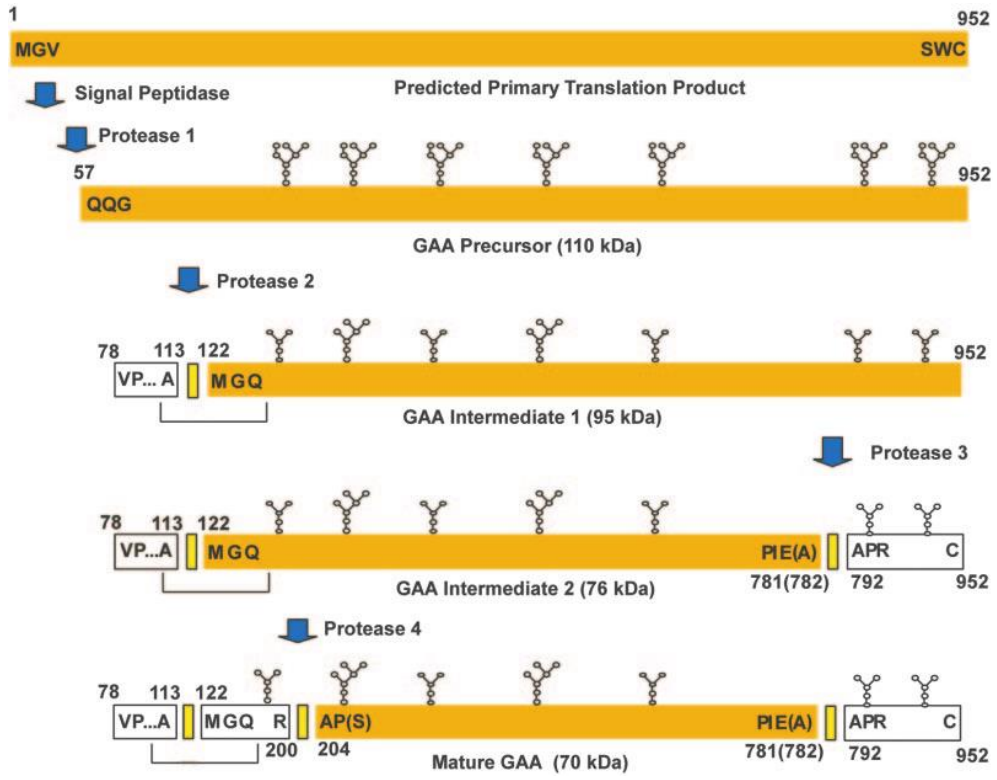


Figure 3.4: Model for the maturation of GAA protein.

Proteolytic processing appears to be required for GAA enzymatic activity. Indeed, the 76/70-kDa species have a 7-10 times higher affinity for glycogen compared to the 110-kDa precursor [8]. The identities of the proteases involved in the maturation of GAA are still unknown.

Genetic mutation affecting the activity of acid alpha-glucosidase leads to metabolic disorder, defined Pompe Disease, responsible for abnormal intra-lysosomal glycogen storage with subsequent extensive damage of muscles.

The age of onset, the organ involvement, and the rate of disease progression are heterogeneous. According to the clinical manifestations, PD is divided into two wide categories: infantile and late-onset forms. The infantile form includes classic and non-classic variants. Classic infantile disease is a rapidly progressive disorder manifesting with cardiomegaly, hypotonia, or mild hepatomegaly that leads to death for cardiorespiratory failure within the first year of life. In the non-classic infantile form, the symptoms appear early, but the disease shows a slower progression and less severe cardiomyopathy [9]. All the infantile forms show a complete or near-complete GAA deficiency [10]. The late-onset forms include childhood, juvenile, and adult-onset variants. Both childhood and juvenile diseases appear any time after infancy with mild or no cardiac involvement. In the adult form, the presence of a residual enzymatic activity leads to a late-onset, occurring between the second and sixth decade of life. A slow progressive skeletal muscle myopathy constitutes the main manifestation of the late-onset form, and it is often associated to a proximal

muscle weakness with a massive involvement of respiratory muscles: indeed, respiratory failure is the principal cause of significant morbidity and mortality in this form.

In Pompe Disease, the glycogen accumulation in lysosomes promoted by muscle activity associated to the lack of the enzymatic degradation, triggers the breaking of lysosomes, with the consequent release of potentially toxic substances into the cytosol and the degenerations of skeletal muscle (Figure 3.5).

Unlike the other cells, in skeletal muscles of PD patients, the enlarged lysosomes are found in a limited inter-myofibrillar space, which facilitates the physical disruption of the lysosomal membrane, leading to the loss of myofibrillar material and the complete destruction of the fiber [11].

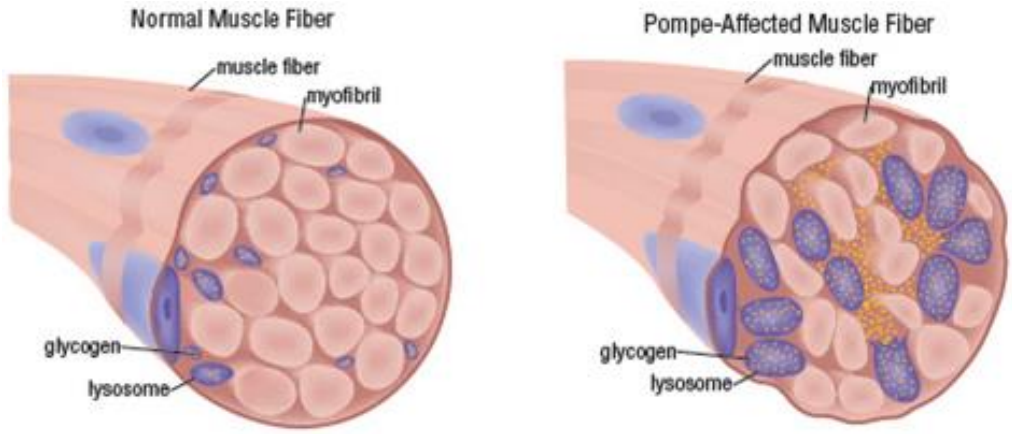


Figure 3.5: Representation of a normal muscle fiber (left) and a fiber muscle affected by PD (right).

3.1.4 Possible Therapies in LSDs

Until about two decades ago, the treatment of LSDs was based on multidisciplinary palliative therapies. In the last 25 years, the advances in knowledge of LSD pathophysiology and the improvement of highly innovative technologies have led to the development of several therapeutic approaches. Most of them have been designed to increase the enzymatic residual activity, such as: hematopoietic stem cell transplantation (HSCT), gene therapy (GT), pharmacological chaperone therapy (PCT), therapy based on substrate reduction (SRT) and enzyme replacement therapy (ERT), [1][12].

3.1.4.1 Hematopoietic Stem Cell Transplantation

Hematopoietic stem cell transplantation (HSCT) has been the first therapeutic approach for the treatment of LSDs in which hematopoietic stem cells derived from a healthy donor are used as therapeutic agents [12]. The bone marrow has been traditionally used as source of cells to be transplanted for this procedure. However, in recent years, many patients have been treated with

umbilical cord blood transplants from uncorrelated donors, thus achieving a quick and easy access to transplantation [12]. This approach has a dual effect: both repopulate specific tissues and secrete functional lysosomal hydrolases in the extracellular space and into blood circulation. The enzyme secreted by transplanted cells is taken up by the recipient host cells thus cross-correcting the enzymatic defect. This approach is not effective for all LSD forms [1].

3.1.4.2 Gene Therapy

Gene therapy (GT) is a therapeutic aimed to restore the defective enzyme activity by delivering a wild-type copy of the defective gene in patients' cells and tissues by using viral vectors as carriers [1]. LSDs are monogenic disorders and then constitute excellent candidates for gene therapy, because, in principle, it is possible to cure the disease by correcting the gene defect. In addition, small increases in enzymatic activity (<10%) may be sufficient to produce clinical benefit and phenotypic correction of LSDs.

The GT, when applicable, constitutes the best therapeutic approach, which allows the long-term expression of the native protein in patients affected by serious forms of disease, for which the alternative ERT would be a commercially not advantageous.

Although GT is a promising approach for LSDs treatment, the main limitation is the safety in the employment of viral vectors, for their potential pathogenic nature and the possibility to integrate into the host genome in uncontrolled way.

Besides the well-known technical difficulties in the GT set up, other strong limitations in the employment of GT concern: 1) the choice of the suitable vector to get the best possible correction in specific organs and tissues, 2) the doses to be used, 3) the risk of an unexpected immune reaction against the enzyme produced by GT, which may be recognized as an "exogenous" protein by immune system [12].

3.1.4.3 Enzyme Replacement Therapy

Enzyme replacement therapy (ERT) is considered the standard of care for several LSDs. This approach is based on periodic intravenous infusions of human recombinant enzyme, produced and purified on a large scale from various sources by recombinant DNA techniques. Once injected, the recombinant enzymes are internalized by cells and directed to the lysosomal compartment, where they have to compensate for the defective enzyme dysfunction.

ERT is an excellent example of how progress in the field of lysosomal biology stimulated the development of novel therapeutic approaches. The logic behind ERT evolved from pivotal studies on the molecular mechanisms implicated in the sorting of newly synthesized lysosomal enzymes

by way of the M6P and MPR pathways. These studies have shown that lysosomal hydrolases have on their structure some markers, namely mannose or mannose-6-phosphate, which they are recognized by specific receptors in the cell, which direct the enzyme to the lysosomal compartment, where it must perform its function. The mannose or mannose-6-phosphate receptors are also expressed on the plasma membrane. Thus, the recombinant lysosomal enzymes produced and purified in the laboratory can be "captured" by the cells, and following the path of the endocytic route, be properly transported to lysosomes (Figure 3.6) [13].

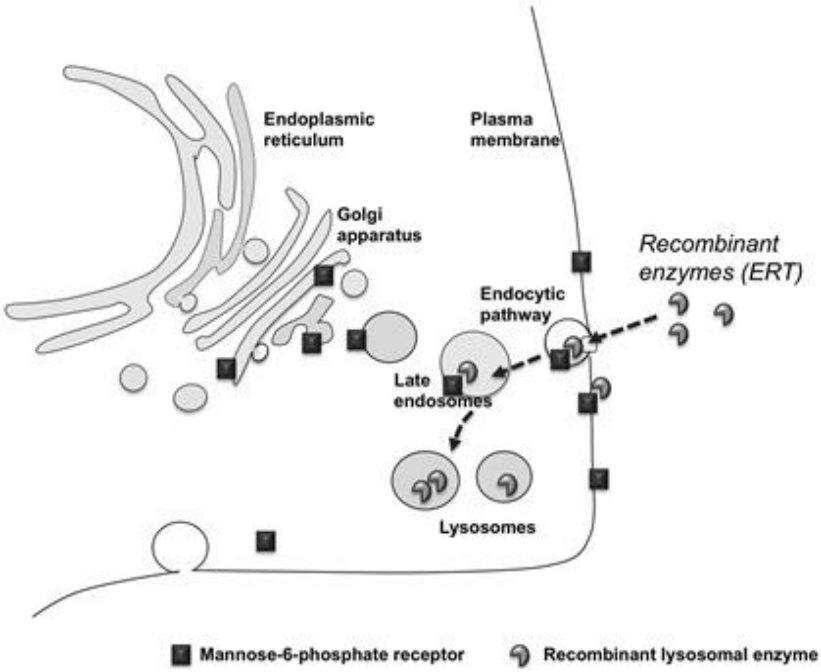


Figure 3.6: Recombinant enzymes are recognized by receptors on the cell surface and transported to lysosomes.

Since 2006, Pompe disease is treated with Myozyme (alglucosidase alfa). The human recombinant enzyme GAA (rhGAA) is produced biotechnologically and administered intravenously in the form of a precursor of 110 kDa containing groups of mannose-6-phosphate, which allow the enzyme to bind to the receptor on the cell surface.

The first clinical tests with rhGAA have shown the benefits of ERT in patients with Pompe disease, but the treatment is not a cure because only a few patients showed significant improvements in motor function. The effectiveness of treatment is high in the early stages of diagnosis.

3.1.4.4 Pharmacological Chaperone Therapy

Pharmacological Chaperone Therapy (PCT) is a therapy based on the employment of drugs developed to stabilise native conformation of proteins containing missense mutations that alter the native folding, and allowing them to escaping the degradation. Indeed, misfolded proteins recognized by the quality control systems of the endoplasmic reticulum (ER) might be retained in

the ER, or abnormally glycosylated and mistrafficked or degraded by proteasome. Therefore, in many diseases associated to misfolding, the loss of function is not often due to a reduced catalytic activity, but rather to aberrant protein degradation or to inappropriate trafficking. Following interacting with mutant proteins, molecular ligands with pharmacological chaperones activity, allow the stabilization of protein native conformation, enhance their stability, and restoring the correct trafficking. As a result, the mutant protein reaches the lysosomes and its enzymatic activity is partially rescued (Figure 3.7) [1].

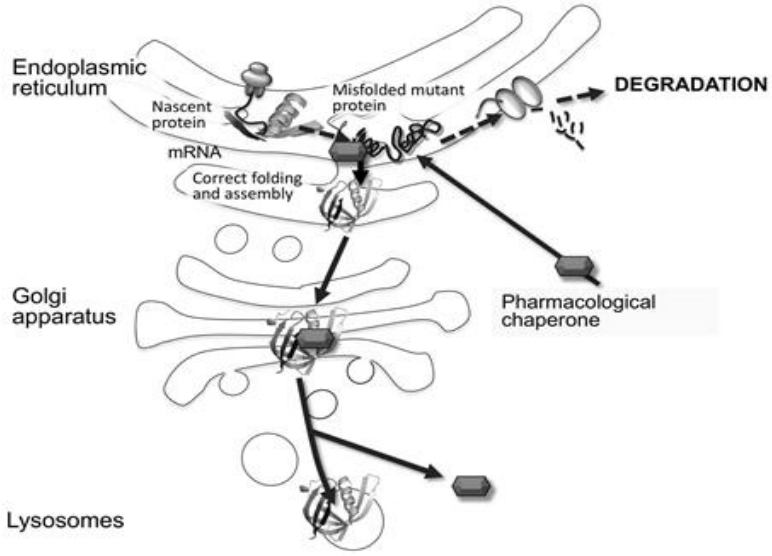


Figure 3.7: Missense mutations in lysosomal enzymes alter their three-dimensional conformation. Thus they are recognized and degraded by the cell. Pharmacological chaperones can interact physically with the mutant protein allowing their correct conformation and reaching the lysosomes.

At present, some iminosugars are used as pharmacological chaperones: 1-deoxynojirimycin (DNJ), its derivative N-alkylated butyl-deoxynojirimycin (NB-DNJ) (Figure 3.8) [14].

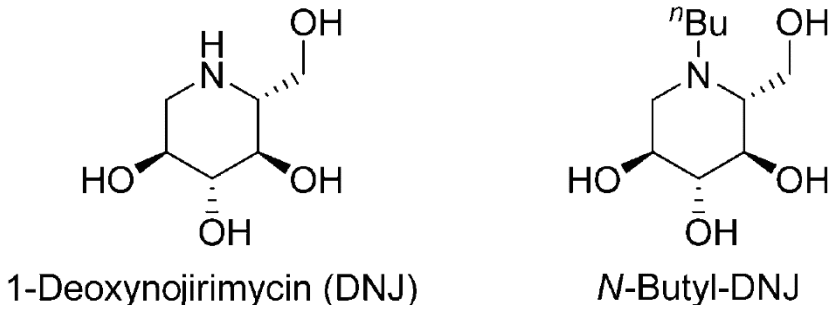


Figure 3.8: Known iminosugars as pharmacological chaperones.

Molecular chaperones have several advantages for LSDs treatment: 1) they can be administrated orally, allowing a non-invasive treatment; 2) they are not immunogenic and do not need to be taken up by cells through receptor mannose-6-phosphate, whose function can be compromise in some of LSDs. In addition, these small molecules diffuse freely across cellular membranes and reach different tissues and systems in therapeutic concentrations.

The real limitations in this therapeutic approach concern its feasibility, because the treatment is effective only in patients carrying few mutations, mostly located in specific structural domains of the enzyme [12].

In addition to PCT, other approaches based on the employment of *proteostasis regulators* have been proposed for the LSDs treatment. This small-molecules class acts at different levels on the so-called cell proteostasis, a complex network of processes that controls protein synthesis, folding, trafficking, aggregation, and degradation. This approach, however, is not so employed because it is still at early stage of testing [1].

3.1.4.5 Substrate Reduction Therapy

Substrate reduction therapy (SRT) is based on inhibition of specific steps in the biosynthetic pathways of substrates to restore the equilibrium between the synthesis of substrates and their degradation by lysosomal enzymes. This task is generally accomplished by using small-molecule enzyme inhibitors involved in substrate biosynthesis [1].

3.1.5 Aim of the Project

Although the biochemistry and genetics of PD are well characterized, the cellular mechanisms involving GAA protein and how glycogen storage leads to muscles atrophy remain poorly understood.

This project was focused on the investigation of the intracellular pathways controlling GAA traffic by the identification of protein partners associated to the enzyme on the route from ER to lysosomes. Similarly, the uptake pathway of recombinant GAA (rhGAA), used in ERT, was also studied in order to define the fate of recombinant protein once got into the cell.

The identification of intracellular factors regulating the traffic of endogenous GAA might shed light on cellular processes impaired in the presence of GAA mutant causative of Pompe Disease. Moreover, the description of the cellular route(s) followed by rhGAA is useful to understand why a low amount of protein correctly reaching the lysosomes leading inefficacy ERT in several cases.

3.2 Materials and Methods

3.2.1 Materials

All chemicals for electrophoresis, 30% Acrylamide Bis-acrylamide solution, Tetramethylethylenediamine (TEMED), sodium dodecylsulfate (SDS), Ammonium Persulphate (APS), Tris-Glicine-SDS (TGS), Tris-HCl and the mixture of molecular known weight proteins (total blue) used as standard were from BIO-RAD.

In addition, from BIO-RAD were purchased chemicals and devices for Western Blot: Trans-Blot Turbo Transfer Pack, Tween 20, PBS and nonfat dry milk used for blocking solution, Triton X-100, as well as the "BioRad Protein Assay" kit and bovine serum albumine (BSA) employed in Bradford assay.

Protease inhibitors (Complete mini EDTA-free) were ROCHE. Reagents for ECL, photographic plates and ECL anti-mouse and anti-rabbit were purchased from GE HEALTHCARE. The Colloidal Coomassie Blue Brilliant was PIERCE.

Dithiothreitol (DTT), trypsin, sodium chloride (NaCl), potassium chloride (KCl), trifluoroacetic acid (TFA) and acetonitrile (ACN) LC-MS Grade were SIGMA-ALDRICH products.

Beads for pre-cleaning and immunoprecipitation (Dynabeads protein G) were from Invitrogen. The iodoacetoamide (IAM) and ammonium bicarbonate (AMBIC) are from FLUKA. Formic acid (HCOOH) and acetonitrile HPLC Grade were Baker products.

Anti-GSN primary antibody was a SIGMA-ALDRICH product. Anti-TPM3 primary antibody was provided by PROTEIN-TECH. MYO1C primary antibody was a SANTA CRUZ product. Anti α-GAL primary antibody is a NOVUS BIOLOGICALS product. The antibodies anti-GAA and anti-MYO6 were provided by collaborators of TIGEM (Telethon Institute of Genetics and Medicine).

3.2.2 Methods

3.2.2.1 Cell Lysis and Protein Quantification

The present project was carried out in collaboration with Prof. Giancarlo Parenti's research group of TIGEM, that provided human fibroblasts and anti-GAA and anti-MYO6 antibodies.

Protein extracts were prepared starting from normal fibroblasts expressing wt-GAA, fibroblasts from PD patient biopsy not expressing GAA, fibroblasts from the same patient but incubated for two and six hours with the recombinant enzyme (concentration of 50 ng/mL) named Myozyme. This latter was expressed in Chinese hamster ovary cells (CHO) and used as a drug in the enzymatic replacement therapy.

Lysis buffer composition was:

- ✓ 50mM Tris HCl pH=6.5
- ✓ 150mM NaCl
- ✓ 0.1% Triton X-100
- ✓ 2.5mM KCl
- ✓ 1 mini pill of protease inhibitor for 10 mL of lysis buffer

A ratio of 1:5 (v/v) pellet/buffer was used for lysis of each cell pellet. After buffer addition, pellets were left for 10 minutes in ice, and successively shacked on the wheel at 4°C for 30 minutes. Finally, the suspensions were centrifuged for 30 minutes at 13000 rpm in order to separate protein extracts from debris.

Bradford assays were carried out for a quantitative evaluation of protein extracts (Table 3.2).

Bovine serum albumin (BSA) was used as standard for the calibration curve construction, obtained by reporting the absorbance measurements at 595 nm in function of concentrations of BSA standard solutions.

Table 3.2: Total protein concentration from Bradford assay.

Experiments	Total protein concentration
Endogenous GAA	9.8µg/µL, (6mg of total proteins)
rh-GAA, sample treated for 2h	6.5 µg/µL, (3.8mg total protein)
rh-GAA, control untreated (2h)	6.4 µg/µL, (3.6mg total protein)
rh-GAA, sample treated for 6h	3.6 µg/µL, (6.7mg total protein)
rh-GAA, control untreated (6h)	4.8 µg/µL, (4.6mg total protein)

3.2.2.2 Isolation of Protein Complexes by Immunoprecipitation

Proteins extract from different fibroblast preparations (normal, PD, and PD treated with Myozyme) were subjected to the immunoprecipitation procedure.

As first step, cell extracts were pre-cleared in order to remove unspecific proteins adsorbed on beads. Pre-cleaning was performed by incubating protein extracts with Dynabeads Protein-G (25µL of slurry resin per mg of protein extract), previously conditioned in lysis buffer, for 3h at 4°C on the wheel.

Pre-cleaned extracts were then incubated with specific antibody anti-GAA (policlonal rabbit) (2µg of antibody per mg of protein extract), over-night at 4°C on the wheel.

Finally, the samples were transferred in new tubes and incubated with Dynabeads Protein-G (50µL of slurry resin per mg of protein extract), for 3h at 4°C under gently agitation on the wheel in order to capture antigen-antibody complexes.

The unbound proteins were then removed, and the beads were repetitively washed with lysis buffer containing either low (150mM) or high (300mM) NaCl concentrations, in order to remove unspecific proteins. Proteins retained were eluted by using Laemli Buffer (100mM Tris HCl pH=6,8, 4% SDS, 20% glycerol, bromophenol blue) for 10 min at 99°C in thermomixer. Beads employed in the pre-cleaning steps were treated in parallel with the same procedure, and the eluted proteins constituted the negative control.

3.2.2.3 Preparative SDS-PAGE

100mM DTT (dithiothreitol) was added to each samples following 10 minutes of boiling at 99°C.

Sample and control eluates were loaded on a 16x16 cm, 10% polyacrylamide gel for mono-dimensional SDS-PAGE fractionation. The electrophoretic runs were performed at initial voltage of 200 V.

The gels were stained by Colloidal Blue Comassie over-night and the excess of dye was removed by extensive washing with deionized water.

3.2.2.4 Hydrolysis in situ

Refer to Chapter 2, section 2.2.2.4.

3.2.2.5 LC-MS/MS analysis

Each peptide mixture was resuspended in 12µL of 0,2% HCOOH, and analyzed by LC-MS/MS, using a system *LTQ Orbitrap XL (Thermo Fisher)* equipped with a nano-HPLC. After loading, the peptide mixture was first concentrated and desalinated in the pre-column (C18 EasyColumn L=2cm, 5µm, ID=100µm, *Thermo Fisher Scientific*). Each peptide sample was then fractionated on a C18 reverse-phase capillary column (C18 EasyColumn L=10cm, ID=75µm, 3µm, *Thermo Fisher Scientific*) working at a flow rate of 300nL/min, with a gradient of eluent B (0,2% formic acid, 95% acetonitrile LC-MS Grade) and eluent A (0,2% formic acid, 2% acetonitrile LC-MS Grade) from 5 to 60% in 70 mins.

Peptide analysis was performed using data-dependent acquisition (DDA) of one MS scan (mass range from 400 to 1800 m/z) followed by MS/MS scans of the five most abundant ions in each MS scan, obtained by CID (*Collision Induced Fragmentation*) fragmentation in the ion trap.

3.2.2.6 Proteins Identification

Raw data from nano-LC-MS/MS analyses were processed and converted in .mgf files to be introduced into the MASCOT software (Matrix Science Boston, USA) to search a non-redundant protein database. Both peptide mass data and the data obtained by fragmentation spectra were included in the *peak list* for protein identification.

The research was done by setting the following parameters:

- ✓ Database: NCBI
- ✓ Taxonomy: Homo sapiens
- ✓ Enzyme: Trypsin
- ✓ Fixed modification: carbamidomethyl (C)
- ✓ Variable modification: oxidation (M), pyro-Glu (N-term Q), pyrocarbamidomethyl (N-term C)
- ✓ Peptide tolerance: 10 ppm
- ✓ MS/MS tolerance: ± 0,6 Da
- ✓ Ions charge: +2, +3
- ✓ Instrument: ESI-FTICR

It has also been used the software ExPASy Blast Form to assign to each protein a code that uniquely identify each sequence in the annotated database UniProt KB.

3.2.2.7 Co-Immunoprecipitation and Western Blotting Experiments

Co-immunoprecipitation (Co-IP) experiments were carried out in order to verify the physiological interaction between endogenous GAA and several potential interacting proteins among those identified by functional proteomic approach. These GAA putative interactors are: Gelsolin (GSN), Unconventional myosin-IC (MYO1C), Unconventional myosin-VI (MYO6), Tropomyosin alpha-3 chain (TPM3).

The protocol used for the Co-IP was the same employed for the preparative IP (refer to section 3.2.2.2), although Co-IP experiments were carried out on lower protein amounts (800μg).

The ratio antibody / total protein extract was: 2μg of antibody per mg protein extract. This condition was used for each Co-IP experiment.

Finally, eluates from IP, controls and 100µg of total extract (INPUT) were diluted in Loading Buffer, with the addition of 100mM DTT (dithiothreitol), boiled for 10 minutes at 99°C and fractionated by SDS-PAGE into a 10% Acrylamide Bis gel (size 10x10 cm and thickness of 1,5 mm).

Proteins separated by SDS-PAGE were then electro-transferred to nitrocellulose membranes with Trans-Blot Turbo Blotting system for mini-gel. Membranes were stained with Ponceau Red (0,1% solution in 1% acetic acid) to visualise the effective transfer of the proteins. After destaining in 1x PBS (1mM sodium phosphate, 15mM NaCl, pH 7.4) the membranes were immersed in a blocking solution, consisting of 5% nonfat dry milk in 1x PBS for 1h at room temperature, on a shaking plate. Then, membranes were incubated with specific primary antibody (Table 3.3), diluted in 1% nonfat dry milk in 1x PBS, 0,1% Tween20.

Table 3.3: Primary antibodies dilutions employed for Western blot assays.

Antibody	Ratio I antibody	Incubation condition
Anti-GAA (rabbit)	1/2000	2 hours, rt
Anti-GSN (mouse)	1/500	2 hours, rt
Anti-TPM3 (rabbit)	1/1000	Over night, 4°C
Anti-MYO6 (rabbit)	1/750	Over night, 4°C
Anti-MYO1C (rabbit)	1/500	Over night, 4°C

The incubation with primary antibodies was followed by three washes with 1x PBS and 0,1% Tween20 (washing solution) and then by 45 minutes incubation with the II antibody (1/5000) at room temperature, in a solution of 1% nonfat dry milk in 1x PBS, 0,1% Tween20. The II antibody is covalently conjugated to the horseradish peroxidase (HRP), an enzyme that, in presence of hydrogen peroxide (H_2O_2), oxidizes its chemiluminescent substrate, luminol, to acid 3-amminoftalic with concomitant emission of light, which is impressed on a photographic plate. Following three washes to remove the excess of II antibody, the enzymatic development was carried out by using an ECL kit consisting of luminol and H_2O_2.

3.3 Results

The investigation of molecular mechanism involving GAA trafficking was carried out by functional proteomic approach. The identification of protein partners interacting with endogenous GAA in normal and PD fibroblasts allowed us to describe the route followed by GAA from ER to lysosomes. This led to formulate a hypothesis on where this route is interrupted when mutants of GAA causative of Pompe Disease occur. By using the same approach, the traffic of recombinant protein (rhGAA) from outside the cell to lysosomes was also investigated. The final aim of the project was to unravel the cellular processes involving GAA at molecular level in order to understand the defects in the physiological traffic towards its natural sub-cellular compartment where the enzyme exerts its biological function. All the information collected during this study will be useful to develop new protocols or to improve existing procedures to correct the aberrant traffic, providing more effective therapies against Pompe Disease.

3.3.1 Investigation of Wild Type GAA Interactome

The investigation of wild type GAA interactome was carried out using normal human fibroblasts, provided by our co-workers from Professor Parenti's laboratory.

Fibroblasts were lysated as reported above and 6 mg of total protein extract were immunoprecipitated by using anti-GAA antibody. In order to remove false positive interactors (proteins with affinity for the matrix, magnetic beads derivatised with protein G), total cellular extract was pre-cleared by preliminary incubation with magnetic beads in the absence of specific antibodies. Proteins retained by the beads during the pre-cleaning step were eluted and employed as control.

The pre-cleared extract was, then, incubated first with anti-GAA antibody over night and then for 2 hours with the Dynabeads derivatized with the G protein; the supernatant containing all the unbound proteins was then discarded and the beads retaining the immunoprecipited proteins were extensively washed with lysis buffer.

Finally, proteins specifically bound to the antibody (GAA containing complexes) or to the beads (control) were eluted in denaturing conditions (Laemli buffer) and the individual proteins fractionated by SDS-PAGE and stained by colloidal blue coomassie (Figure 3.9). 33 protein bands for the control and sample lanes were excised from the gel, digested in situ with trypsin and the resulting peptide mixtures were directly analyzed by LC-MS/MS, using a *LTQ Orbitrap XL* mass spectrometer *(Thermo Fisher)* equipped with a nano-HPLC.

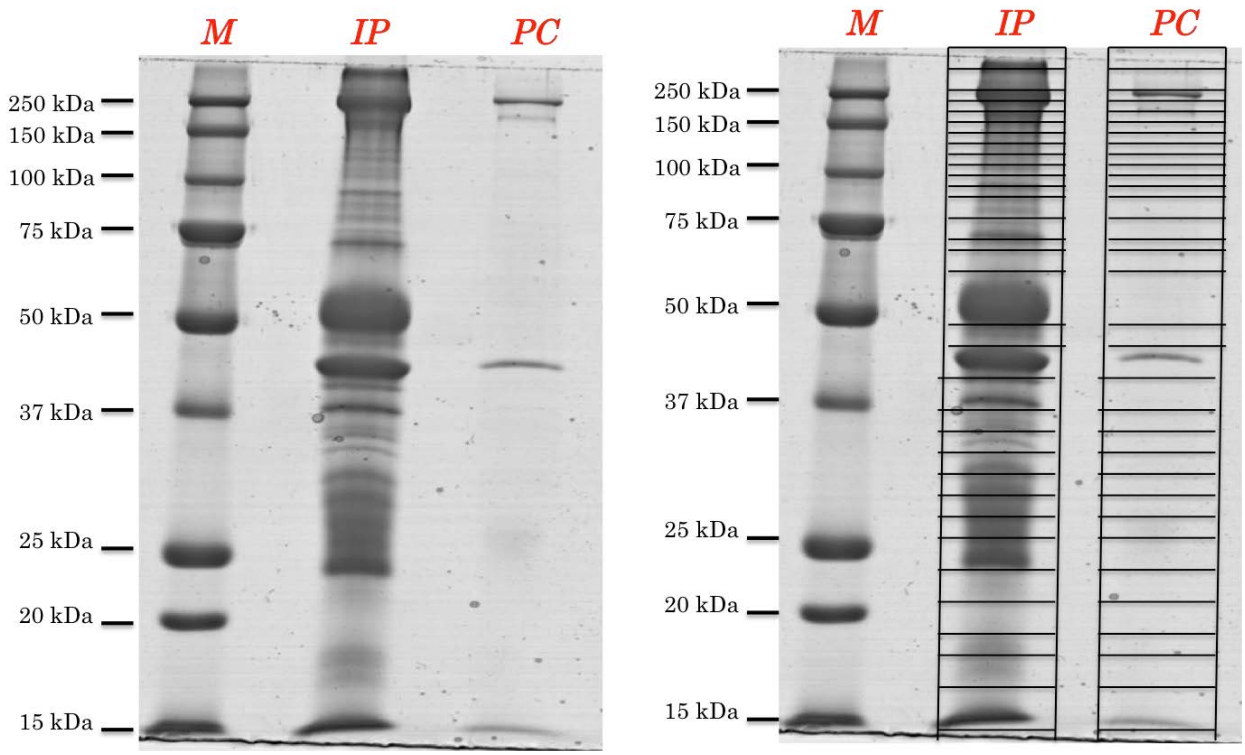


Figure 3.9: SDS-PAGE of sample (IP) and control (PC), M represents a standard mixture of proteins with known molecular weight (left panel). Scheme of protein bands excised from the gel (right panel).

The final list of GAA putative interactors (Table 3.4) was obtained by discording all the proteins occurring both in the sample and in the control (PC). For each protein, the molecular weight, the total number of identified peptides (in parentheses those with a score higher than the reliability threshold defined by Mascot), the UniProt code, the gene name, keywords for GO Molecular Functions and Biological process are also reported.

Table 3.4: List of putative GAA protein partners. Protein name, gene name, number of identified peptides, in parentheses those with scores above the threshold, UniProt code and the key words for GO Molecular Functions and Biological Process of the identified protein are reported. In bold are reported the identification parameters of GAA (the bait).

MW (kDa)	Protein name	Gene	Peptides	UniProt code	GO Molecular Function	GO Biological Process
> 250	Plectin	PLEC	10(3)	Q15149	Structural constituent of muscle	Hemidesmosome assembly
150-100	Unconventional myosin-VI	MYO6	3(1)	Q9UM54	Actin Binding	Actin filament-based movement
	Unconventional myosin-Ic	MYO1C	22(11)	O00159	Microfilament motor activity	Actin filament-based movement
	Unconventional myosin-Ib	MYO1B	7(5)	O43795	Actin-dependent ATPase activity	Actin filament-based movement
100	**Lysosomal alpha-glucosidase**	**GAA**	**6(4)**	**P10253**	**alpha-1,4-glucosidase activity**	**Glucose metabolic process**
100-75	Gelsolin	GSN	16(8)	P06396	Actin and Myosin II Binding	Actin filament polymerization
	Unconventional myosin-Id	MYO1D	8(6)	O94832	Actin filament binding	Actin filament-based movement
75-50	Coronin-1C	CORO1C	5(3)	Q9ULV4	Actin filament binding	Actin cytoskeleton organization
50-37	Tropomodulin-3	TMOD3	4(2)	Q9NYL9	Tropomyosin binding	Actin cytoskeleton organization
37-25	Tropomyosin alpha-1 chain	TPM1	18(9)	P09493	Actin binding	Actin cytoskeleton organization
	Tropomyosin alpha-4 chain	TPM4	14(11)	P67936	Actin filament binding	Actin filament organization
	Tropomyosin alpha-3 chain	TPM3	10(8)	P06753	Actin filament binding	Actin filament organization
	Tropomyosin beta chain	TPM2	4(2)	P07951	Actin binding	Actin filament organization
20-15	60S ribosomal protein L12	RPL12	2(2)	P30050	rRNA binding	Translation

As shown by the Gene Ontology categories (Molecular Functions and Biological Process), almost all the GAA interactors belong to the actin cytoskeleton, including Gelsolin (GSN) and tropomyosin (TPM) or are involved in protein trafficking, such as unconventional myosins.

Co-immunoprecipitation (Co-IP) experiments have been performed in normal fibroblasts in order to validate the interaction between GAA and four putative interactors: Gelsolin (GSN), Unconventional myosin-1C (MYO1C), Unconventional myosin–VI (MYO6), Tropomyosin-3 (TPM3).

The IP experiments were carried out on lower protein amounts (800µg) using the same condition of the preparative experiment (i.e. lysis buffer, beads, pre-cleaning step, washes); the ratios between each antibody and total protein extracts are reported in 3.2.2.7 section.

The analyses of the IPs were then performed by western blot, by using specific antibodies in order to reveal the proteins of interest. Co-IP experiments were carried out both directly, immunoprecipiting GAA and revealing for the interactors and reverse, namely immunoprecipiting the interactors and revealing for GAA (Figure 3.10-3.13).

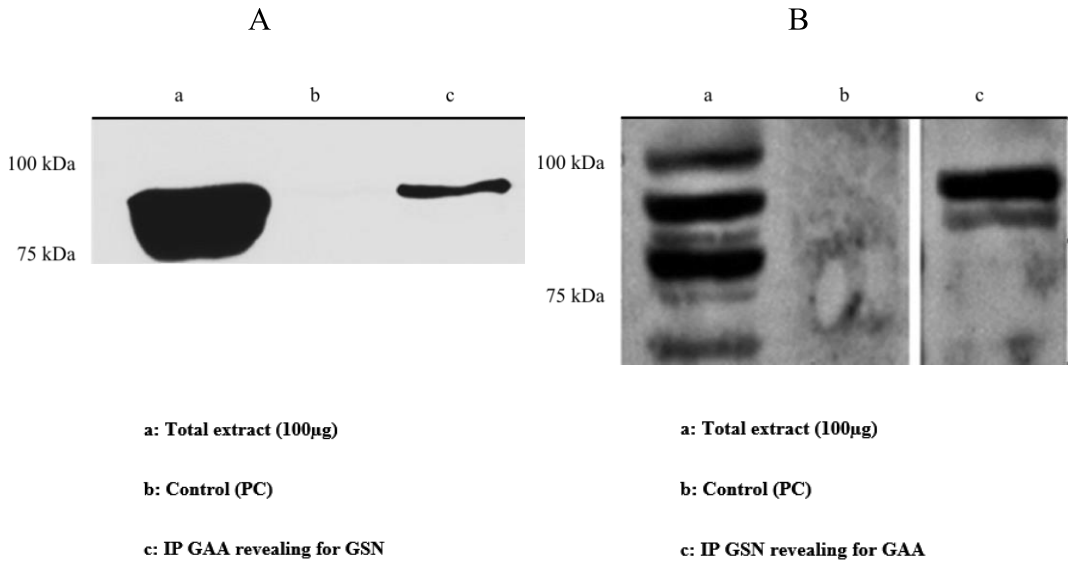


Figure 3.10: Western Blotting to confirm the GAA-GSN interaction.

Panel A: Immunoprecipitation with anti-GAA antibody and detection with anti-GSN antibody. **Panel B:** Reverse experiment, Immunoprecipitation with anti-GSN antibody and detection with anti-GAA antibody. In both panels are present in the IP lanes a specific band at the molecular weight expected for the immunoprecipited protein.

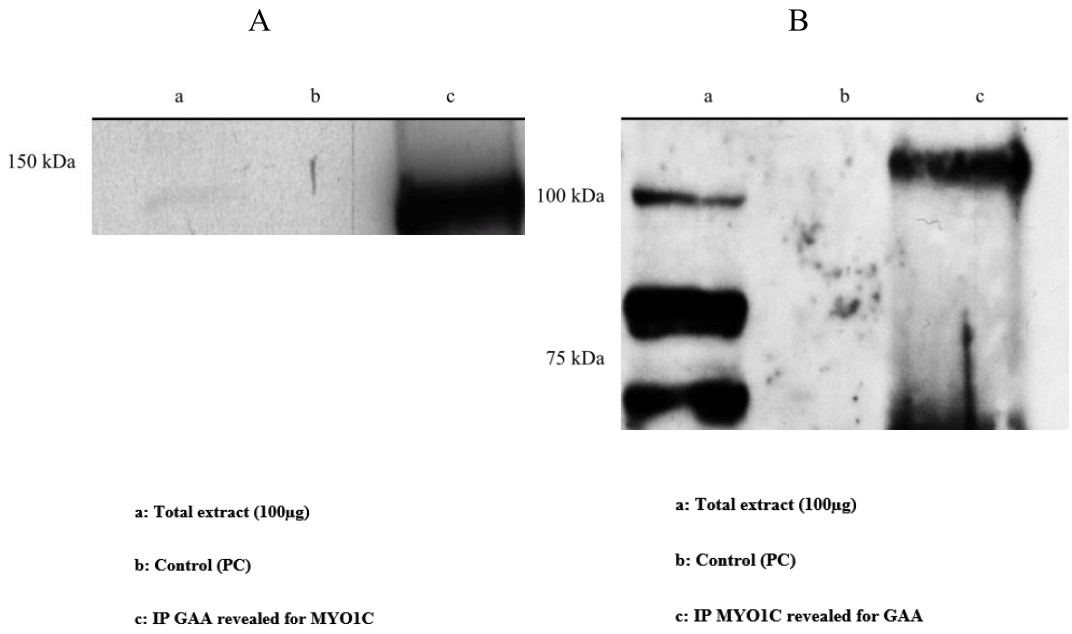


Figure 3.11: Western Blotting to confirm the GAA-MYO1C interaction

Panel A: Immunoprecipitation with anti-GAA antibody and detection with anti-MYO1C antibody. **Panel B:** Reverse experiment, Immunoprecipitation with anti- MYO1C antibody and detection with anti-GAA antibody. In both panels are present in the IP lanes a specific band at the molecular weight expected for the immunoprecipited protein.

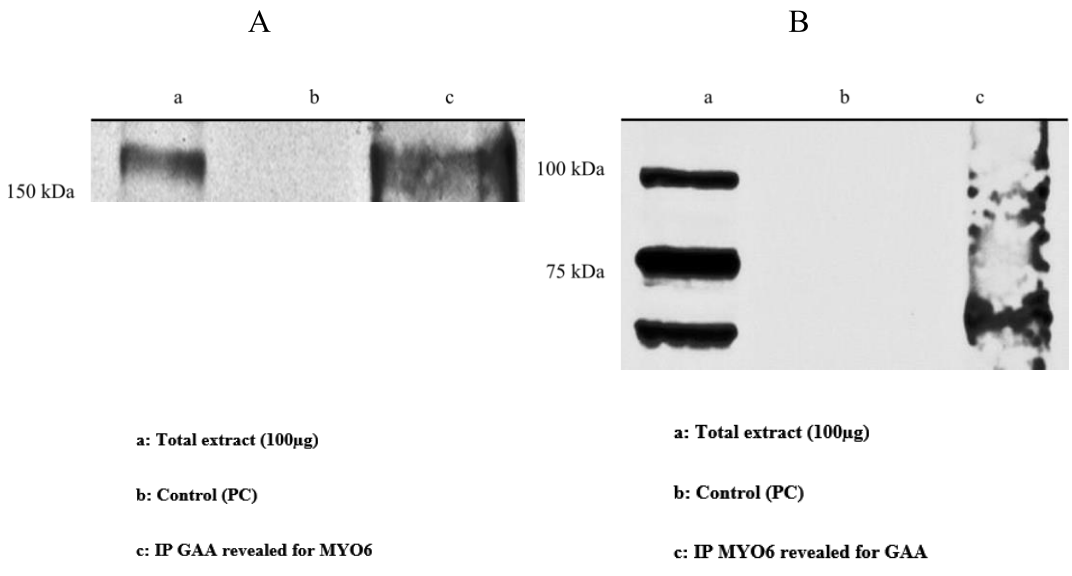


Figure 3.12: Western Blotting to confirm the GAA-MYO6 interaction

Panel A: Immunoprecipitation with anti-GAA antibody and detection with anti-MYO6 antibody. **Panel B:** Reverse experiment, Immunoprecipitation with anti-MYO6 antibody and detection with anti-GAA antibody. In both panels are present in the IP lanes a specific band at the molecular weight expected for the immunoprecipited protein.

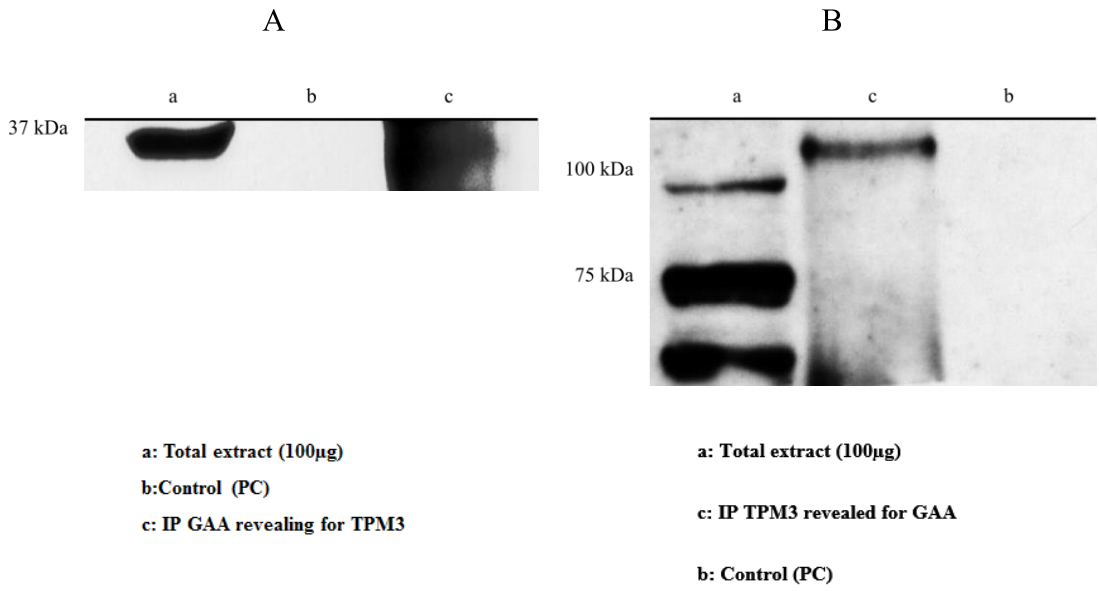


Figure 3.13: Western Blotting to confirm the GAA-TPM3 interaction

Panel A: Immunoprecipitation with anti-GAA antibody and detection with anti-TPM3 antibody. **Panel B:** Reverse experiment, Immunoprecipitation with anti-TPM3 antibody and detection with anti-GAA antibody. In both panels are present in the IP lanes a specific band at the molecular weight expected for the immunoprecipited protein.

As shown in previous figures, all direct co-IP experiments confirmed that GAA interacts with GSN, MYO1C, MYO6, and TPM3 as suggested by proteomic experiments. Analogous results were obtained from the reverse experiments (Fig. 3.10-3.13, panels B). In all western blots developed with anti-GAA antibody, several protein bands whose electrophoretic mobility ranges between 100 and 70kDa were observed. These bands correspond to different GAA isoforms, differing for the level of proteolytic maturation (Figure 3.4). Surprisingly, the western blot assays showed that the interaction of GAA with each partner involves only one specific isoform among all those existing in vivo. In details, TMP3 and MYO1C interact with the GAA zymogene at 110 kDa (Fig 3.11B and 3.13B), while GSN and MYO6 specifically interact with the mature forms at 95 (Fig. 3.10B) and 70kDa (Fig. 3.12B), respectively.

These data suggest that GAA is involved in different supra-molecular complexes according to its maturation level. It is plausible that all these complexes occur in different moments of GAA time life, accompanying the enzyme in the different steps along the route from ER to lysosomes.

The interaction between GAA and GSN was further validated *in vivo* in normal fibroblasts by *Proximity Ligation Assay* (in situ PLA) (Figure 3.14). This advanced immunoassay technique is able to localise proteins with high accuracy, to detect protein-protein interactions and protein modifications with high specificity and sensitivity. The interaction between GSN and ACTIN was used as positive control, being GSN an actin-binding protein, while the interaction between GSN and LAMP2 constituted the negative control, being LAMP2 a lysosomal marker.

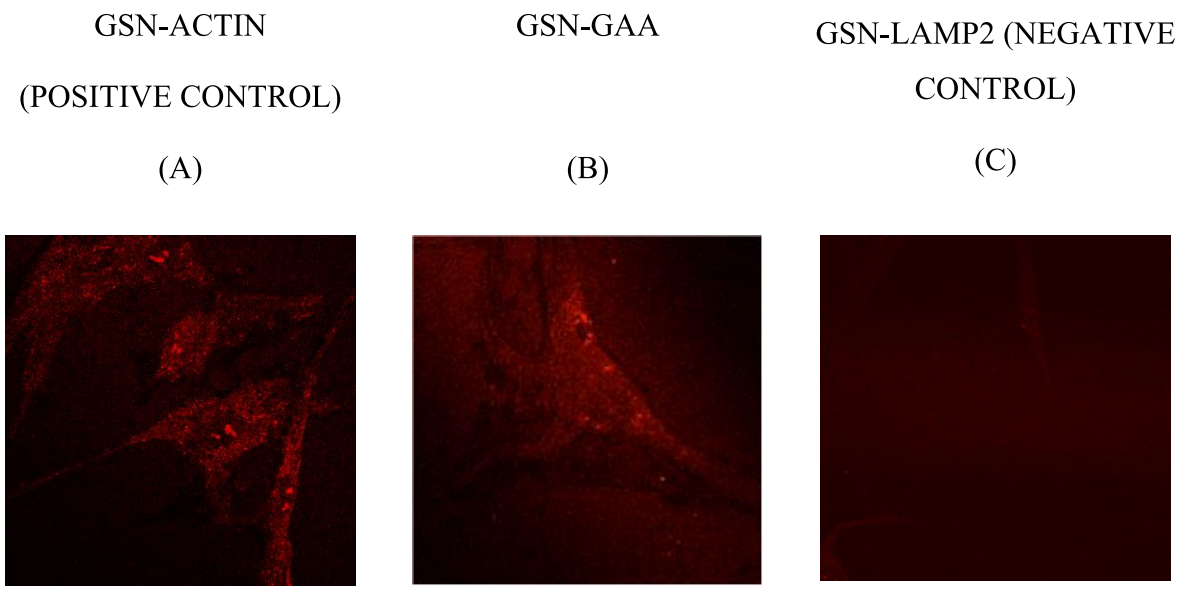


Figure 3.14: Proximity Ligation Assay technique shows protein-protein interaction (in red): (A) Gelsolin-Actin interaction is the positive control; (C) Gelsolin-Lamp2 interaction is the negative control; (B) Gelsolin-Alpha Glucosidase interaction.

As shown in the Fig. 3.14B several red emission points are present, suggesting the occurrence of GAA-GSN interaction inside the cell, although this interaction is not as evident as the one occurring between ACTIN and GSN.

A question arises whether the interaction between GSN and GAA is specific for this enzyme, or GSN is able to interact with other lysosomal enzyme, including α-galactosidase A (α-GAL), associated to Fabry Disease. Co-immunoprecipitation experiments carried out immunoprecipiting α-GAL and revealing for GSN demonstrated the occurrence of an interaction between α-galactosidase A and Gelsolin (Figure 3.15).

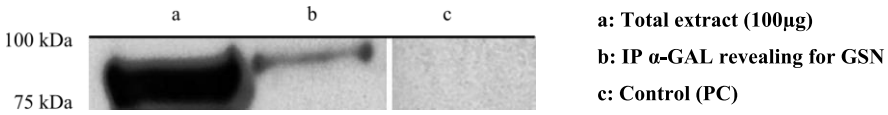


Figure 3.15: Western Blotting to confirm the α-GAL-GSN interaction.

These data confirm that both GAA and α-GAL interact with GSN, suggesting general mechanism of intracellular trafficking with this interaction being functional for all enzymes directed to lysosomes, occurring during the route towards lysosome. In order to deeply understand the role of GSN in GAA trafficking, the interaction between GNS and the L552P GAA mutant associated to a PD phenotype was investigated by Co-IP.

Immunoprecipitation of endogenous L552P-GAA and GSN were performed starting from total fibroblasts protein extract of a PD patient homozygote for the L552P variant. Both direct and

reverse immunoprecipitation, experiments demonstrated that the mutant the mutant GAA was not able to interact with GSN (Figure 3.16).

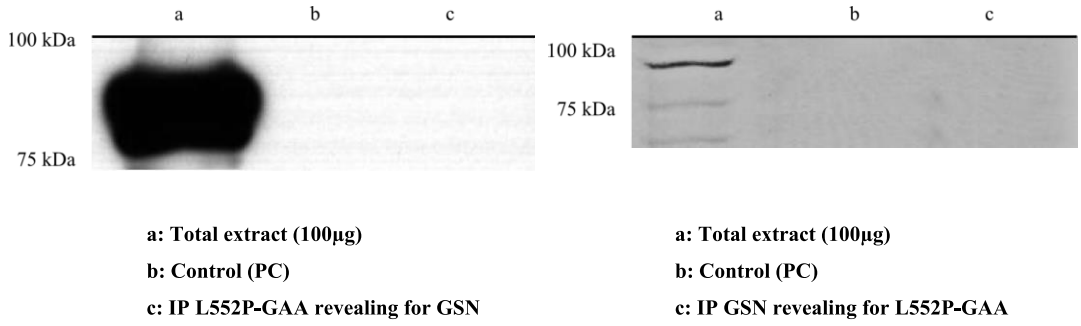


Figure 3.16: Western Blotting to confirm the L552P-GAA- GSN interaction.

These data confirm that the native interaction between GAA and GSN is impaired in PD, strongly suggesting that the association between GAA and GSN might be crucial for the correct traffic of GAA towards lysosome and that its absence might be critical for the onset of a pathological state.

3.3.2 Investigation of rhGAA Interactome

An analogous proteomic approach was employed to investigate the molecular mechanisms controlling the uptake of recombinant GAA (rhGAA) occurring during the enzymatic replace treatment of PD patients. Fibroblasts derived by a patient affected by Pompe Disease and totally deficient GAA enzyme were treated for either two or six hours with Myozyme drug (rhGAA), while untreated fibroblasts of the same patient were used as control.

3.6 mg of protein extract from each sample and control were immunoprecipitated by using anti-GAA antibody according to the protocol reported in section 3.2.2.2.

Recombinant GAA protein partners were fractionated by SDS-PAGE (Figure 3.17); 32 bands were excised from the gel, digested in situ with trypsin and analysed by nanoLC-MS/MS using LTQ Orbitrap XL (Thermo Scientific) and identified by MASCOT, as described above.

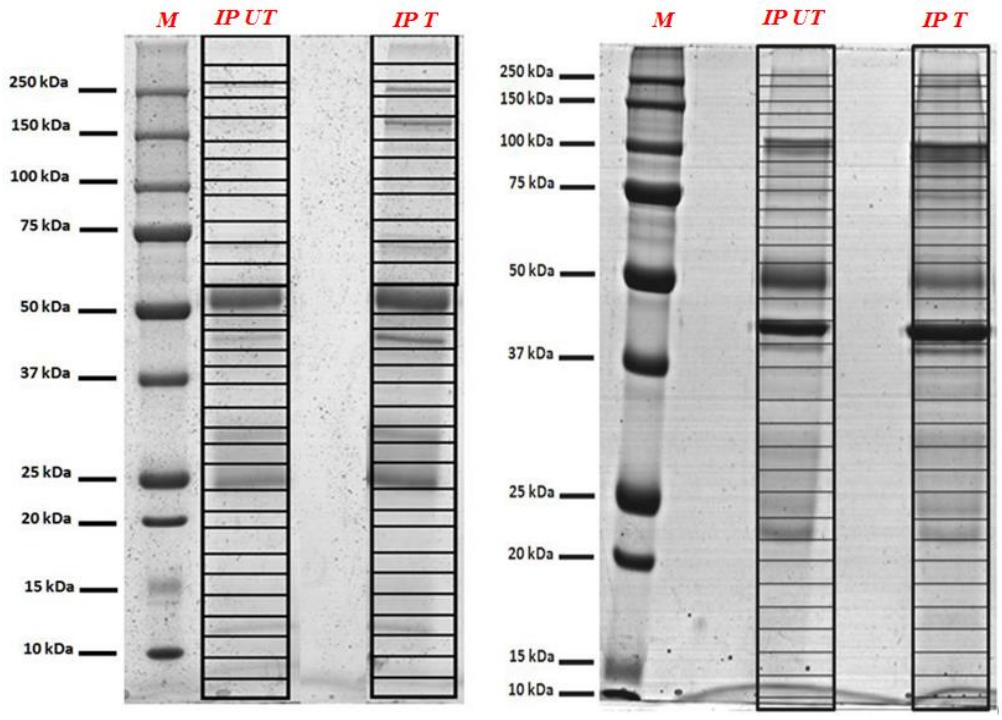


Figure 3.17: SDS-PAGE of the immunoprecipitated samples from fibroblasts treated with Myozyme for 2h (left) and for 6h (right). Lanes indicated with "IP UT" are the controls (untreated), while the ones indicated with "IP T" are the samples (treated). M represents a mixture of known molecular weight proteins used as standards.

The list of rhGAA putative interactors (Table 3.5) was obtained by discording the proteins identified in both in the sample and in the control (IP UT) and considering those only occurring in sample (IP T).

Table 3.5: Putative rhGAA protein partners identified at 2h and 6 h of incubation. Protein name, gene name, number of identified peptides in both experiments, in parentheses those with scores above the threshold, UniProt code and GO Biological process are reported.

MW (kDa)	Protein	Gene	Peptides 2h	Peptides 6h	UniProt code	GO Biological Process
> 250	Filamin-B	FLNB	14 (5)	-	O75369	Actin cytoskeleton organization
	Neuroblast differentiation-associated protein AHNAK	AHNAK	10 (7)	-	Q09666	Protein oligomerization
250-150	Myosin-10	MYH10	38 (28)	-	P35580	Actin filament-based movement
	Ras GTPase-activating-like protein IQGAP1	IQGAP1	16 (6)	-	P46940	Signal transduction
	Clathrin heavy chain 1	CLTC	2 (1)	-	Q00610	Intracellular protein transport
150-100	**Lysosomal alpha-glucosidase**	GAA	7 (5)	10 (7)	P10253	Glucose metabolic process
	Unconventional myosin-Ic	MYO1C	20 (10)	10 (6)	O00159	Protein transport

MW (kDa)	Protein	Gene	Peptides 2h	Peptides 6h	UniProt code	GO Biological Process
150-100	Unconventional myosin-Id	MYO1D	7 (3)	6 (2)	O94832	Early endosome to recycling endosome transport
	Major vault protein	MVP	14 (5)	-	Q14764	Protein transport
	Unconventional myosin-Ib	MYO1B	5 (2)	-	O43795	Post-Golgi vesicle-mediated transport
	Cytospin-A	SPECC1L	6 (4)	-	Q69YQ0	Actin cytoskeleton organization
	Liprin-alpha-1	PPFIA1	4 (2)	-	Q13136	Signal transduction
	SERPINB	SERPINB3	-	9 (4)	P29508	Autocrine signaling
	ProteinS100-A8	S100A8	-	7 (6)	P05109	Apoptotic process
	Protein S100-A9	S100A9	-	5 (5)	P06702	Apoptotic process
	Protein S100-A7	S100A7	-	4 (4)	P31151	Innate immune response
	Protein-glutamine gamma-glutamyltransferase K	TGM1	-	6 (2)	P22735	Keratinocyte differentiation
	Alpha-2-macroglobulin-like protein 1	A2ML1	-	5 (1)	A8K2U0	Regulation of endopeptidase activity
	Protein POF1B	POF1B	-	3 (3)	Q8WVV4	Organization of epithelial monolayers
	Serpin B12	SERPINB12	-	4 (2)	Q96P63	Negative regulation of protein catabolic process
100	**Lysosomal alpha-glucosidase**	GAA	13 (9)	8 (5)	P10253	Glucose metabolic process
	Endoplasmin	HSP90B1	11 (9)	-	P14625	ER-associated ubiquitin-dependent protein catabolic process
	Protein phosphatase 1 regulatory subunit 12A	PPP1R12A	2 (2)	-	O14974	Cellular response to drug
	Phostensin	PPP1R18	13 (7)	-	Q6NYC8	Target protein phosphatase 1 to F-actin cytoskeleton
	Actin filament-associated protein1	AFAP1	6 (3)	-	Q8N556	Regulation of cytoskeleton organization
	Ribosome-binding protein1	RRBP1	3 (2)	-	Q9P2E9	Translation
	Dynamin-2	DNM2	3 (2)	-	P50570	Endocytosis
100-75	**Lysosomal alpha-glucosidase**	GAA	4 (2)	7 (5)	P10253	Glucose metabolic process

MW (kDa)	Protein	Gene	Peptides 2h	Peptides 6h	UniProt code	GO Biological Process
100-75	ATP-dependent 6-phosphofructokinase, liver type	PFKL	7 (6)	-	P17858	Glycolytic process
	Protein-glutamine gamma-glutamyltransferase 2	TGM2	-	3 (1)	P21980	Apoptotic cell clearance
75	Moesin	MSN	19 (10)	-	P26038	Positive regulation of early endosome to late endosome transport
	Stress-70 protein, mitochondrial	HSPA9	12 (5)	-	P38646	Protein folding
	*Glycine-tRNA ligase	GARS	13 (7)	-	P41250	tRNA aminoacylation for protein translation
	Synaptopodin-2	SYNPO2	6 (2)	-	Q9UMS6	Positive regulation of actin filament bundle assembly
	Annexin A6	ANXA6	5 (3)	-	P08133	Protein homooligomerization
	Transforming growth factor-beta-induced protein ig-h3	TGFBI	4 (2)	-	Q15582	Cell adhesion
75-50	**Lysosomal alpha-glucosidase**	GAA	4 (3)	7 (6)	P10253	Glucose metabolic process
	5'-nucleotidase	NT5E	2 (2)	7 (4)	P21589	Negative regulation of inflammatory response
	Palladin	PALLD	11 (7)	-	Q8WX93	actin cytoskeleton organization
	PDZ and LIM domain protein 5	PDLIM5	6 (2)	-	Q96HC4	Cell growth involved in cardiac muscle cell development
	Dihydropyrimidinase-related protein 2	DPYSL2	2 (2)	-	Q16555	Cytoskeleton organization
	T-complex protein 1 subunit alpha	TCP1	6 (3)	-	P17987	Protein folding
	Prolyl 4-hydroxylase subunit alpha-1	P4HA1	5 (3)	-	P13674	Collagen fibril organization
	Ser/thr-protein phosphatase 2A 65 kDa regulatory sub A	PPP2R1A	5 (1)	-	P30153	Apoptotic process
	Prolyl 4-hydroxylase sub α-2	P4HA2	2 (2)	-	O15460	oxidation-reduction process
	Pyruvate kinase PKM	PKM	6 (4)	-	P14618	Glycolytic process
	Protein disulfide-isomerase	P4HB	6 (3)	-	P07237	Cell redox homeostasis

MW (kDa)	Protein	Gene	Peptides 2h	Peptides 6h	UniProt code	GO Biological Process
75-50	Methylcrotonoyl-CoA carboxylase beta chain, mitochondrial	CCC2	3 (2)	-	Q9HCC0	Biotin metabolic process
50	Coronin-1C	CORO1C	18 (8)	-	Q9ULV4	Actin cytoskeleton organization
	Alpha-enolase	ENO1	10 (7)	-	P06733	Glycolytic process
	Protein disulfide-isomerase A6	PDIA6	5 (4)	-	Q15084	Cell redox homeostasis
	Protein disulfide-isomerase A3	PDIA3	7 (4)	-	P30101	Cell redox homeostasis
	Fascin	FSCN1	6 (4)	-	Q16658	actin cytoskeleton organization
	Ca/calmodulin-dependent protein kinase type II sub delta	CAMK2D	4 (4)	-	Q13557	Calcium ion transport
	Protein-lysine 6-oxidase	LOX	5 (2)	-	P28300	Oxidation-reduction process
	Septin-11	SEPT11	3 (3)	-	Q9NVA2	Cell cycle
37-25	Synaptic vesicle membrane protein VAT-1 homolog	VAT1	3 (2)	-	Q99536	Calcium-regulated keratinocyte activation
	Perilipin-3	PLIN3	2 (2)	-	O60664	Vesicle-mediated transport
	Phosphoglycerate kinase1	PGK1	7 (7)	-	P00558	Glycolytic process
	F-actin-capping protein sub α-1	CAPZA1	2 (2)	2 (1)	P52907	Barbed-end actin filament capping
	PDZ and LIM domain protein 1	PDLIM1	15 (13)	-	O00151	Regulation of transcription
	Annexin A1	ANXA1	7 (6)	-	P04083	Actin cytoskeleton reorganization
	Annexin A2	ANXA2	8 (5)	-	P07355	Protein targeting to plasma membrane
	60S acidic ribosomal protein P0	RPLP0	4 (4)	-	P05388	Translation
	Ser/thr-protein phosphatase PP1-α catalytic sub	PPP1CA	6 (4)	-	P62136	Glycogen metabolic process
	Tropomyosin alpha-4 chain	TPM4	5 (4)	-	P67936	Actin filament organization
	EF-hand domain-containing protein D2	EFHD2	3 (2)	-	Q96C19	Cadherin binding involved in cell-cell adhesion

MW (kDa)	Protein	Gene	Peptides 2h	Peptides 6h	UniProt code	GO Biological Process
37-25	Coatomer subunit delta	ARCN1	-	2(1)	P48444	Intracellular protein transport
25-20	Triosephosphate isomerase	TPI1	10 (5)	-	P60174	Glycolytic process
	Alpha-crystallin B chain	CRYAB	4 (2)	-	P02511	Protein folding
	Proteasome subunit beta type-6	PSMB6	-	4 (4)	P28072	Protein polyubiquitination
	Proteasome subunit beta type-5	PSMB5	-	5 (3)	P28074	Protein polyubiquitination
20-15	40S ribosomal protein S18	RPS18	3 (2)	5 (4)	P62269	Translation
	Heat shock protein beta-1	HSPB1	5 (2)	-	P04792	Regulation of apoptotic process
	40S ribosomal proteinS13	RPS13	3 (2)	-	P62277	Translation
	Peptidyl-prolyl cis-trans isomerase A	PPIA	3 (2)	-	P62937	Protein folding
	40S ribosomal proteinS25	RPS25	2 (2)	-	P62851	Translation
15-10	Histone H4	HIST1H4A	-	5(5)	P62805	Transcription regulation
	Fatty acid-binding protein, epidermal	FABP5	-	6 (2)	Q01469	Lipid metabolic process
	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-12	GNG12	-	4 (1)	Q9UBI6	Signal transduction
	Calmodulin-like protein5	CALML5	-	3 (1)	Q9NZT1	Signal transduction

3.4 Discussion

Pompe disease (PD) is an inherited disorder caused by mutation the acid alpha-glucosidase (GAA) gene, mapped to the long arm of chromosome 17 (location 17q25.2-q25.3). As autosomal recessive disorder, PD only occurs when an individual inherits two mutant alleles, one from each parent. To date, almost 300 distinct GAA mutations have been identified, although not all are considered pathogenic [15].

In general, genotype-phenotype correlation is not well understood and significant clinical heterogeneity exists among patients with similar or identical mutations. One exception is the presence of two null mutations, resulting in a complete absence of the GAA enzyme activity. This genotype results in very early disease onset during infancy and severe, rapid disease progression.

More studies, however, are needed in order to better understand genotype-phenotype relations, because the molecular basis underlying the cellular abnormalities of PD are still not clear. It is supposed that aberrant protein-protein interactions occurring in the presence of GAA variants may contribute to the loss of function by affecting GAA regular trafficking and cation independent-mannose-6-phosphate receptor (CI-MPR) distribution in PD.

The investigation of GAA interactome allowed us to shed light on the physiological steps of GAA traffic from endoplasmic reticulum to lysosomes and to understand the key role of some of these interactions in physiological conditions, when the enzyme localizes in the expected cell compartments.

Once the route followed by native GAA was characterised, the attention focused to the investigation of both mutant GAA enzyme trafficking and the route followed by the recombinant enzyme rhGAA used during the enzyme replace therapy (ERT) in PD fibroblasts not expressing GAA. The aim of this study was the definition of molecular mechanisms involved in protein internalization thus to speculate about the low rescue of GAA biological function in many ERTs. The identification of proteins contacting rhGAA following its internalization, has allowed us to obtain information about the route followed by rhGAA from the entrance into the cell until its localization in lysosomes or other cell organelles.

Bioinformatic analysis with Cytoscape tool was performed in order to highlight unique and common interactors between wild type GAA and rhGAA proteins. A graphical representation of the molecular interactions networks for GAA, rhGAA at 2 and 6 hours of incubation is shown in Figure 3.18. The presence of several common interactors between endogenous GAA and rhGAA, suggests that the route followed inside the cell from both proteins is partially overlapped.

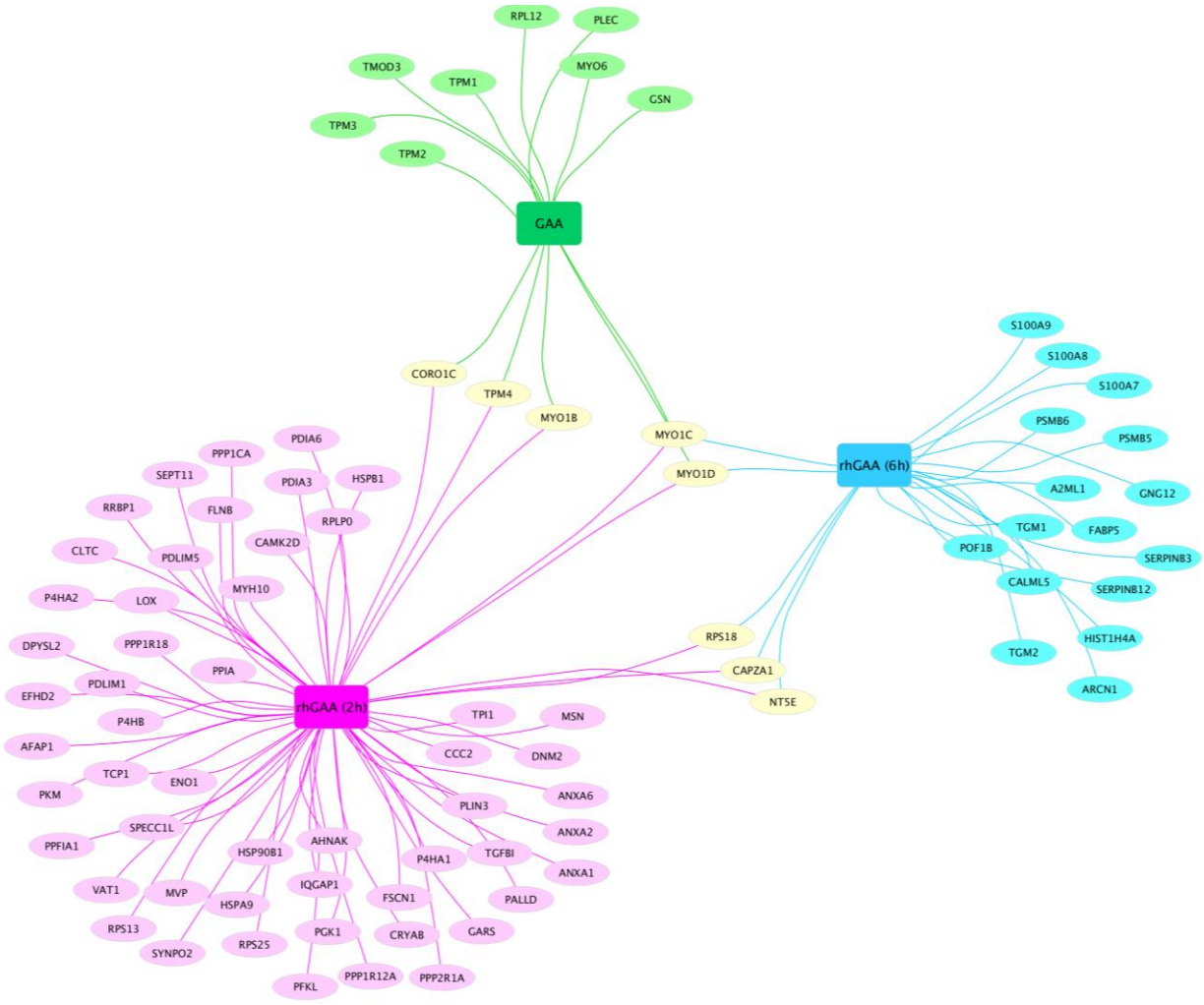


Figure 3.18: Graphical representation of protein interaction networks specific for endogenous GAA in green, rhGAA (2h) in pink, rhGAA (6h) in light blue, and those in common in yellow, obtained by using Cytoscape software.

3.4.1 Investigation of Wild Type GAA Interactome

The investigation of wild type GAA interactome in normal fibroblasts led to the identification of 12 putative interacting proteins, classified according to the main biological function by using Gene Ontology, STRING and/or information reported in literature (Figure 3.19).

A large number of GAA protein partners belongs to the family of actin interacting protein, such as gelsolin (GSN), several tropomyosins (TPM) and unconventional myosins; these latter are actin-based motor molecules involved in intracellular movements.

Bioinformatic analysis with STRING tool revealed that the identified proteins belong to the so-called "*cytoskeleton based processes*" network (Figure 3.20).

These results strongly suggest that the GAA fate is strictly connected to cellular processes concerning the rearrangement of cytoskeleton and protein traffic.

Among identified GAA interactors some proteins belong to the unconventional myosin family, such as *Myo1B, Myo1C, Myo1D, Myo6.*

It is known that *Unconventional myosins* are a superfamily of actin-based molecular motors playing several roles including intracellular vesicle trafficking, mechanical supports, force sensing and transmission. Unconventional myosins are able to bind to specific cargoes including proteins and lipid vesicles by their variable neck and tail domains (Figure 3.21) [16]. Co-IP experiments indicated that interactions occurred between GAA protein partners and specific GAA isoforms. These data were also supported by recent literature reports and allowed us to define the individual step of GAA journey from ER to lysosomes.

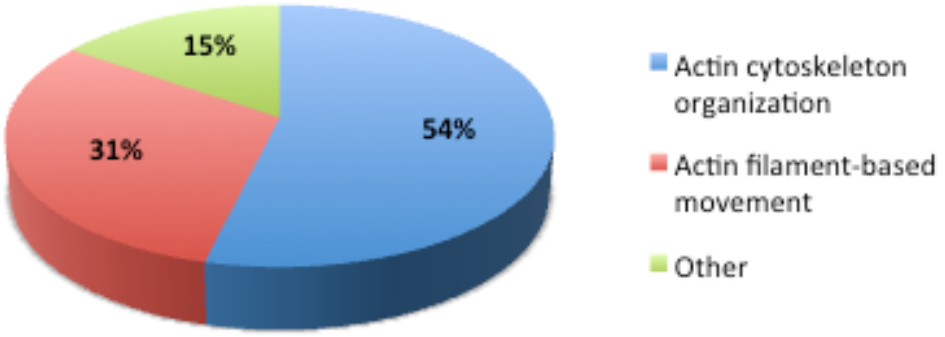


Figure 3.19: Functional classification of GAA putative interactors.

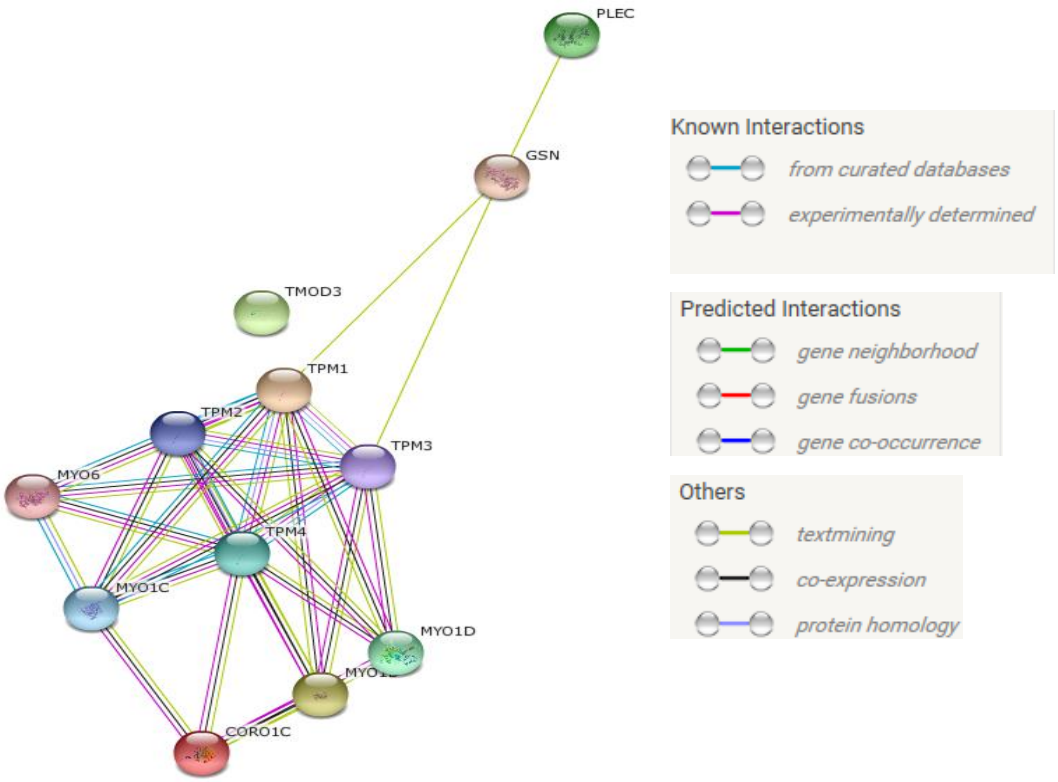


Figure 3.20: "Cytoskeleton based processes" obtained by analyzing all GAA interactome by String software. The interaction among all proteins belonging to the present network are indicated with different color code.

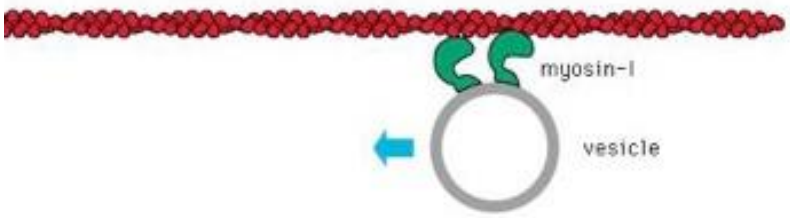


Figure 3.21: Unconventional Myosin along actin cytoskeleton.

Among all unconventional myosin identified as GAA interactors, *Myo1C* captured our attention because it has been found to be involved in the exocytosis of Glucose Transporter 4 (GLUT4), a glycoprotein that maintains glucose homeostasis by regulating protein translocation. In muscle and adipocytes, insulin stimulates the glucose uptake by activating the signalling for the translocation of GLUT4 from intracellular membranes to the cell periphery [17]. The translocation of GLUT4 occurs on actin cytoskeleton and its rapid movement along linear tracks is mediated by Myo1C motor, which was found to be associated to GLUT4 vesicles and its presence was demonstrated to be essential for the terminal events of GLUT4 vesicle exocytosis [18]. Other GAA interacting proteins affect GLUT4 transport process: *Tropomodulin-3 (TPMOD3) and Tropomyosin 3 (TPM3)*. TPMOD3 is a pointed-end actin-capping protein, which promotes the insulin-dependent GLUT4 translocation and exocytosis through an actin remodelling [19].

By co-immunoprecipitation experiment and western blot analysis, we demonstrated that Myo1C interacts with the GAA 110kDa precursor (Figure 3.12). This data suggests that the interaction between GAA and Myo1C (and probably TPMOD3) occur at the earliest phase of enzyme translocation, prior to the beginning of the proteolytic processing. The complex involving Myo1C and TPMOD3 might favour the movement of the GAA precursor from ER lumen to the Golgi apparatus, along actin cytoskeleton, in a process similar that involving GLUT4. The 100kDa GAA precursor interacts also with *TPM3,* as demonstrated by co-IP. TPM3, as well as *TPM1, TPM2,* and *TPM4,* belongs to tropomyosin family, actin-associated proteins that can form head-to-tail polymers along the major groove of F-actin. Different actin–tropomyosin interactions support a wide range of contractile and transport processes, each driven by either a specific member of the non-muscle myosin-2 family or an unconventional myosin.

TPM3 is reported to affect GLUT4 trafficking to the plasma membrane by balancing the amount of cortical F-actin that can interact with various myosin isoforms. Consistent with this, Tpm3.1 containing actin filaments promotes the engagement of non-muscle myosin-IIA (MyoIIA), generating a contractile force that may promote vesicle penetrance of the cortical mesh. Simultaneously TPM3.1 acts as Myo1C antagonist inhibiting its interaction with actin by competing with Myo1C for binding to actin filaments [20]. Consequently, the balance between

TPM3.1-containing and TPM3.1-free actin filaments determines the efficiency of movement of GLUT4 through the cortical actin network to plasma membrane (Figure 3.22) [21].

The interaction of GAA with both proteins suggests that they contribute to protein translocation along actin filaments as well as for GLUT4 traffic. Both proteins interact with the longest GAA form, suggesting that these competitive interactions occur at the earliest phase of enzyme translocation, prior or alternatively to the beginning of the proteolytic processing.

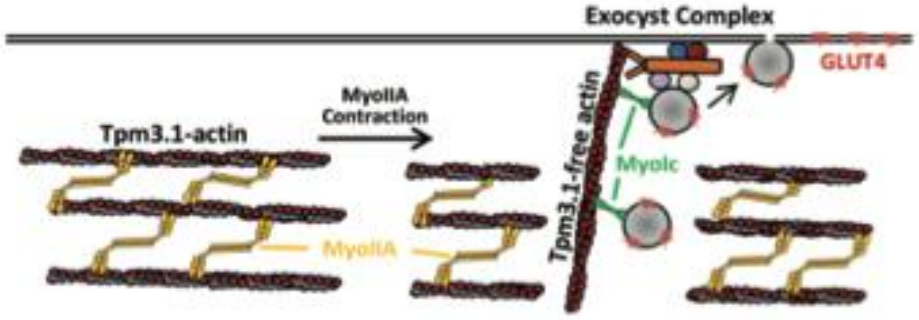


Figure 3.22: Schematic representation of how the balance between Tpm3.1/MyoIIA and Tpm3.1-free/Myo1c actin filament populations may determine the efficiency of movement and/or fusion of GLUT4 vesicles with the plasma membrane.

Differently, the first product of proteolytic maturation, the 95kDa isoform, was found to interact with Gelsolin (GSN), a calcium-, pH- and lipid-dependent actin severing and capping protein: its main function is to regulate the assembly of the actin cytoskeleton. GSN is composed of six conserved domains (S1–6) that have four distinct actin-binding sites, within S1, S2, S4 and S6. The increase of intracellular Ca^{2+} concentration induces conformational changes in GSN structure that allows the severing process to proceed. After severing, GSN is bound to the newly formed actin + end (barbed end) of one of the two shorter filaments. Further uncapping of actin filaments requires binding of gelsolin to phosphatidylinositol lipids (PPIs). This process exposes the barbed end for polymerization (Figure 3.23) [22][23][24].

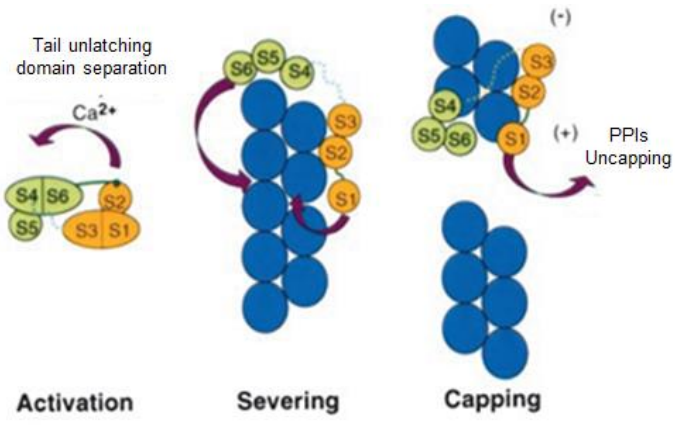


Figure 3.23: Schematic model for GSN binding, severing, and capping of an actin filament. Left, Gelsolin in the process of Ca^{2+} activation. Severing occurs when sufficient actin-actin bonds are broken. Right, the terminal actin in each strand is capped by GSN. PIP_2 induces Gelsolin to dissociate from the ends in a process called uncapping.

In addition, it was demonstrated that at low pH values, within the physiologically normal range, Gelsolin binds lipids in the absence of PPI becoming partially embedded in lipid vesicles and simultaneously vesicle-bound Gelsolin is also able to bind actin filaments [22].

Previous studies showed the Gelsolin involvement in Minute Virus of Mice (MVM)-induced actin-network remodelling and vesicular egress. Gelsolin activity is indispensable not only to trim actin filaments but also to drive virus export from perinuclear region to the periphery of infected cells. The delivery is mediated by vesicles bearing protein markers of lysosomes and/or late endosomes, suggesting that GSN may play a role in the assembly, loading, and/or trafficking of these vesicles [25].

These previous data suggest that, following the beginning of maturation process and formation of 95 kDa intermediate, the GAA might translocate by moving in vesicles along actin cytoskeleton associated to GSN, as confirmed in vivo by Proximity Ligand Assay (Figure 3.14).

Myo1B is involved in the traffic of cargo along the endocytic pathway, regulating the formation of post-Golgi carriers by modulation of actin assembly at trans-Golgi network (TGN). In this region, moreover, Myo1B co-localized with CI-mannose-6-phosphate receptor (CI-MPR) that carries cargos from TGN to sorting endosomes, transports lysosomal enzymes to lysosomes and recycles back to TGN. It was demonstrated that Myo1B expression regulates the cellular distribution and TGN exit of CI-MPR. Indeed, the depletion of Myo1B inhibits CI-MPR the delivery of lysosomal hydrolases is affected, as demonstrated for β–hexosaminidase, one of the lysosomal enzymes [26]. The interaction between GAA and Myo1B might have a potential role, even indirectly, in GAA trafficking from TGN to lysosomes.

Another identified unconventional myosin among GAA interactors was *Myo6*, a multifunctional motor protein involved in a wide variety of intracellular processes including cargo sorting and membrane delivery in the endocytic pathway, Golgi network organization and exocytosis along actin cytoskeleton [27]. Previous studies demonstrated that Myo6 participates to preliminary phases of autophagy, by mediating the delivery of endocytic cargo to autophagosomes and facilitating the autophagosome maturation and fusion with the lysosomes [28].

Autophagy is the major intracellular degradation system by which cytoplasmic materials are delivered to and degraded in the lysosome, contributing to the maintenance of cellular homeostasis in physiological conditions. A crucial step in this degradation pathway is the fusion of the double-membrane autophagosomes with lysosomes, where they deliver their contents for degradation. Myo6 plays an important role in this process; indeed its lack leads to an accumulation of autophagosomes due to a lysosomal fusion defect [28].

Co-immunoprecipitation and western blot analysis showed that Myo6 interacts selectively with the 70 kDa mature form of GAA (Figure 3.12).

The functional proteomic approach employed for identification of GAA interactome, showed that that the enzyme interacts with actin cytoskeleton associated proteins, such as several unconventional myosins, tropomyosins, tropomodulins and gelsolin. In addition, co-IP experiments and western blot assays allowed to identify which GAA isoform interacts with each protein partners.

On the basis of all these data and considering that GAA maturation occurs in specific cell compartments, it is possible to speculate about the route and the proteins accompanying GAA in the path from the endoplasmic reticulum to lysosomes (Figure 3.24).

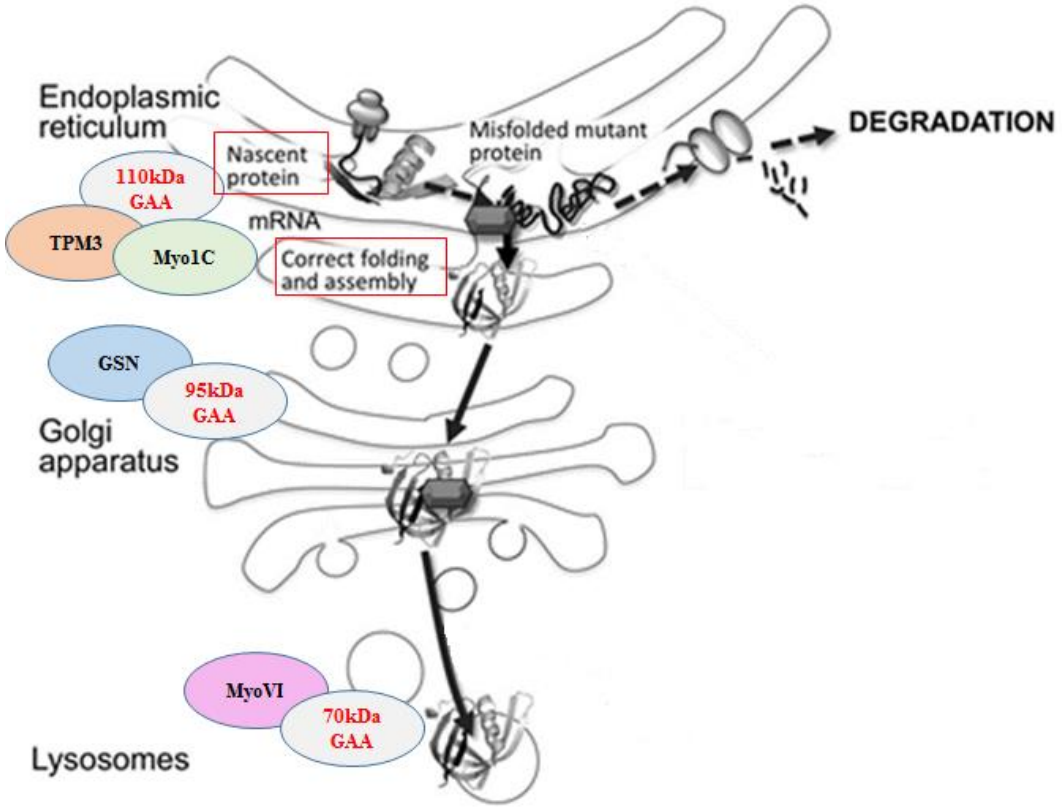


Figure 3.24: Representation of how GAA traffic from ER to lysosomes might occur in presence of confirmed protein partners.

Myo1C and TPM3 interact with the GAA 110 kDa precursor, existing in ER lumen and as demonstrated for GLU4 trafficking, it is possible to hypothesize that these interactions are alternative to one another. The exit of 95 kDa GAA isoform from TGN might involve GSN, a protein controlling actin polymerization and that might have a key role in the early stages of GAA trafficking toward lysosomes. Intact this does not occur in presence of L552P-GAA variant. The loss of interaction with GSN might lead to consequent accumulation of the variant in ER, from

where it is redirected towards degradation by ER quality control machine. In order to validate this hypothesis, immunofluorescence experiments are still in progress and the creation of a mouse knock out for GNS for functional studies in vivo. The occurrence of the interaction between GSN and α-galactosidase A, another lysosomal enzyme involved in Fabry disease confirms the cruciality of this protein in the general traffic process of many enzymes that have to be addressed to lysosomes. If this hypothesis will be confirmed by further experiments, the restoring of interaction with GSN might constitute a target for the development of broad-spectrum therapies for LSDs.

Finally, the delivery of GAA to lysosomes might be controlled by Myo6, as demonstrated by its interaction with the 70 kDa active form of GAA, localised specifically in lysososmes.

3.4.2 Investigation of rhGAA Interactome

Proteomics analysis was carried out in PD fibroblasts treated for 2 and 6 hours with exogenous recombinant enzyme rhGAA, Myozyme drug. Putative interactors were classified according to their main biological function (Figure 3.25-26).

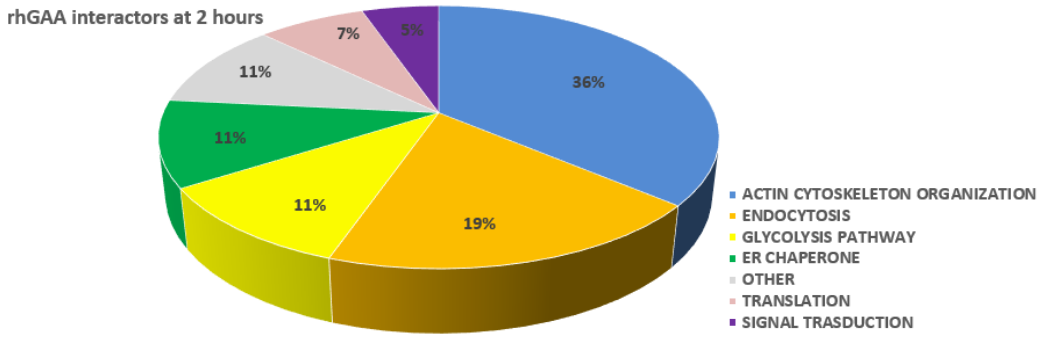


Figure 3.25: Functional classification of rhGAA putative interactors identified at 2 hours of treatment with Myozyme**.**

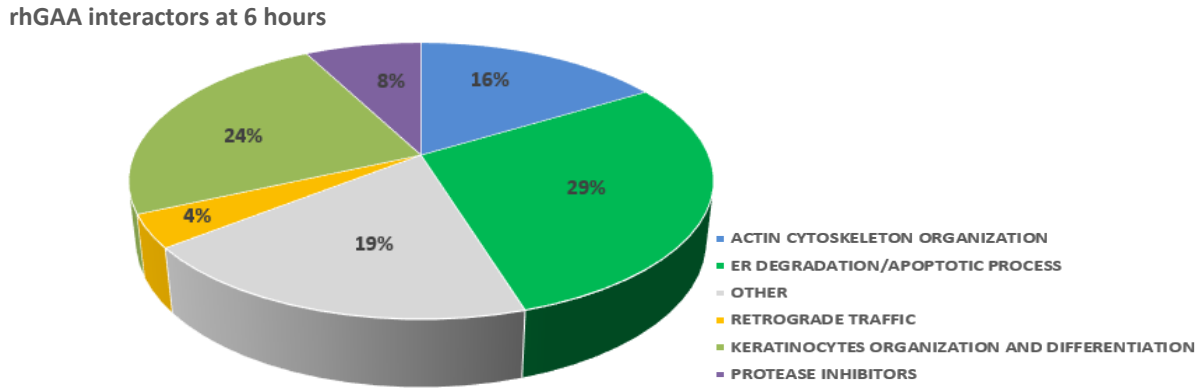


Figure 3.26: Functional classification of rhGAA putative interactors identified at 6 hours of treatment with Myozyme**.**

After 2 hours of treatment, several proteins involved in vesicular trafficking processes within endocytic pathway, i.e. *Clathrin heavy chain 1* and *Dynamin-2* have been identified, as well as a

large number of proteins involved in actin cytoskeleton rearrangement and polymerization, a process essential for the internalisation of exogenous molecules.

Endocytosis is the process through which the cell incorporates external macromolecules by invaginations of plasma membrane, resulting in the production of vesicles able to fuse with endosomes and inserting into the endo-lysosomal membrane system. Endocytosis also regulates the recycling of plasma membrane lipids and trafficking proteins, and the uptake or degradation of cell-surface receptors [29].

In the endocytosis mediated by Clathrin-coated vesicles (CCVs) (Figure 3.27), the vesicles assemble at the plasma membrane, select the load, invaginate and then are detached. Moving into the cell, the vesicle lose their coat and, once uncoated, they fuse with the early endosome to deliver their content [29][30].

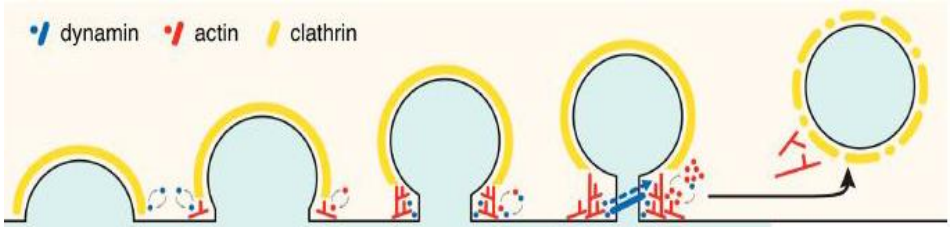


Figure 3.27: Role of Dynamin in Clathrin-coated vesicles formation.

In mammals, endocytosis process involves actin cytoskeleton, whose recruitment at the nascent vesicle favours the formation and internalization of most of the "coated pits" and mature vesicles through the development of dynamic polar filamentous structures (F-actin), which form an important part of the cytoskeleton. F-actin associates with numerous proteins that regulate the rapidity of assembly and disassembly promoting both the closing and movement of the vesicle through an effect called "comet-like tail" whose driving force is related to hydrolysis of ATP (Figure 3.28). Dynamin, a GTPase that mediates plasma membrane fission during clathrin-mediated endocytosis through the energy generated from the hydrolysis of GTP, acts on the generation, length and regularity of the movement of comet tails that move the intracellular vesicles [31].

Figure 3.28: The effect of "comet-like tail" of actin in the formation and movement of coated clathrin vesicles.

Dynamin is both involved at an early stage of endocytosis, by promoting the detachment of the vesicles from plasma membrane, and in a later stage by participating in the assembly of proteins involved in actin polymerization and "comet tail" generation [32].

Many protein partners play an important role in cytoskeleton organization, such as *Palladin (PALLD), Myosin-10 (MYH10), Filamin-B (FLNB), Cytospin-A (SPECC1L)*, *Fascin (FSCN1), Actin filament-associated protein (AFAP1), Ras GTPase-activating-like protein IQGAP1 (IQGAP1).* In addition, *Moesin (MSN),* a member of the Ezrin protein family, has been found to act as linkers between the plasma membrane and the cytoskeleton, playing an important role in connections of major cytoskeletal structures to the plasma membrane [33].

These proteins have been identified at 2 hours of treatment, but not at 6 hours, suggesting that cytoskeleton reorganization is crucial for the first phase of internalisation.

Moreover, at shorter time of incubation, several members of Annexins family, *Annexin-1 -2* and *-6* were found among rhGAA protein partners.

Annexins are calcium- and phospholipid-binding proteins that have multiple roles in membrane traffic. In endocytosis, Annexins play a role increasing the efficiency of internalization process from plasma membrane regulated by calcium [34].

Previous studies showed that Annexins 2 and 6 interact with both mu1 and mu2 subunits of clathrin assembly proteins complex AP-2. The mu2 subunit is mainly involved in trafficking events at plasma membrane, and mu1 controls movements of vesicles between internal compartments in the endocytic pathway. These components of Annexin family might act a level of both processes, controlling the entrance of rhGAA and its translocation inside the cell. In contrast, Annexin 1 is only involved in trafficking between intracellular compartments [35].

Annexin 2 has also been found to interact with the large protein *AHNAK*, another rhGAA partner at 2 hours of treatment. Recruiting AHNAK, Annexin 2 might play the role of scaffold protein linking plasma membrane domains with the actin polymerization machinery [36]. In addition, when Annexin 2 localized at early endosomes, it exerts an essential role in multi vesicular endosome/bodies (MVBs) formation. MVBs are also known as endocytic carrier vesicles and transport material from early to late endosomes. The so-called early MVBs have few internal vesicles and still contain recycling proteins such as transferrin receptor, while late MVBs have accumulated more internal vesicles and recycling proteins have been largely removed. Given the ability of Annexin 2 to interact with actin, its effects on MVB formation could be mediated through the actin cytoskeleton. Annexin 2 plays a role in regulating the morphology of the early endosome such that its depletion can affect both stable vacuole/MVB formation and efficiency of recycling.

Also Annexin 1 was found associated with MVBs. It can mediate inward vesiculation and might act in closing the invagination and/or vesicle scission.

Annexin 6 has been localized to many places throughout the cell including the plasma membrane and late endosomes, where plays a role in endosome–lysosome fusion (Figure 3.29) [34].

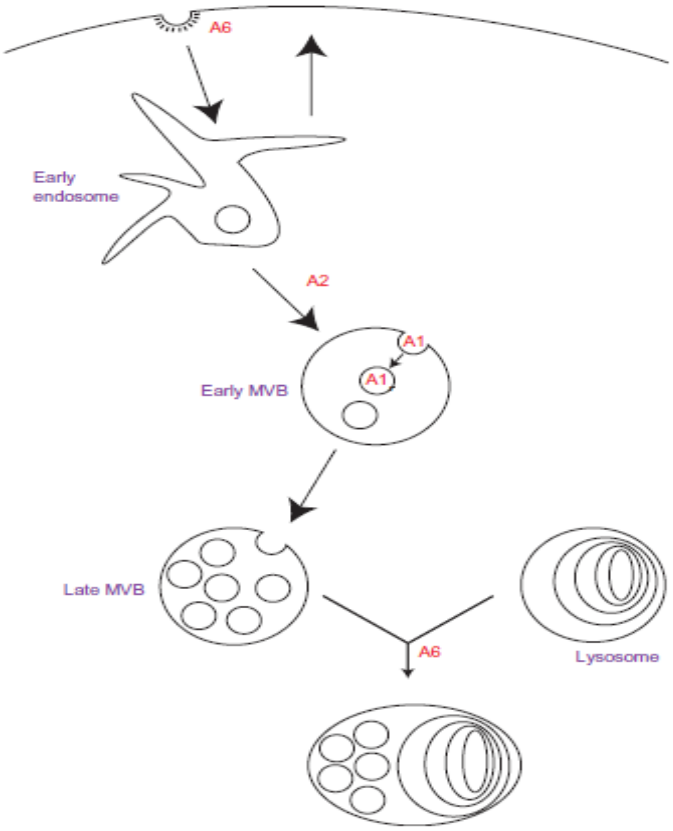


Figure 3.29: Possible sites of action of Annexins on the endocytic pathway.

Among rhGAA interactors there is *Perilipin-3*, also called *TIP47*. It is a 47kDa protein, that binds selectively to the cytoplasmic domains of cation-independent and cation-dependent Mannose 6-phosphate receptors (MPRs). TIP47 is present in cytosol and on endosomes and it is necessary for MPR transport from endosomes to the trans–Golgi network. MPRs transport newly synthesized lysosomal hydrolases from the Golgi to prelysosomes and then return to the Golgi for another round of transport. In order to escape lysosomal degradation, MPR would need to be retrieved from this prelysosomal compartment. TIP47 binds to MPR cytoplasmic domains and thereby triggers their efficient inclusion into newly forming transport vesicles departing from endosomes [37].

Moreover, studies have demonstrated that Dynamin has a role in late endosome dynamics and trafficking of the CI-MPR. It could be involved in the recycling of CI-MPR from endosomes to

the TGN, probably by playing a crucial role in the generation of recycling vesicles from tubular processes of endosomes [38].

Among the identified cytoskeleton proteins there are some unconventional myosins, Myo1B, Myo1C, Myo1D, involved in vesicle trafficking already present in endogenous GAA partners list. These interactions suggest that once rhGAA is internalized into the cell, the route to reach lysosomes is partially superimposable to the endogenous GAA pathway.

The above described proteins have been found as putative interactors of recombinant enzyme in the treatment for two hours. Their involvement in processes ranging from the formation of vesicles at plasma membrane, until the fusion of late endosomes into lysosomes suggests the path that rhGAA already accomplished in two hours by enzyme administration to reach to final fate, the lysosome, where the enzyme plays its physiological function.

The rhGAA interactors involved in the early phase of endocytic pathway and found at 2 hours of incubation with Myozyme drug, are absent at longer time of treatment. This suggests that after six hours the internalization process is almost complete.

Although it has been shown that following treatment with rhGAA, the formation of its mature forms takes place in a manner similar to the endogenous GAA, often the enzyme activity achieved is lower than the expected. The data obtained in these experiments might suggest a hypothesis to explain this partial failure.

As shown in the chart pie (Figure 3.25), 11% of rhGAA interactors are ER proteins, such as several *PDIs* and *Heat shock protein beta-1,* suggesting that a portion of recombinant enzyme reaches also this compartment, implying a not complete transport of rhGAA to lysosomes. The identification of *Endoplasmin*, a protein that operates in the endoplasmic reticulum associated to the quality control and degradation machinery (ERAD), suggests that the fate of a portion of rhGAA might be addressed to degradation.

The identification of *Proteasome subunit beta type-6 e Proteasome subunit beta type-5*, at 6 hours of Myozyme treatment, confirms the previously hypothesis. Indeed, the ubiquitin-proteasome pathway is one of the systems employed by eukaryotic cells to accomplish the turnover of proteins and to degrade defective or damaged proteins. The retro-traslocation of an amount of internalised enzyme toward the ER and proteasome where it will take place the partial degradation of Myozyme could explain the unsatisfactory yield of rhGAA in terms of the total effectiveness in the ERT.

In addition, after longer treatment time, among the rhGAA interactors there are some proteins belonging to S100 family: *Protein S100-A7, Protein S100-A8* and *Protein S100-A9*. These proteins

are defined "calgranulins", reflecting their calcium-binding properties and high expression in granulocytes. Little is known about their biological function, but it has been demonstrated that S100-A8 and S100-A9 levels increase in cellular stress conditions and are often strongly expressed in inflamed tissues. These proteins are also associated with the activation of apoptotic processes that may contribute to lowering of ERT efficacy [39].

All together these data suggest the possible pathways followed by rhGAA, as described by the known functions and cell localization of its protein partners (Figure 3.30).

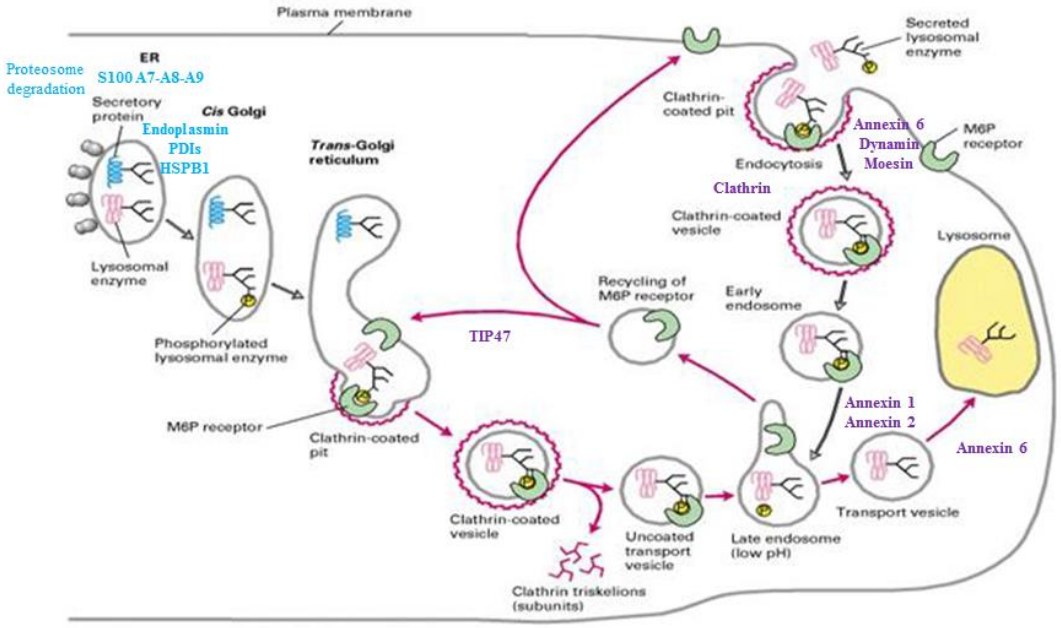


Figure 3.30: Hypotheses of rhGAA uptake pathways followed by Myozyme in PD fibroblasts.

In fibroblasts, the rhGAA uptake begins from the Myozyme recognition by mannose-6-phosphate receptor localized on the plasma membrane. The internalization process occurs through a clathrin-mediated mechanism, in which the plasma membrane invaginates and subsequent detachment of coated vesicle reaches completeness through Dynamin, Annexin 6 and actin activities. The largest amount of vesicles reaches the early endosomes, and the fusion with the late endosomes involves several annexins. The mannose-6-phosphate receptor is recycled to trans-Golgi by TIP47. This step might be the same by means an amount of internalized rhGAA retrotranslocates to the ER. At present, no other information is available to confirm this hypothesis or better describe how a portion of rhGGA may contact ER resident enzymes, such as Endoplasmin and several PDIs. Once in the ER, this amount of rhGAA does not pass the quality control and then it is addressed to the proteasome degradation. Further investigations, including co-IP experiments for each relevant rhGAA interactor, will be performed in order to confirm and support our hypothesis about rhGAA uptake process and its routes once internalized in the cell, and to define the maturation process steps of exogenous protein once internalized.

3.5 Conclusions

Proteomics results have showed that, in physiological condition, endogenous GAA interacts with different cytoskeleton proteins and in particular, the interactions are isoforms specific, suggesting that they might occur at different step during the traffic of GAA from endoplasmatic reticulum to lysososmes.

Gelsolin, a validated GAA protein partner, has a key role in the early stages of GAA trafficking resulting a potential target for definition of new therapeutic strategy in Pompe disease.

Moreover, proteomics experiments carried out in PD fibroblasts treated for 2 and 6 hours with rhGAA have provided information about both early steps of internalization process and the route followed by the Myozyme towards lysosomes. Furthermore, our data suggest that the only partial effectiveness of rhGAA observed in patients treated with Enzyme Replace Therapy might be explained considering that an amount of internalized protein is destined to degradation.

3.6 References

[1] Parenti G, Andria G, Ballabio A, Lysosomal storage diseases: from pathophysiology to therapy. *The Annual Review of Medicine, Vol. 66:471-86, 2015.*

[2] Braulke T, Bonifacino JS, Sorting of lysosomal proteins. *Biochimica et Biophysica Acta, Vol. 1793: 605–14, 2009.*

[3] Reczek D, Schwake M, Schroder J, et al., LIMP-2 is a receptor for lysosomal mannose-6-phosphate independent targeting of β-glucocerebrosidase. *Cell; Vol. 131(4): 770–83, 2007.*

[4] Parenti G, Fecarotta S, Moracci M, Andria G., Pharmacological chaperone therapy for lysosomal storage diseases. *Future Medicinal Chemistry Vol. 6(9): 1031–45, 2014.*

[5] Ballabio A, Gieselmann V, Lysosomal disorders: from storage to cellular damage. *Biochimica et Biophysica Acta, Vol. 1793: 684-96, 2009.*

[6] Manganelli F, Ruggiero L, Clinical features of Pompe disease. *Acta Myologica Vol. XXXII: 82-4, 2013.*

[7] Dasouki M, Jawdat O, Almadhoun O, Pasnoor M, McVey AL, Abuzinadah A, Herbelin L, Barohn RJ, Dimachkie MM, Pompe Disease: Literature Review and Case Series. *Neurologic Clinics, Vol.32(3): 751-76, 2014.*

[8] Moreland RJ, Jin X, Zhang XK, Decker RW, Albee KL, Lee KL, Cauthron RD, Brewer K, Edmunds T, Canfield WM, Lysosomal acid α-glucosidase consists of four different peptides processed from a single chain precursor. *Journal of Biological Chemistry, Vol. 280, N. 8, 6780–91, 2005.*

[9] Slonim AE, Bulone L, Ritz S, et al., Identification of two subtypes of infantile acid maltase deficiency. *Journal of Pediatrics, Vol. 137:283–85, 2000.*

[10] Kishnani PS, Howell RR, Pompe disease in infants and children. *Journal of Pediatrics, Vol. 144: S35–S43, 2004.*

[11] Fukuda T., Ashley Roberts, Paul H. Plotz and Nina Raben, Acid alpha-glucosidase deficiency (Pompe disease). *Current Neurology and Neuroscience Reports, Vol. 7: 71–7, 2007.*

[12] Parenti G, Porto C, Della Casa R, Nuovi approcci terapeutici alle malattie da accumulo lisosomiale. *Prospettive in pediatria Vol. 42, N. 168: 209-18, 2012.*

[13] Fukuda T, Ewan L, Bauer M, Mattaliano RJ, Zaal K, Ralston E, Plotz PH, Raben M, Dysfunction of endocytic and autophagic pathways in a lysosomal storage disease. *Annals of Neurology, Vol. 59: 700–8, 2006.*

[14] Porto C, Cardone M, Fontana F, Rossi B, Tuzzi MR, Tarallo A, Barone MR, Andria C, Parenti G, The Pharmacological Chaperone N-butyldeoxynojirimycin Enhances Enzyme Replacement Therapy in Pompe Disease Fibroblasts. *Molecular Therapy, Vol. 17 N. 6, 2009.*

[15] Kroos ML, Pomponio RJ, van Vliet L, Palmer RE, Phipps M, Van der Helm R, Halley D, Reuser A, Update of the Pompe disease mutation database with 107 sequence variants and a format for severity rating. *Human Mutation, Vol. 29(6):E13-26. 2008.*

[16] Li J, Lu Q, Zhang M, Structural Basis of Cargo Recognition by Unconventional Myosins in Cellular Trafficking, *Traffic, Vol. 17 (8): 822-38, 2016.*

[17] Bose A, Guilherme A, Robida SI, Nicoloro SMC, Zhou QL, Jiang ZY, Pomerleau DP, Czech MP, Glucose transporter recycling in response to insulin is facilitated by myosin Myo1c. *Nature, Vol. 420: 821-4, 2002.*

[18] Holman GD, Sakamoto K, Regulating the motor for GLUT4 vesicle traffic. *Cell Metabolism, Vol. 8(5): 344–6, 2008.*

[19] Lim C, Bi X, Wu D, Kim JB, Gunning PW, Hong W, Han W, Tropomodulin3 is a novel Akt2 effector regulating insulin-stimulated GLUT4 exocytosis through cortical actin remodeling. *Nature Communication, Vol. 6: 5951, 2015.*

[20] Manstein DJ, Mulvihill DP, Tropomyosin-Mediated Regulation of Cytoplasmic Myosins. *Traffic, Vol. 17: 872–7, 2016.*

[21] Kee1 AJ, Yang1 L, Lucas1 CA, Greenberg MJ, Martel N, Leong GM, Hughes WE, Cooney GJ, James DE, Ostap EM, Han W, Gunning PW, Hardeman EC, An Actin Filament Population Defined by the Tropomyosin Tpm3.1 Regulates Glucose Uptake. *Traffic, Vol. 16: 691–711, 2015.*

[22] Méré J, Chahinian A, Maciver SK, Fattoum A, Bettache N, Benyamin Y, Roustan C, Gelsolin binds to polyphosphoinositide-free lipid vesicles and simultaneously to actin microfilaments. *Biochemical Journal, Vol. 386: 47-56, 2005.*

[23] P. Silacci, L. Mazzolaib, C. Gaucia, N. Stergiopulosa, H. L.Yinc, D. Hayozb, Gelsolin superfamily proteins: key regulators of cellular functions. *Cellular and Molecular Life Sciences, Vol. 61 (19-20): 2614-23, 2004.*

[24] Sun HQ, Yamamoto M, Mejillano M, Yin HL, Gelsolin, a Multifunctional Actin Regulatory Protein. *Journal of Biological Chemistry, Vol. 274 (47): 33179–82, 1999.*

[25] Bar S, Daeffler L, Rommelaere J, Nuesch JPF, Vesicular Egress of Non-Enveloped Lytic Parvoviruses Depends on Gelsolin Functioning. *PLoS Pathogens, Vol. 4, Issue 8, 2008.*

[26] Almeida CG, Yamada A, Tenza D, Louvard D, Raposo G, Coudrier E, Myosin 1b promotes the formation of post-Golgi carriers by regulating actin assembly and membrane remodelling at the trans-Golgi network. *Nature Cell Biology, Vol. 13 (7): 779-89, 2011.*

[27] Sahlender DA, Roberts RC, Arden SD, Spudich G, Taylor MJ, Luzio JP, Kendrick-Jones J, Buss F, Optineurin links myosin VI to the Golgi complex and is involved in Golgi organization and exocytosis. *Journal of Cell Biology, Vol. 169 (2): 285-95, 2005.*

[28] Tumbarello DA, Waxse BJ, Arden SD, Bright NA, Kendrick-Jones J, Buss F, Autophagy receptors link myosin VI to autophagosomes to mediate Tom1-dependent autophagosome maturation and fusion with the lysosome. *Nature Cell Biology, Vol. 14 (10): 1024-35, 2012.ù*

[29] Elkin SR, Lakoduk AM, Schmid SL, Endocytic pathways and endosomal trafficking: a primer. *Wiener Medizinische Wochenschrift, Vol. 166 (7-8): 196-204, 2016.*

[30] Smythe E, Ayscough KR, Actin regulation in endocytosis. *Journal of Cell Science, Vol. 119: 4589-98, 2006.*

[31] Sundborger AC, Hinshaw JE, Regulating dynamin dynamics during endocytosis. *F1000 Prime Reports, Vol. 6: 85, 2014.*

[32] Khaitlina SY, Intracellular transport based on actin polymerization. *Biochemistry, Vol. 79, No. 9: 917-27, 2014.*

[33] Vaheri A, Carpén O, Heiska L, Helander TS, Jaaskelainen J, Majander-Nordenswan P, Sainio M, Timonen T, Turunen O, The ezrin protein family: membrane-cytoskeleton interactions and disease associations. *Current Opinion in Cell Biology, Vol. 9: 659-66, 1997.*

[34] Futter CE, White IJ, Annexins and Endocytosis. *Traffic, Vol. 8: 951–8, 2007.*

[35] Creutz CE, Snyder SL, Interactions of annexins with the mu subunits of the clathrin assembly proteins. *Biochemistry, Vol. 44: 13795–806, 2005.*

[36] Hayes1 MJ, Rescher U, Gerke V, Moss SE, Annexin–actin interactions. *Traffic, Vol. 5: 571–6, 2004.*

[37] Dıaz E, Pfeffer SR, TIP47: a cargo selection device for mannose-6-phosphate receptor trafficking. *Cell, Vol. 93: 433–43, 1998.*

[38] Nicoziani P, Vilhardt F, Llorente A, Hilout L, Courtoy PJ, Sandvig K, van Deurs B, Role for dynamin in late endosome dynamics and trafficking of the cation independent mannose-6-phosphate receptor. *Molecular Biology of the Cell, Vol. 11: 481–95, 2000.*

[39] Goyette J, Geczy CL, Inflammation-associated S100 proteins: new mechanisms that regulate function. *Amino Acids, Vol. 41: 821–42, 2011.*

Chapter 4 –Investigation of Molecular Mechanisms Impaired in Amyloidosis

4.1 Introduction

4.1.1 Protein Misfolding Diseases: Amyloidosis

The amyloidoses are a large group of diseases caused by extracellular deposition of insoluble protein aggregates, called *amyloid*, derives from misfolding of several precursor proteins that aggregate in highly ordered abnormal cross β-sheet conformation.

In the cell, the folding process of a newly synthesized protein occurs through a rapid sequence of conformational changes. According to the *folding energy landscape theory*, the process follows a funnel-like pathway in which the conformational intermediates progressively merge into a final species (Figure 4.1) [1].

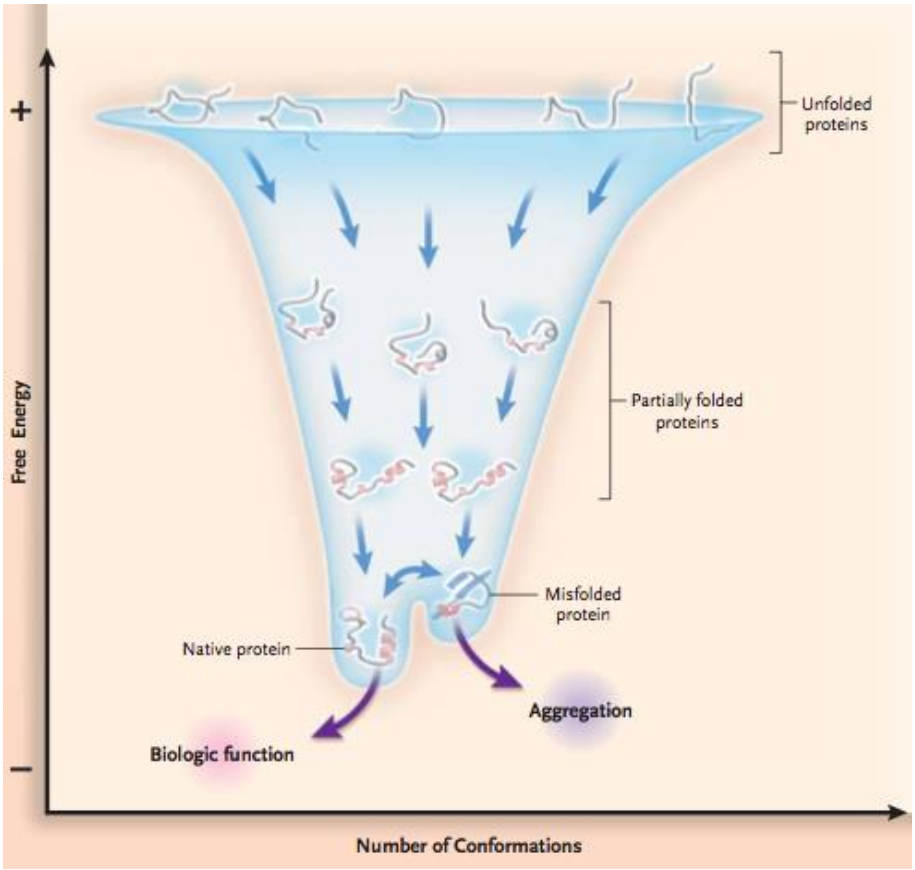


Figure 4.1: The process of protein folding: the unfolded polypeptide enters a funnel-like pathway in which the conformational intermediates become progressively more organized, resulting in the most stable native state. In this state, there is a minimum of free energy, which results from the balance between the level of enthalpy, the internal energy that in folded protein is mainly determined by the kind and number of intramolecular bonds, and the level of conformational entropy, the level of randomness of the polypeptide in solution.

The folding process is subjected to a quality control by molecular chaperones and protease machinery.

Misfolded proteins are either rescued by chaperones or eliminated by the ubiquitin-proteasome system or lysosome-mediated autophagy [2]. These mechanisms are normally sufficient to prevent the accumulation of misfolded proteins, but under certain pathological conditions, the system of intracellular quality control appears to be incapable of recognizing and removing dangerous misfolded proteins.

Once the folding process has been completed and the native proteins overcome the ER quality control, they are addressed to intra o extra-cellular places where accomplish their biological functions. All native proteins remain in dynamic equilibrium with a partially unfolded state, which is favourited in presence of genetic mutation and or in presence of abnormal environmental conditions.

These pathological conditions are known as *protein misfolding diseases* (or *protein conformational disease*) because the pathological protein undergoes structural changes that trigger self-association, aggregation and tissue deposition, causing severe functional impairment [3]. These structures are defined as amyloid fibrils (or amyloid plaques) when they accumulate in the extracellular space, whereas the term "intracellular inclusions" has been suggested as more appropriate when fibrils accumulate inside the cell [4].

The molecular mechanisms involved in fibrillogenesis, are not clear yet. The knowledge of how protein aggregation process occurs requires the characterization of the conformational changes and of oligomeric structures adopted by the polypeptide chain during the process.

The identification of residues and/or regions of the amino acidic sequence that promote the aggregation is also determinant for the comprehension of the process and for the perspective of the development of new therapeutic strategies.

It is known that the formation of amyloid fibrils is a polymerization nucleation-dependent process (Figure 4.2). The time course of the conversion of a peptide or protein into its fibrillar form typically includes a *lag phase*, the rate-limiting step, followed by a rapid exponential *growth phase*. The lag phase represents the time required to form"nuclei". Once a critical nucleus has been generated, it represents a "seed" which accelerates the growth of fibril with very fast kinetics. Any available monomers in the amyloidogenic conformation quickly become trapped in the fibril. As in other processes dependent on a nucleation step, the addition of preformed fibrillar species to a sample protein under aggregation conditions (seeding) determines the shortening of the lag phase [1][5].

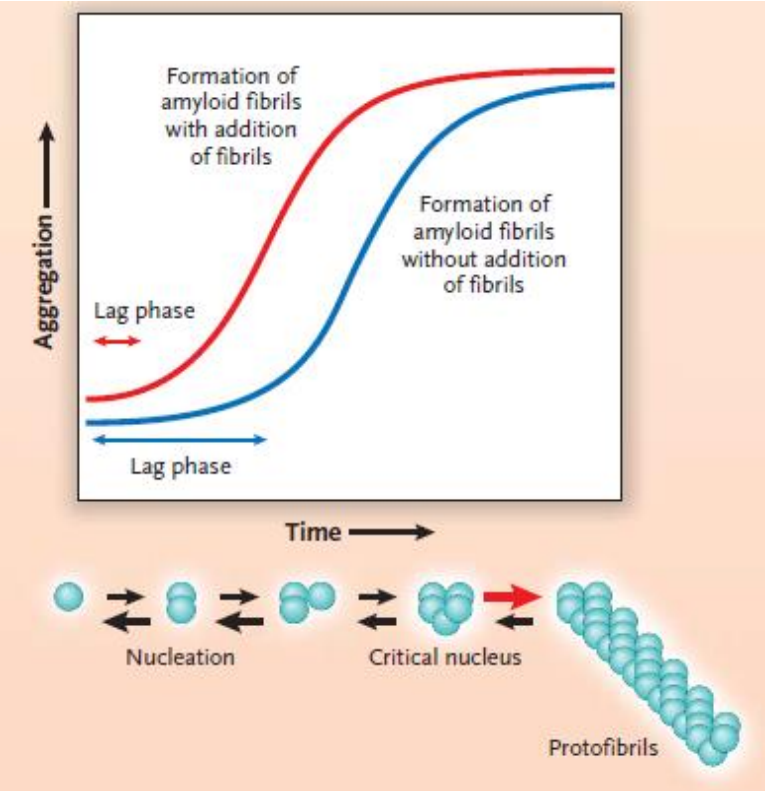


Figure 4.2: Kinetics of amyloid fibril formation: the process is described by a sigmoidal curve. The blue line indicates the formation of amyloid fibrils, beginning from a solution of monomeric proteins, which may temporarily assume an amyloidogenic conformation. Following rate-limiting step (lag phase) aggregate growth proceeds rapidly by further addition of monomers to form larger polymers (plateau phase). The red line represents a similar condition in which preformed fibrils are added, thus making the lag phase much shorter.

As mentioned before, the occurrence of specific environmental conditions, the achievement of a higher local concentration, the interaction with extracellular matrix might synergistically induce the deposition of amyloid plaques.

Both *in vivo* and *in vitro* fibrils when stained by Congo red give rise to a specific apple-green birefringence under a polarized light microscope; this flurescence is so specific that constitutes the preliminary method for amyloidosis diagnosis. The presence of rigid, non-branching fibrils is also detectable by electron microscopy and atomic force microscopy (Figure 4.3) [1].

All amyloid deposits contain an amount of heparansulfate and dermatan sulphate proteoglycans and glycosaminoglycan chains, some of which are strongly bound to the fibrils. These glycan molecules may contribute to amyloid fibrillogenesis as well as the stabilization of fibril structure. Heparansulfate proteoglycans are components of the extracellular matrix, as well as perlecan, laminin, entactin, and collagen IV. These molecules can have a scaffold role, facilitating the initial phase of fibril nucleation, and could have a targeting role in the localization of amyloid deposits in tissue. Ubiquitous components of amyloid deposits are the glycoprotein Serum Amyloid P component (SAP) and other proteins, such as Apolipoprotein E and Apolipoprotein A-IV, which are considered as amyloid protein signatures [1][6].

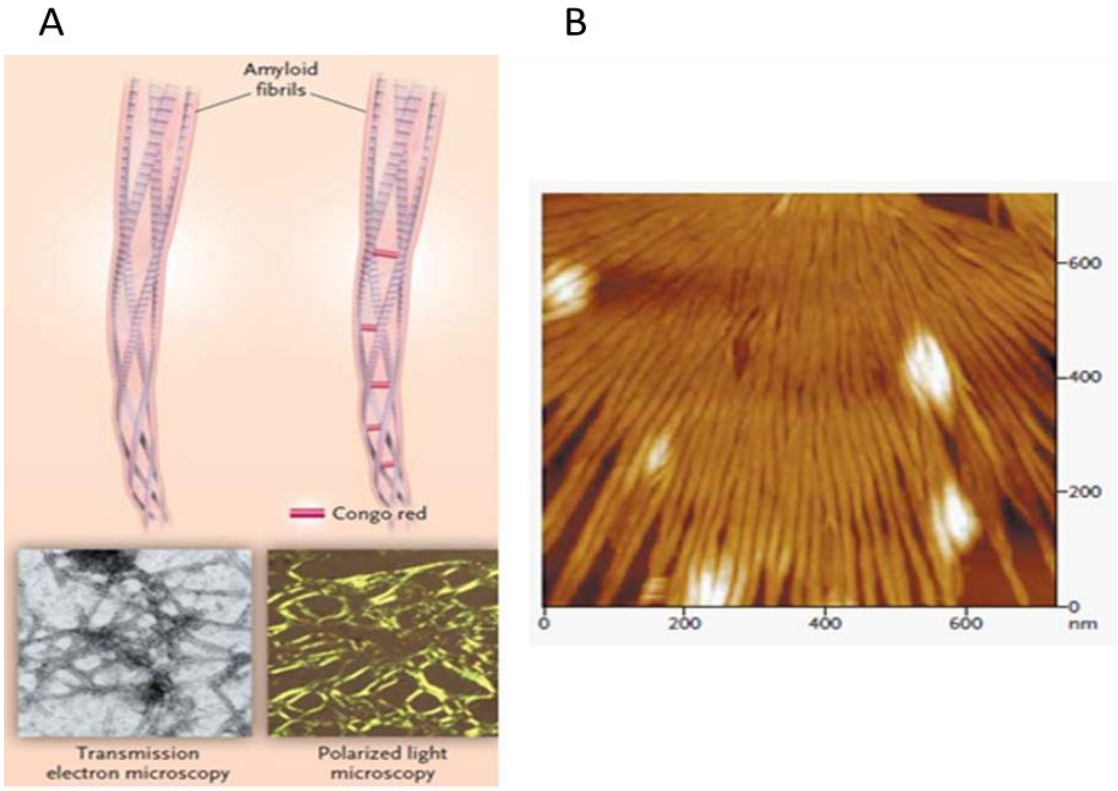


Figure 4.3: Image of amyloid fibrils analysed by Congo red, electron microscopy (A) and atomic force microscopy (B).

Amyloidosis may be systemic disease and can arise either spontaneously or in a hereditary fashion. Pathogenic misfolded proteins can form in different ways. There are some proteins that have an intrinsic propensity to assume a pathologic conformation, which becomes evident with aging, as well as normal transthyretin (TTR) in patients with senile systemic amyloidosis, or at persistently high concentrations in serum, as well as β-2-microglobulin in patients undergoing long term haemodialysis.

In addition, proteolytic remodelling of protein precursor can lead to fibrils formation, as in the case of β-amyloid precursor protein (APP) in Alzheimer's disease.

Hereditary amyloidosis is caused by mutant genes encoding variant proteins whose structure makes them amyloidogenic. The most common causes of hereditary amyloidosis (systemic amyloidosis) are associated to mutations in genes encoding TTR, Apolipoproteins A-I and A-II, Fibrinogen A α-chain, Gelsolin, and Lysozyme [1].

In some cases, amyloidosis occurs secondary to an underlying haematological disorder or chronic inflammation or infection: for istance light chain (AL) amyloidosis is often associated to multiple myeloma, as well as the serum amyloid A (AA) amyloidosis is associated to inflammation [6]. Almost 30 amyloidogenic proteins have been identified, some of which are shown in Table 4.1 [1].

Table 4.1: The most common amyloid proteins, their pathogenic precursors, distribution and type of amyloidosis.

AMYLOID PROTEIN	PRECURSOR	DISTRIBUTION	TYPE
AB	Aβ protein precursor	Localized	Acquired
		Localized	Hereditary
APRP	Prion protein	Localized	Acquired
		Localized	Hereditary
ABRI	ABri protein precursor	Localized or systemic	Hereditary
ACYS	Cystatin C	Systemic	Hereditary
AB2M	β 2-microglobulin	Systemic	Acquired
AL	Immunoglobulin light chain	Systemic or localized	Acquired
AA	Serum amyloid A	Systemic	Acquired
ATTR	Transthyretin	Systemic	Hereditary
		Systemic	Acquired
AAPOAI	Apolipoprotein A-I	Systemic	Hereditary
AAPOAII	Apolipoprotein A-II	Systemic	Hereditary
AGEL	Gelsolin	Systemic	Hereditary
ALYS	Lysozyme	Systemic	Hereditary
AFIB	Fibrinogen A α-chain	Systemic	Hereditary

4.1.2 Examples of Systemic Amyloidosis: Clinical Features, Diagnosis and Therapies

AL amyloidosis **–** Immunoglobulin light chain (AL) amyloidosis is characterized by a clonal population of bone marrow plasma cells that produces a monoclonal light chain of lambda or kappa type as either an intact molecule or a fragment. Light chain kappa or lambda fibrils accumulate in tissues and interfere with organs function [7].

AL amyloidosis is the most commonly diagnosed form of cardiac amyloidosis. One or many organs may be involved: heart, kidneys, liver, peripheral and autonomic nervous systems and soft tissues [8]. Clinical features depend on organs involved, but can include restrictive cardiomyopathy, nephrotic syndrome, hepatic failure, peripheral/autonomic neuropathy, and atypical multiple myeloma [7].

The chemotherapy is used for the treatment of AL amyloidosis. In addition, most patients with cardiac AL amyloidosis receive a cyclic combination of the alkylating agent cyclophosphamide, the proteasome inhibitor bortezomib and dexamethasone [7][8].

ATTR amyloidosis – Transthyretin (TTR) is homo-tetrameric plasma protein produced almost exclusively by the liver, but also in small amounts within the choroid plexus andretina. Its physiological function is transportation of retinol-binding protein and thyroxine. Normal TTR (ATTRwt) is an inherently amyloidogenic protein. ATTRwt amyloidosis is almost exclusively a disease that affects elderly individuals (more than 70 years of age) with a 10-fold male predominance. In most patients, the heart is the only vital organ that is clinically affected, but carpal tunnel syndrome is common and may precede the development of cardiac symptoms by 10 years [8].

Hereditary ATTR amyloidosis is usually associated with a single amino acid substitution caused by a point mutation in the TTR gene. Amyloidogenic TTR variants are less stable than their wild-type counterpart, underlying their increased propensity to form amyloid fibrils. To date, more than 120 point mutations in TTR gene have been identified, most of which are associated with amyloidosis. The most common mutations causative of hereditary ATTR amyloidosis are Val30Met, Thr60Ala, Ser77Tyr and Val122Ile, and their occurrence is strongly dependent by geographical area.

So far the prevalent therapeutic treatment for ATTR amyloidosis include liver or heart transplantation in the cases of significant cardiac involvement from either ATTRwt or ATTR variant. In these last years several alternative therapeutic methods addressed to the synthesis of molecules able to stabilize TTR tetramer, or to inhibit hepatic synthesis of both wild-type and variant TTR, or to inihibit/destroy fibril formation have been developed to tackle TTR amyloidosis [8][9].

AApoAI amyloidosis - Apolipoprotein A-I (ApoAI) is the main component of high density lipoproteins (HDL) and is mostly involved in the removal of cholesterol from peripheral tissues and its transfer, via the plasma, to the liver, where it is either recycled back to plasma as a component of newly formed lipoproteins, or excreted from the body via bile [10].

ApoAI amyloidosis can be present as a non-hereditary form with wild-type protein deposits in atherosclerotic plaques or as a hereditary form with the variant protein depositing more systemically in specific target organs, depending by the nature and localization of mutation [11]. For example, L75P-AApoAI amyloidosis is characterized by preferential accumulation of fibrils in the kidneys and liver, while fibrils from L174S-AApoAI occur preferentially in the heart, skin, testes and larynx [12][13].

Patients affected by hereditary ApoAI amyloidosis typically present hypertension, proteinuria, and renal impairment and frequently develop extensive visceral deposits that affect the liver, spleen, and kidneys, with occasional involvement of the nerves, larynx, and gastrointestinal tract [14].

4.1.3 About Amyloidosis Diagnosis Issues

Regarding the therapeutic strategies, an accurate definition of amyloid type is a crucial step in the diagnosis of systemic amyloidosis since the amyloid type dictates the therapy [15]. In fact, a precise identification of the amyloid fibril protein is essential for administration of appropriate therapy underlined by the fact that AL amyloidosis is treated with systemic chemotherapy, which is not efficacious and may have deleterious effects in other forms of systemic amyloidosis [16].

Proteomic analysis of amyloid deposits provides a chemical characterization of fibrillar constituents, representing a complementary strategy to genetic sequencing and immunohistochemistry analysis. However, there are ambiguous cases, in which more than one potentially amyloidogenic protein is found within the patient's biopsy and an accurate diagnosis becomes challenging [17]. Typical examples are those in which both light chains and transthyretin are detected. This problem is becoming more common in aged individuals suffering of monoclonal gammopathies associated to a high secretion of monoclonal light chains: in several cases the amyloid plaques contain both proteins, the precursor of AL amyloid and the ATTR, leading ambiguous the diagnosed of amyloidosis [18]. Indeed, a recent study showed that up to 20% of patients with ATTR amyloidosis have an incidental monoclonal protein in serum or urine [19]. In addition, the increasing use of fine needle aspiration fat biopsies aggravates the issue for unambiguous typing of amyloid. Fine needle fat aspiration has the advantages to be a safe, simple and non-invasive technique to extract accumulated amyloid from tissue. The disadvantage is that the proteome of fine needle aspiration biopsy typically contains more than one potential amyloid fibril protein, thus complicating the achievement of a proper diagnosis.

4.1.4 Apolipoprotein A-I: From Physiological Function to Amyloid Fibrils Formation

ApoAI plays a critical role in lipid metabolism [20], both in delivering cholesterol to steroidogenic tissues and in transporting it from the periphery to liver for catabolism in reverse cholesterol transport. High density lipoproteins (HDL) particles containing ApoAI are antiatherogenic and protect against atherogenesis via reverse cholesterol transport and play an important role in antiinflammatory and antithrombotic response [21].

ApoAI protein is synthesized by the liver and the small intestine; the mature form of 243 amino acids (28 kDa) is secreted in the plasma, where it is either associated to lipids, or in a lipid-free state (5-10%) [11][22].

More than 50 ApoAI variants have been identified, and about half of them are associated with HDL reduced levels [11].

To date, 19 mutations are known to be associated with hereditary AApoAI amyloidosis (Table 4.2). Most of the mutants are generated by single point mutations, but deletions or deletions/insertions have also been described [23][24]. For these variants, amyloid fibrils isolated *ex vivo* were found to be mainly constituted by 90-100 residues long N-terminal fragments of ApoAI. Thus, ApoAI variants represent the precursors of N-terminal fibrillogenic fragments, which show a high propensity to aggregate.

Most of mutations are present within the N- terminal portion of the protein that is eventually found in fibrils (inside mutations). Other mutations occur in positions located outside this region of the polypeptide sequence (outside mutations) [14][25].

Table 4.2: The mutations occurring in ApoAI sequence, associated to amyloid pathology. The affected organs or tissues are also indicated.

INSIDE MUTATIONS	INVOLVED ORGAN OR TISSUE
G26R	Kidneys, liver, peripheral nerves, GI tract
E34K	Kidneys, liver
W50R	Kidneys, liver, GI tract
L60R	Kidneys, liver, testes, heart
L64P	Kidneys, liver
Δ60-71+ VT INS	Liver
Δ70-72	Kidneys, liver, choroid
F71Y	Liver, palate
N74K frameshift	Kidneys, uterus, ovaries, pelvic lymph nodes, GI tract
L75P	Kidneys, liver, testes
L90P	Skin, heart, larynx

OUTSIDE MUTATIONS	INVOLVED ORGAN OR TISSUE
ΔL107	Circulatory sistem
A154 frameshift	Kidneys
H155M frameshift	Kidneys
L170P	Larynx
R173P	Kidneys, skin, heart, larynx
L174S	Skin, testes, heart, larynx
A175P	Larynx, testes
L178H	Larynx, skin, heart, nerves

The structures proposed for ApoAI as lipid bound or lipid-free protein are different. In the absence of lipids, ApoAI assumes a compact four-helix bundle structure. In the lipid-bound form ApoAI exists as a dimer. It known that Δ(185-243)ApoAI C-terminally truncated forms a dimer composed of two antiparallel molecules adopting a semi-circular conformation, with a diameter comparable to that of HDL particle. The structure is stabilized by two symmetric four-helix bundles at the opposite ends of the dimer (Figure 4.4) [26].

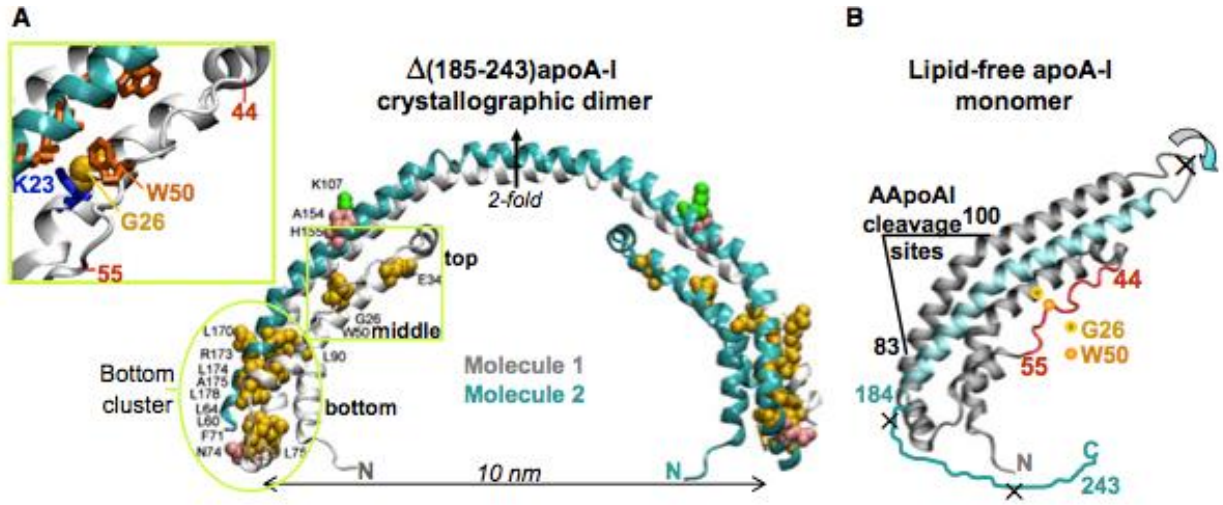


Figure 4.4: Locations of the AApoAI mutations in the structure of free ApoAI. (A) X-ray crystal structure ofC-terminally truncated human lipid-free Δ (185–243) ApoAI with the sites of all known AApoAI mutations mapped The structure is a crystallographic dimer stabilized by two symmetry-related four-helix bundles. (B) Proposed structure of lipid-free ApoAI monomer.

Recently, a model has been proposed, prompting that AApoAI mutations perturb one of the 4 predicted hot spots of the protein (*14–22*, *53–58*, *69–72* and *227–232*), with consequent full-length protein aggregation via the N-terminal region. Consequently, the proteolytic cleavage at the exposed loop (83–100) occurs and the N-terminal fibrillogenic polypeptide is released (Figure 4.5) [26]. The latter has been found to be the main protein constituent of ex-vivo fibrils [27]. It has been assumed that perturbation in the hot spots is one of the prerequisite for AApoAI misfolding, even though other factors may influence this process [26].

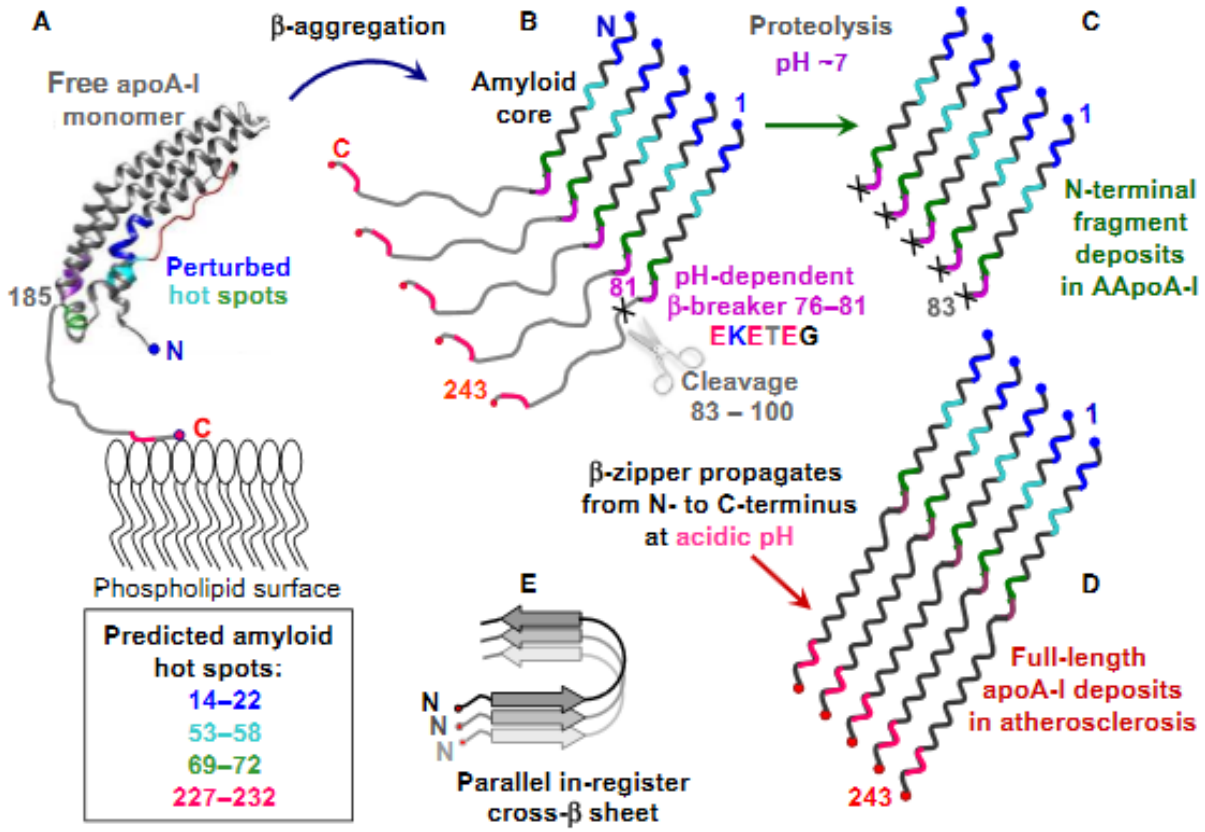


Figure 4.5: Proposed pathway of ApoAI misfolding in amyloidosis. (A) The presence of a mutation in ApoAI reduces its affinity for lipid and promotes the shift from HDL-bound to lipid-poor/free monomeric ApoAI. Once dissociated from HDL, ApoAI forms a four-helix bundle containing residues 1–184 (helix ribbon) followed by the partially unstructured C-terminal residues 185–243 (thin line) containing the primary lipid-binding site. Four predicted amyloid hot spots are color-coded. The single point mutations may perturb the native structure in these hot spots and allow β-aggregation. This leads to the formation of the protease-resistant amyloid core in the first 75 residues containing N-terminal amyloid hot spots (in blue, teal and green) (B). In the parallel in-register β -sheet predicted for ApoAI, each hot spot (color-coded) stacks in register against its counterparts from the adjacent molecules. At about pH 7, the β-breaking EKETEG residues 76–81 probably impede the amyloid core propagation towards the C-terminus and hamper the in-register stacking of the similarly charged groups. As a result, amyloid core stops at residues 76–81, leading to protein cleavage shortly there after. This explains why amyloid core in the patient-derived AApoAI deposits contains residue segments 1–83 to 1–100 (C). (D) At low pH, partial protonation of the adjacent Glu is expected to diminish the β -breaking potential of the EKETEG motif. This facilitates the β -sheet propagation towards the C-terminus and the formation of fibrils containing full-length ApoAI, such as those found in atherosclerotic plaques. In the resulting parallel in-register β -sheet structure, different hot spots within the molecule act in synergy (D), whereas similar secondary structural elements from different molecules (in different shades of gray) (E) are stacked in register.

4.1.5 Aims of the Project

In the field of systemic amyloidosis, the present research project was focused on two main goals:

1. The investigation of the effects of "inside mutation" L75P and the "outside mutation" L174S on AApoAI propensity to undergo an amyloidogenic pathway in order to elucidate the effect of two different mutation on the protein stability. To this end, a complementary proteolysis approach, coupled with mass spectrometry (LC-MS), was performed.

This project was carried out in collaboration with Professor Renata Piccoli and Professor Daria Maria Monti of University of Naples Federico II.

2. The development of an innovative proteomic strategy for typing amyloid deposits. The procedure is based on methodologies of tissues decellularization addressed to an unambiguous identification by mass spectrometry (LC-MS/MS) of the fibrillogenic protein.

This research activity has been carried out at Wolfson Drug Discovery Unit, Centre for Amyloidosis and Acute Phase Proteins, Division of Medicine, University College London (Royal Free Campus), London, UK, during the last year of my PhD course, under supervision of Professor Vittorio Bellotti.

4.2 Investigation of ApoAI Conformational Perturbations in Hereditary Amyloidosis

4.2.1 Materials and Methods

4.2.1.1 Materials

Trifluoroacetic acid (TFA), trypsin, chymotrypsin, endoproteinase Glu-C used for proteolysis experiments were SIGMA- ALDRICH products.

Tris-HCl used to solubilize ApoAI protein was provided by BIO-RAD.

Formic acid and acetonitrile HPLC Grade used for LC/MS analysis were Baker products.

Professor D.M. Monti research group of University of Naples "Federico II" provided the recombinant his-tagged ApoAI proteins: wild-type, L75P, L174S ApoAI.

4.2.1.2 Methods

Complementary proteolysis experiments were carried out by using 0.4 mg/mL solutions of proteins in presence of trypsin (E:S= 1:8000), chymotrypsin (E:S= 1:8000), and endoproteinase Glu-C (E:S= 1:10000) in 50mM Tris-HCl, pH 7.4.

The proteolysis experiments were carried out at 37°C and monitored on a time-course basis by sampling the peptide mixture at 15 and 30 minutes and blocking the reactions with 2% trifluoroacetic acid (TFA).

Peptide mixtures were analyzed by liquid chromatography-mass spectrometry (LC–MS) onto a QuattroMicro LC- MS system (Micromass, Waters) interfaced with a 1100 HPLC (Agilent Technologies, Palo Alto, CA). Peptide mixtures from the different proteolysis experiments were fractionated by reverse-phase HPLC on a Phenomenex Jupiter C18 column (250 mm × 2.1 mm, 300 Å pore size) and eluted by using a step gradient from 5% to 60% of solvent B (5% formic acid and 0.05% TFA in acetonitrile) over 60 min and from 60% to 95% in 5 min (solvent A 5% formic acid and 0.05% TFA in water). The flow rate was 200 μL/min and directly introduced in the mass spectrometer ESI source. Horse heart myoglobin was used as standard to calibrate the instrument (average molecular mass 16,951.5 Da) at 5 scans/s.

4.2.2 Results and Discussion

The aim of this study was to clarify the different impacts of two single point mutations, L75P and L174S, on ApoAI conformational stability by complementary proteolysis approach coupled with mass spectrometry identification of cleavage sites (LC–MS) [28].

Complementary proteolysis refers to limited proteolysis experiments carried out in conditions able to give rise to a single proteolytic event on the protein, which originates two "complementary" peptides. In these conditions, flexible and exposed regions will be preferentially hydrolysed instead of buried or rigid regions. Further cleavages occurring on the two complementary peptides are not considered informative for conformational analysis. By using an array of proteolytic enzymes with different specificity, it has been possible to discriminate exposed regions from the regions not accessible to proteases. Conformational changes induced by mutations and affecting protein topology were monitored by the appearance of preferential cleavage sites located in exposed and flexible regions of the proteins, which become more susceptible to proteolysis [29][30].

Complementary proteolysis experiments were performed on the three proteins, wild-type, L75P-, L174S-ApoAI by using three proteases: trypsin, chymotrypsin and endoprotease Glu-C, as conformational probes. The reactions were carried out in parallel using the same enzyme:substrate ratio and were monitored on a time-course basis by sampling the incubation mixture after 15 and 30 minutes of hydrolysis. Following this time course approach, primary and secondary cleavage sites, always occurring as a single proteolytic event, were only defined on kinetic basis.

Peptide mixtures were analyzed by liquid chromatography-mass spectrometry (LC-MS) onto a QuattroMicro LC- MS system (Micromass, Waters) interfaced with a 1100 HPLC (Agilent Technologies, Palo Alto, CA).

The LC-MS profiles showed that both variants were more susceptible to proteases than the wild-type protein, as demonstrated by the higher number of peptides released at the same hydrolysis time (Figures 4.7- 4.9). This observation suggests that both AApoAI variants are more flexible and/or less structured than the wild-type protein.

Molecular weight (MW) of intact proteins and complementary peptides originated by proteolysis with the relative MW are shown in Tables 4.3- 4.5.

Recombinant proteins have been expressed with a His tag stretch at N-terminus (Figure 4.6) and the molecular weights expected for each one are following reported: wild type ApoAI 29798.5 Da, L75P- AApoAI 29782.5, and L174S- AApoAI 29772.4.

HIS- WT ApoAI

```
  1 MHHHHHHGLV PRGSIDDPPQ SPWDRVKDLA TVYVDVLKDS GRDYVSQFEG SALGKQLNLK LLDNWDSVTS TFSKLREQLG
 81 PVTQEFWDNL EKETEGLRQE MSKDLEEVKA KVQPYLDDFQ KKWQEEMELY RQKVEPLRAE LQEGARQKLH ELQEKLSPLG
161 EEMRDRARAH VDALRTHLAP YSDELRQRLA ARLEALKENG GARLAEYHAK ATEHLSTLSE KAKPALEDLR QGLLPVLESF
241 KVSFLSALEE YTKKLNTQ
```

HIS- L75P ApoAI

```
  1 MHHHHHHGLV PRGSIDDPPQ SPWDRVKDLA TVYVDVLKDS GRDYVSQFEG SALGKQLNLK LLDNWDSVTS TFSKLREQLG
 81 PVTQEFWDNP EKETEGLRQE MSKDLEEVKA KVQPYLDDFQ KKWQEEMELY RQKVEPLRAE LQEGARQKLH ELQEKLSPLG
161 EEMRDRARAH VDALRTHLAP YSDELRQRLA ARLEALKENG GARLAEYHAK ATEHLSTLSE KAKPALEDLR QGLLPVLESF
241 KVSFLSALEE YTKKLNTQ
```

HIS- L174S ApoAI

```
  1 MHHHHHHGLV PRGSIDDPPQ SPWDRVKDLA TVYVDVLKDS GRDYVSQFEG SALGKQLNLK LLDNWDSVTS TFSKLREQLG
 81 PVTQEFWDNL EKETEGLRQE MSKDLEEVKA KVQPYLDDFQ KKWQEEMELY RQKVEPLRAE LQEGARQKLH ELQEKLSPLG
161 EEMRDRARAH VDALRTHLAP YSDELRQRSA ARLEALKENG GARLAEYHAK ATEHLSTLSE KAKPALEDLR QGLLPVLESF
241 KVSFLSALEE YTKKLNTQ
```

Figure 4.6: Sequences of His-tagged wild-type, L75P-, L174S- AApoAI. His-tag stretch at N-terminus is in black. The amino acid substitutions in L75P and L174S variants are reported in bold.

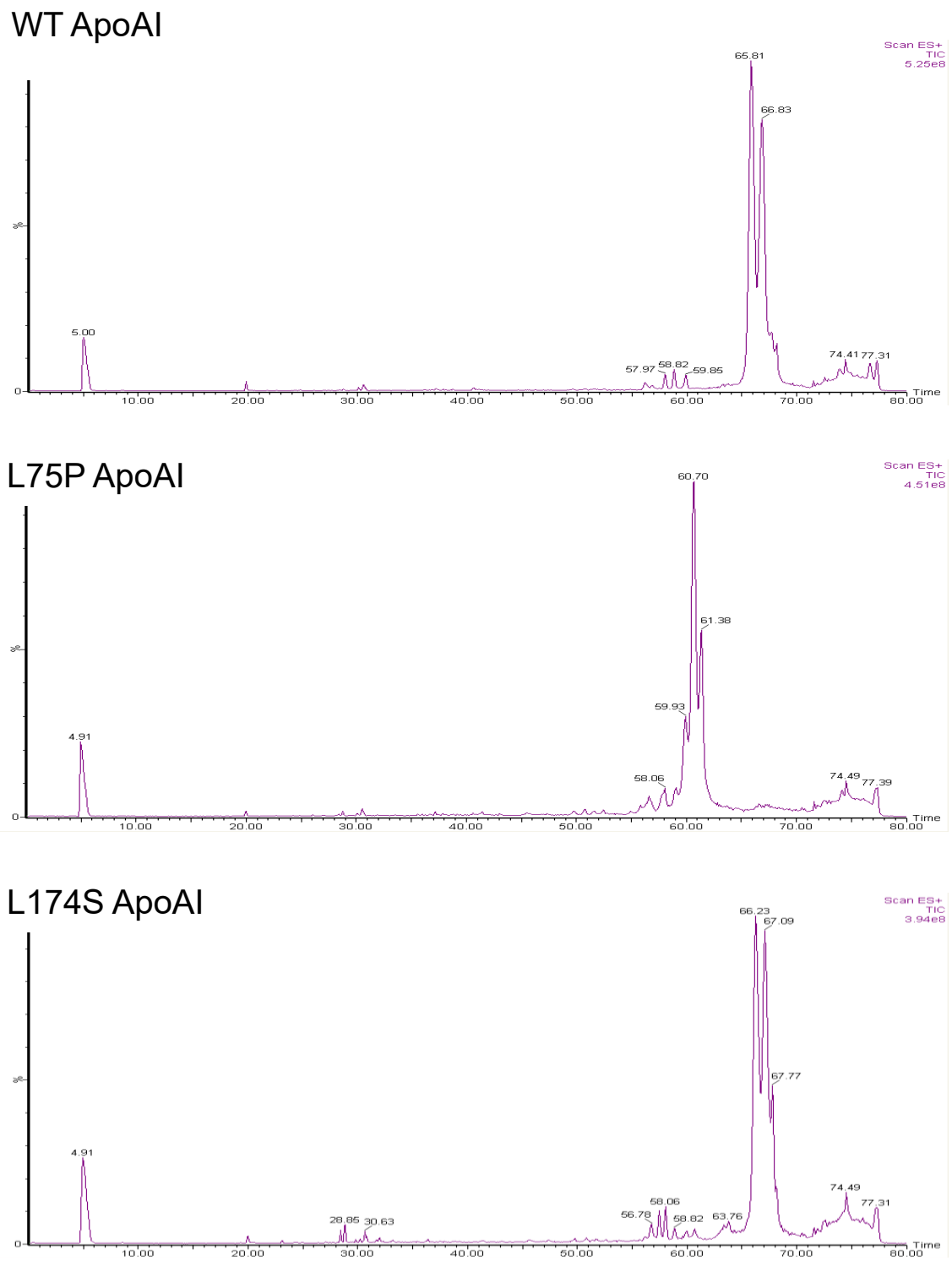


Figure 4.7: Chromatographic profiles (30 min) of proteolysis experiments of wild type, L75P, L174S ApoAI in presence of trypsin 1/8000.

Table 4.3: Wild type, L75P, L174S ApoAI proteolytic cleavages carried out by using trypsin. Peptides released, retention time (RT), observed and theoretic MW (Da), cleveage types are reported.

ApoAI	Peptides released	RT (min)	Experimental MW (Da)	Theoretic MW_{AV} (Da)	Cleavages
Wild-type	1-4 5-243	20.00 66.83	1494.0±0.01 28322±1.98	1494.70 28321.83	Primary
	1-131 132-243	67.68 56.18	17102.25±3.82 12715.93±0.86	17101.08 12715.45	Primary
	1-133 134-243	67.68 56.18	17357.59±2.22 12459.71 ±1.61	17357.38 12459.15	Secondary
	1-188 189-243	65.81 57.97	23630.38+1.78 6188.03± 1.5	23628.46 6188.07	Secondary
	1-243	65.81	29799.0 ±1.79	29798.5	-
L75P	1-4 5-243	20.00 61.38	1494.1±0.02 28305.63±3.22	1494.70 28305.79	Primary
	1-96 97-243	59.08 55.84	12749.80±3.08 17050.25±2.33	12751.15 17049.33	Secondary
	1-188 189-243	59.93 58.06	23614.30±2.79 6188.21±0.71	23612.42 6188.71	Primary
	1-243	60.70	29783±1.5	29782.5	-
L174S	1-4 5-243	19.90 67.09	1494.84±0.02 28295.98±2.37	1494.70 28295.75	Primary
	1-188 189-243	66.23 58.06	23603.91±2.58 6188.06±0.28	23602.38 6188.07	Secondary
	1-153 154-243	67.09 56.78	19747.81±1.35 10044.52±0.41	19747.10 10043.35	Primary
	1-243	66.23	29772.29±1.98	29772.4	-

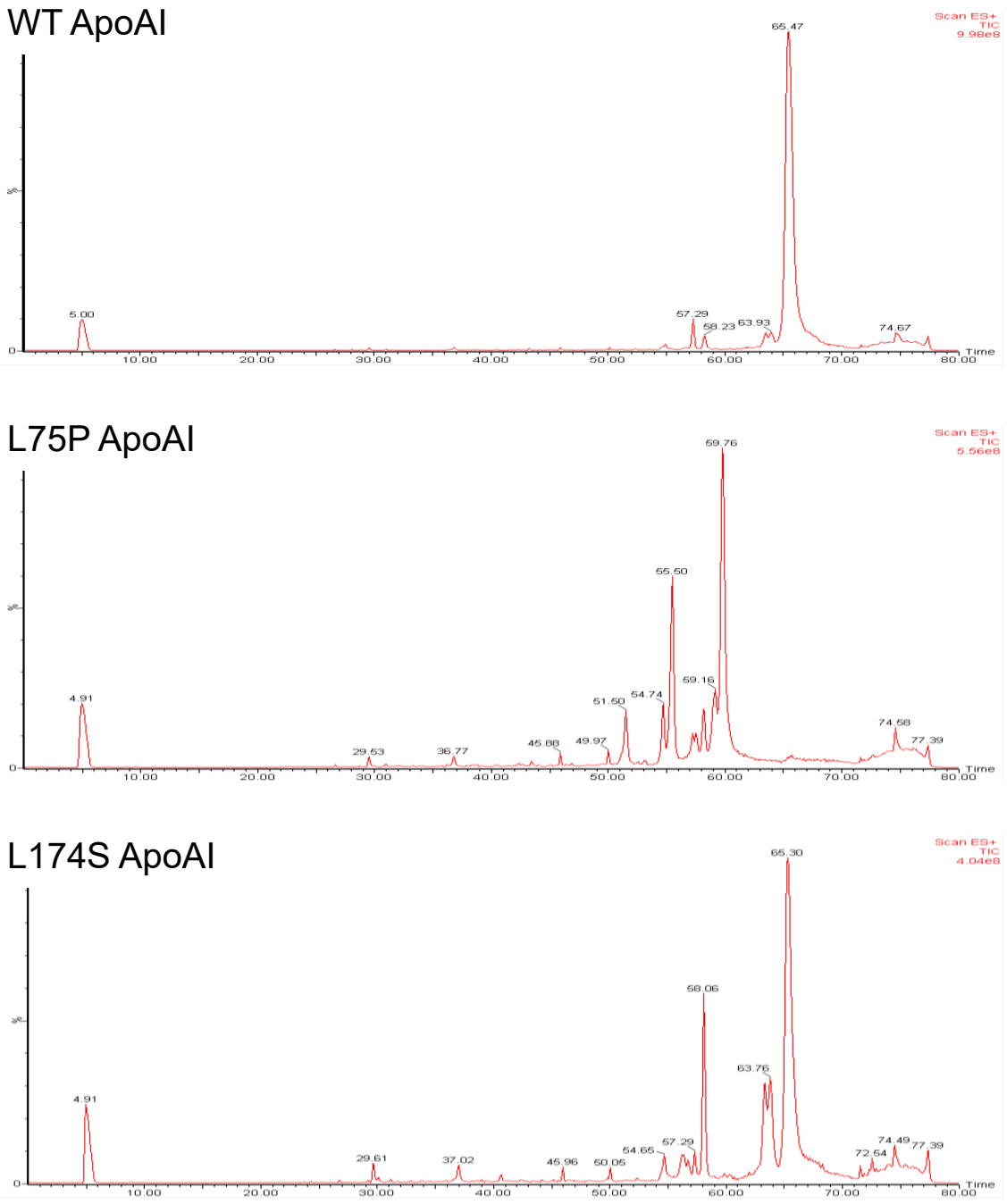


Figure 4.8: Chromatographic profiles (30 min) of proteolysis experiments of wild type, L75P, L174S ApoAI in presence of chymotrypsin 1/8000.

Table 4.4: Wild type, L75P, L174S ApoAI proteolytic cleavages carried out by using chymotrypsin. Peptides released, retention time (RT), observed and theoretic MW (Da), cleveage type are reported.

ApoAI	Peptides released	RT (min)	Experimental MW (Da)	Theoretic MW_{AV} (Da)	Cleavages
Wild-type	1-8 9-243	29.44 63.93	2675.13±0.29 27142.75±2.37	2674.94 27141.59	Primary
	1-18 19-243	36.85 63.51	3836.14±0.16 25981.38±3.28	3836.26 25980.27	Secondary
	1-57 58-243	54.91 58.23	8165.70±1.65 21652.96±2.63	8165.05 21651.48	Secondary
	1-192 193-243	65.47 57.29	24106.92±1.26 5712.45±1.41	24104.99 5711.54	Primary
	1-243	65.47	29800±2.25	29798.5	-
L75P	1-8 9-243	29.53 58.31	2675.03±0.16 27127.99±1.85	2674.94 27125.55	Primary
	1-18 19-243	36.77 57.72	3836.19±0.29 25966.14±2.41	3836.25 25964.23	Primary
	1-57 58-243	54.74 55.50	8165.92±1.06 21637.09±1.66	8165.05 24635.44	Primary
	1-192 193-243	59.16 57.29	24091.81±1.96 5712.42±0.62	24088.94 5711.54	Primary
	1-243	59.76	29783±1.3	29782.5	-
L174S	1-8 9-243	29.61 63.76	2675.09±0.35 27115.44±1.94	2674.94 27115.51	Primary
	1-18 19-243	37.02 63.76	3835.80±0.81 25955.47±2.31	3836.26 25954.19	Primary
	1-57 58-243	54.65 58.06	8164.89±1.20 21627.11±2.13	8165.05 21625.40	Primary
	1-115 116-243	67.08 55.24	15241.96±1.68 14552.21±2.13	15238.96 14551.49	Secondary
	1-192 193-243	65.30 57.29	24079.81±1.87 5711.44±0.71	24078.91 5711.54	Secondary
	1-243	65.30	29773±1.4	29772.4	-

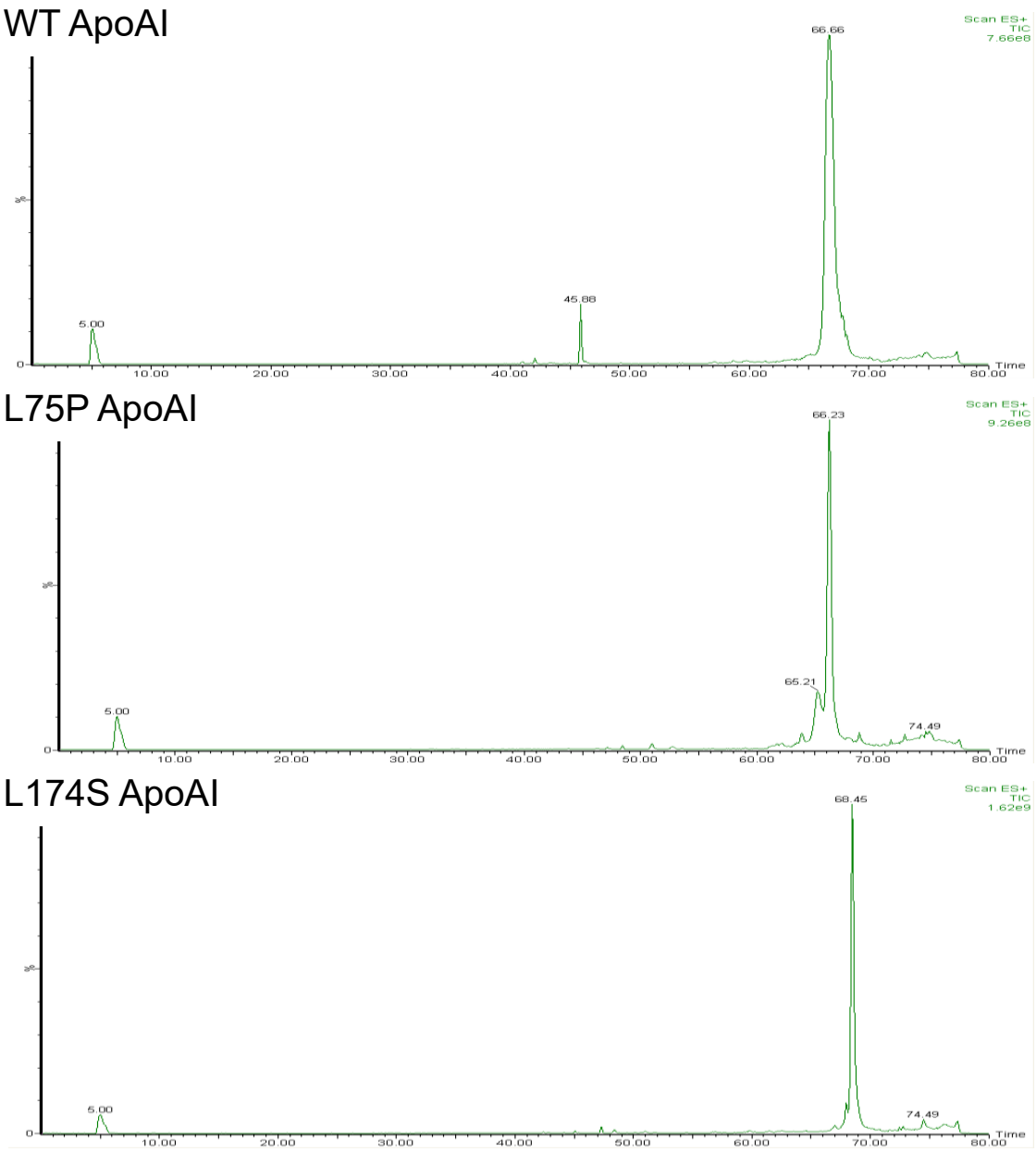


Figure 4.9: Chromatographic profiles (30 min) of proteolysis experiments of wild type, L75P, L174S ApoAI in presence of endoproteinase GluC 1/10000.

Table 4.5: Wild type, L75P, L174S ApoAI proteolytic cleavages carried out by using endoprotease Glu-C. Peptides released, retention time (RT), observed and theoretic MW (Da), cleveage type are reported.

ApoAI	Peptides released	RT (min)	Experimental MW (Da)	Theoretic MW_{AV} (Da)	Cleavages
Wild-type	1-136	67.68	17735.20±2.69	17736.80	Secondary
	137-243	57.12	12079.90±2.75	12079.73	
	1-223	66.66	27484.23±1.84	27482.87	Primary
	224-243	45.88	2333.39±0.10	2333.66	
	1-243	66.66	29799.49±1.8	29798.5	-
L75P	1-34	52.69	5674.68±0.73	5675.25	Primary
	35-243	63.85	24126.94±1.94	24125.24	
	1-223	65.21	27466.41±2.16	27466.83	Secondary
	224-243	48.43	2333.28±0.08	2333.66	
	1-243	66.23	29783.88±2.77	29782.5	-
L174S	1-34	51.16	5674.57±0.93	5675.25	Primary
	35-243	67.85	24113.88±1.72	24115.20	
	1-136	67.85	17735.55±1.62	17736.80	Secondary
	137-243	60.44	12053.22±0.91	12053.65	
	1-243	68.45	29772.29±1.98	29772.4	-

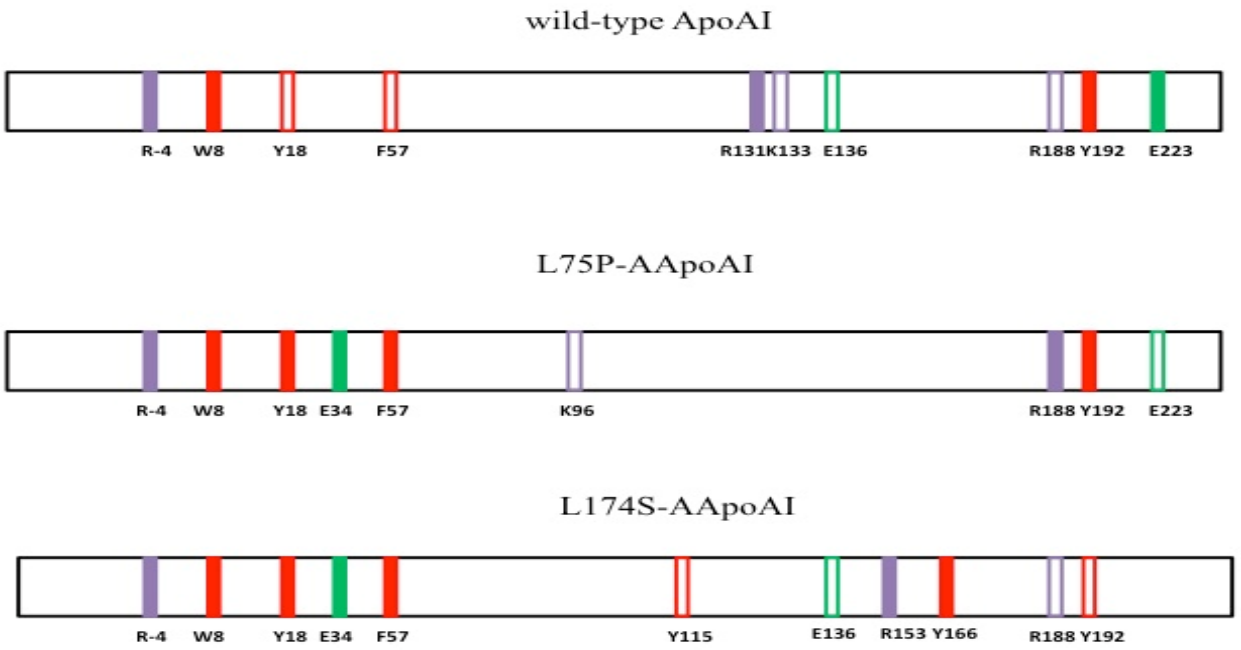


Figure 4.10: Schematic representation of the results obtained from limited proteolysis experiments on wild-type ApoAI (upper), L75P-AApoAI (middle) and L174S-AApoAI (lower). Proteolytic sites are designed as primary (filled) or secondary sites (empty) on kinetic basis. The R-4 cleavage site was not taken in consideration, as it is located in the his-tag peptide sequence occurring in the recombinant proteins.

The distribution of the proteolytic sites of ApoAI, obtained by trypsin, chymotrypsin and endoprotease Glu-C hydrolyses, is shown in Figure 4.10.

In all the proteins a similar pattern of proteolytic accessibility can be observed, with preferential cleavage sites gathering at both the N- and C- terminal regions.

However, proteolytic cleavages at *Y18* and *F57* residues occurred much faster in the mutants than in wild-type ApoAI and *E34* was not recognized in the native protein, suggesting a higher flexibility in the N-terminal region of the variants. Moreover, hydrolysis at residues *K96* and *Y115*, in L75P- and L174S-AApoAI, respectively, suggests a slight conformational change in the region including the putative proteolytic site in AApoAI variants, responsible for the release of the fibrillogenic polypeptide. On the contrary, in wild-type ApoAI this region, as reported in the crystal structure [31], is in predominantly α-helical structure, which protects the protein from proteolytic cleavage.

Also at C-terminal region, wild-type ApoAI and L75P variant are characterized by a similar proteolytic profile, showing accessibility at residues *R188*, *Y192* and *E223*. The disappearance of the *E223* cleavage site and the slower kinetic of hydrolysis at level of *R188* and *Y192* residues in the L174S variant suggest that a slight conformational change occurred in this region.

Further differences in the proteolytic patterns of the three proteins were located in the region including the 131–136 sequence that was exposed in the wild-type ApoAI, as demonstrated by

cleavages at *R131*, *K133* and *E136* sites. On the other hand, this region became completely inaccessible in the L75P variant. The conformation in this region changed also in the L174S mutant, where *E136* was still recognized, but with slow kinetics and the appearance of new cleavage sites at *R153* and *Y166* was detected. It is noteworthy that these cleavages occurred quite close to the mutation site *174*, indicating that a conformational change occurred, making this region less structurated or higher accessible.

Complementary proteolysis support perfectly the molecular dynamic (MD) simulation carried out by our co-workers.

From the MD simulations it emerged that no significant changes in the overall structure from the corresponding crystal structure occurred in wild- type ApoAI, while the two variants show subtle, but significant variations. The replacement of L174 with a Ser has only a moderate effect on the overall structure of the protein. In fact, in the L174S variant, the serine side chain can be well accommodated within the helix 141–180, since it can form a hydrogen bond with the carbonylic oxygen of Y170 (Figure 4.11A). On the contrary, the L75P substitution has a stronger destabilizing effect, since the presence of Pro breaks helix 70–76 and disturbs the stabilizing hydrophobic interactions that play a major role in orienting the helices of the four- segment bundle (Figure 4.11B).

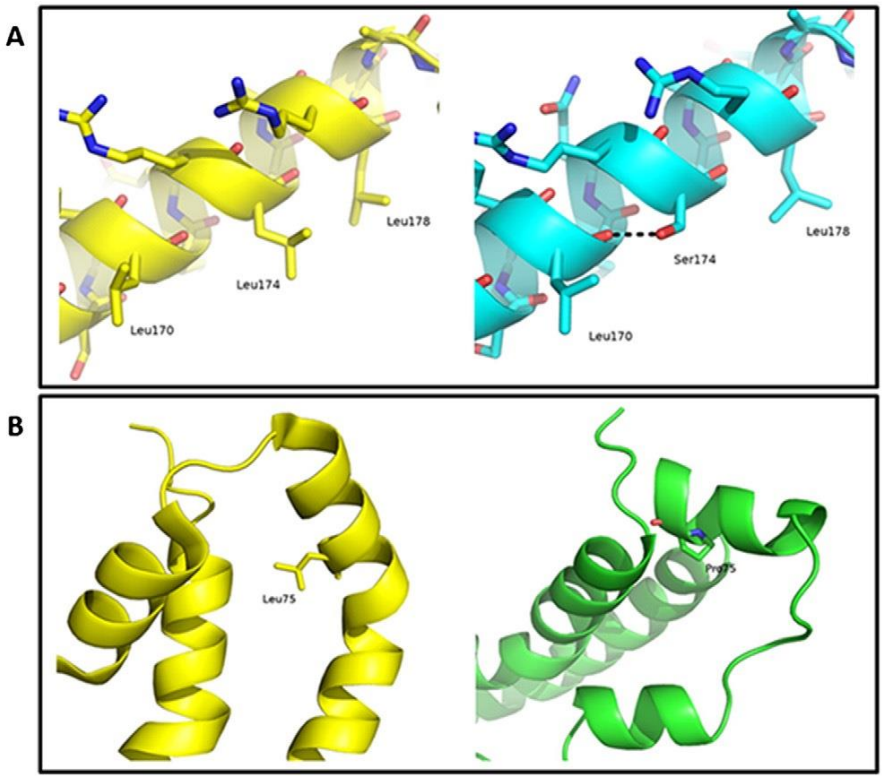


Figure 4.11: Structural features of AApoAIvariants. A, insight of the structure of helix encompassing residues 141–180 in wild-typeApoAI (yellow) and in L174S variant (cyan). B, ribbon representation of the helix bundle in proximity of L75 in the structure of wild-typeApoAI (yellow) and of L75P variant (green). For sake of clarity, one helix from the second chain of the dimer was removed.

All these data clearly indicate that the mutations enhance the overall flexibility of the protein and both L75P- and L174S-AApoAI variants have an increased exposure of the N-terminal region. In particular, L75P-AApoAI shows a lower helical content in correspondence of residues 66–73, 135–137, 157–160 and 172–177, but it is more structured than both wild-type ApoAI and L174S in the regions 125–129 and 140–142. L174S-AApoAI has a lower α-helical content in correspondence of residues 107–110, 140–141, but it is more structured in the regions 77–80 and 158–160.

These simulation data well fit with the results of far-UV CD spectroscopy measurements showing that the wild- type ApoAI and the two amyloidogenic variants were predominantly in an α-helical state, with a slightly lower α-helix content of the variants with respect to the wild-type protein (Table 4.6). When the proteins were incubated at 37 °C over time, no significant changes in the spectra of the wild-type protein were observed; on the contrary, the α-helicity of both variants decreased over time accompanied by an increase in β- strand, indicating an alteration of the protein native conformation towards an aggregation competent state.

Table 4.6: α-Helix and β-strand contents (%) for wild-type ApoAI and its amyloidogenic variants, upon incubation at 37 °C. Data shown are the means from three independent experiments. The accuracy of the secondary structure content estimation is approximately 2% for wild-type ApoAI and 4% for its amyloidogenic variants.

Incubation time (h)	Wild-type ApoAI		L75P-AApoAI		L174S-AApoAI	
	α-Helix (%)	β-Strand (%)	α-Helix (%)	β-Strand (%)	α-Helix (%)	β-Strand (%)
0	49	29	40	36	45	28
24	49	28	34	36	40	21
48	47	29	32	39	35	37
72	40	34	31	39	29	41
168	38	33	29	40	20	53

As supposed, the substitution of L75 with proline has a destabilizing effect, as it breaks helix 70–76. This helix overlaps with one of the three predicted N- terminal amyloid “hot spots”, whose perturbation has been recently suggested to be a prerequisite for misfolding [32]. This leads to the exposure of hydrophobic residues mainly located in the N-terminal region, as demonstrated by the increased intrinsic fluorescence intensity and ANS binding [33], as well as by the exposure of a proteolytic site at Y18, belonging to the major hot spot 14–22, as also found for G26R and L178H variants [34]. Furthermore, the accessibility of further proteolytic sites at E34 and F57 (the latter belonging to the minor hot spot 53–58), both hindered in the native protein, is the consequence of the propagation of the local perturbation (helix 70–76) to the other two predicted hot spots and the cause of the decreased protein compactness. In fact, DLS measurements of the hydrodynamic

diameter carried out by our co-workers [33] demonstrated that the L75P variant is much less compact protein with respect to both the wild-type protein and the L174S variant.

4.2.3 Conclusions

In conclusion, the investigation of effects L75P and L174S mutations on conformational stability of ApoAI contribuited to add knowledge, at a molecular level, to the general mechanism by which amyloidogenic mutations determine ApoAI destabilization and a higher propensity to aggragate triggering systemic amyloidosis.

As a response to the occurrence of the “internal” mutation L75P, the protein acquires a “looser” structure at the N-terminal domain, with significant alteration of protein conformation and compactness, while the “external” mutation L174S is responsible for a less destabilized but more flexible structure to which a more pronounced aggregation-competent state is associated [33].

4.3 Development of an Innovative Proteomic Strategy for Typing Amyloid Deposits

4.3.1 Materials and Methods

4.3.1.1 Materials

Dithiothreitol (DTT), sodium deoxycholate (SDC), acetonitrile LC-MS, methanol LC-MS, formic acid, trifluoroacetic acid (TFA)were provided by SIGMA ALDRICH. Ammonia solution HPLC grade (35%) and water LC-MS were Fisher scientific products. Zwittergent (3-16 detergent) was provided by Merck Millipore.

Trypsin Gold Mass Spectrometry Grade was Promega product.

Sodium chloride, calcium chloride, ethanol (96%), sodium hydroxide were provided by AnalaR NORMAPUR.

ClearVueMountants XYL was Thermo Scientific product. Formal saline solution (10%) was provided by CellPath.

Congo red stain was provided by AMRESCO.

Mayer's Hematoxylin solution and Scott's Tap Water Substitute solution were Pioneer Research Chemicals products.

Adipose tissue biopsies were provided by National Amyloidosis Centre (NAC), University College London, Royal Free Hospital, London, UK.

4.3.1.2 Methods

Adipose tissues decellularization - The protocol for the decellularization of adipose tissue was carried out through sequential 2 min washing steps with 500mM NaCl; 4% sodium deoxycholate (SDC) in the presence of 2mM calcium chloride; Tris-HCl containing 140mM NaCl (TC), 2mM calcium chloride, pH 8, respectively.

All the steps were performedby using a minimal volume of each buffer (500 µL), in a TissueLyser instrument set at a frequency of 25Hz. After decellularization the fat became translucent. Both non-treated and decellularized fat tissues were directly subjected to trypsin digestion.

Adipose tissues digestion - Proteins were extracted from adipose tissues by using buffer solution (50µL/sample) and by heating at 99 ºC for 90 minutes.

Buffer composition:

- ✓ 10mM TrisHCl pH 8
- ✓ 1mM $CaCl_2$ $2H_2O$
- ✓ 0.002% Zwittergent

The samples were allowed to cool, centrifuged at 10,000 rpm for 5 minutes, and sonicated for 60 minutes at room temperature. Each sample was treated with 20µL of 4ng/µL trypsin solution and incubated overnight at 37 °C. The digested samples were reduced with dithiothreitol (50µg/sample) at 99 °C for 5 minutes. Finally, the samples were freeze dried, reconstituted in 0.1 % v/v trifluoroacetic acid in HPLC grade water (100µL/sample) and 5µLof sample were analysed by LC-MS/MS.

LC-MS/MS analysis - Nanoflow liquid chromatography-electrospray tandem mass spectrometry was performed using a Waters nanoACQUITY™ UPLC system (Waters Ltd., Elstree, Hertfordshire, UK) coupled to a Thermo Scientific *OrbitrapVelos Mass Spectrometer* (Thermo Electron, Bremen, Germany) operated in the positive ion mode. Tryptic digests were applied to a trap column (180 µm ID x 20 mm bed, 5 µm Symmetry C18 packing; Waters Corporation, Massachusetts, USA) and separated on a reverse phase column (100 µm ID x 150 mm bed, 5 µm C18 packing; NikkyoTechnos Company Ltd, Tokyo, Japan) using a gradient from5% to 60 % acetonitrile/water over 44 min. Peptide analysis was performed using data-dependent acquisition (DDA) of one MS scan (mass range from 400 to 2000 m/z) followed by MS/MS scans of the ten most abundant ions in each MS scan.

The chromatographic system was washed between each sample using a protocol that includes five washes: 10% formic acid; 0.1% TFA; 1% NH_3 in 50% MeCN; 80% MeOH and 80% MeCN. A control blank run prior to each sample to confirm the absence of any carryover.

Proteins identification - Instrument control and data acquisition used Thermo Scientific Xcalibur Version 2.1. MS data files were analysed using Matrix Science Mascot. Searches were performed by setting following parameters:

- ✓ Database: SwissProt
- ✓ Taxonomy: Homo sapiens
- ✓ Enzyme: Trypsin (with 2 missed cleavages)
- ✓ Variable modification: oxidation (M)
- ✓ Peptide tolerance: 10 ppm
- ✓ MS/MS tolerance: ± 0,6 Da
- ✓ Ions charge: +2, +3, +4
- ✓ Instrument: ESI-TRAP

Protein identities were expressed in terms of MASCOT score probabilities with a significance value set at $P<0.05$.

In addition, Mascot output data files were analyzed and validated using Scaffold 4.6.1 (Proteome Software). This tool uses a Local False Discovery Rate (LFDR)-based scoring system for peptide validation based on a Bayesian approach to confirm peptide probabilities. MS data files were queried using both Mascot and X!Tandem algorithms.The likelihood of peptides is calculated on parent ion mass accuracy and parent ion delta masses. Proteins identified with 2 assigned peptides, protein threshold over the 99 % confidence level, peptide threshold over the 95 % were considered significant.

Congo red staining – Untreated and decellularized adipose tissues (5-10 μm thick sections) were attached to the slides and left to dry for 2 hours. After, the slides were placed for 1 minute in box containing formalin fixative alcoholic solution (85.5% ethanol and 10%of formal saline solution 10%) in order to preserve permanently the tissues. Then dried slides were put in a box with Mayer's hematoxylin stain for 10 minutes. This step was carried out in order to stain the nuclei of cells. Subsequently a wash was carried out in Scott's tap water solution for 5 minutes. The slides were washed with saturated NaCl in 80% ethanol and then with the same solution containing Congo red stain for 20 minutes, in order to visualize the amyloid. Finally, tha slide were washed in 100% ethanol removing the excess of stain, placed in ClearVuecoverslipper machine (Thermo Scientific) and left to dry over-night. The slides were displayed using LEICA DMLB microscope.

4.3.2 Results and Discussion

Proper amyloid typing represents a crucial step in the diagnosis and treatment of systemic amyloidosis. There are some cases in which more than one potentially amyloidogenic protein may be detected and an accurate diagnosis becomes challenging.

The present study was focused on the development of a methodology based on decellularization of human fat biopsies coupled with mass spectrometry analysis (LC-MS/MS) in order to improve the specificity of amyloid typing. Adipose tissue biopsies from patients affected by amyloidosis were provided by National Amyloidosis Centre at Royal Free Hospital-University College London (London, UK).

The research activity was carried out at the laboratoty of Wolfson Drug Discovery Unit, Centre for Amyloidosis and Acute Phase Proteins, University College London (London, UK), where it has recently been developed a strategy for the decellularization of a whole amyloidotic mouse organ demonstrating that structure and functional properties of amyloid fibrils were maintained

within the extracellular space. The procedure included sequential perfusion of water, detergent and DNase [35].

In this study, a specific protocol addressed to the decellularization of human fat biopsies has been set up. Ten cases were analysed. The small size tissue samples permitted to employ few steps of sequential washing using a 1/40 ratio for weight of biopsy/buffer volume. The procedure was fast and the limited number of washing steps minimized the loss of tissue material and allowed the analysis of sample of very limited dimensions. After decellularization, the fat tissues became translucent as shown in Figure 4.12.

Microscopic analysis in bright field demonstrated that the procedure successfully removed the cells without affecting the abundant green birefringence in the Congo-red stained material observed in the tissue before and after decellularization under cross-polarized light (Figure 4.13).

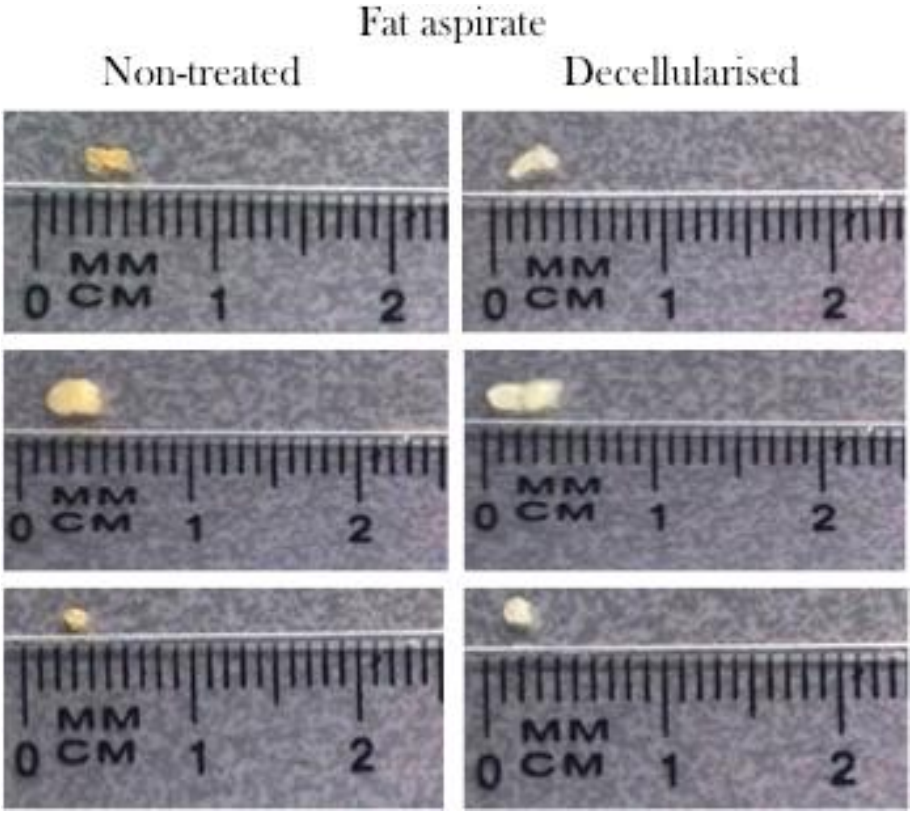
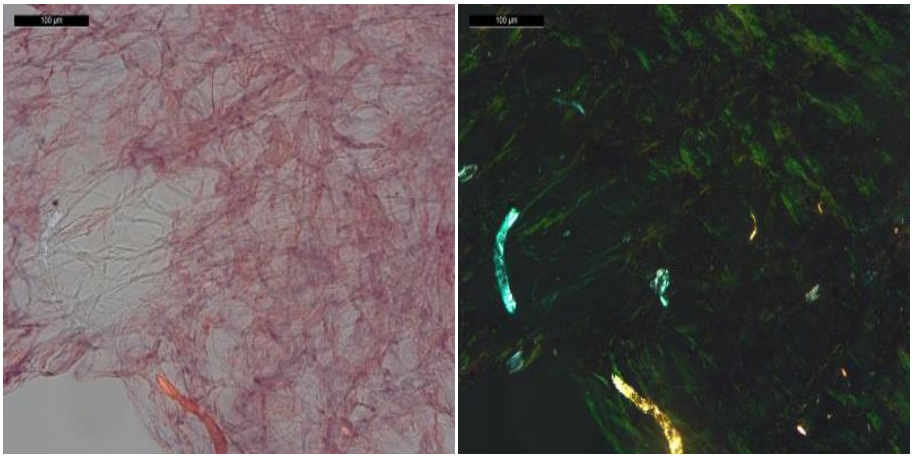


Figure 4.12: Fat aspirate biopsies pre- and post decellularization.

Non-treated



Decellularised

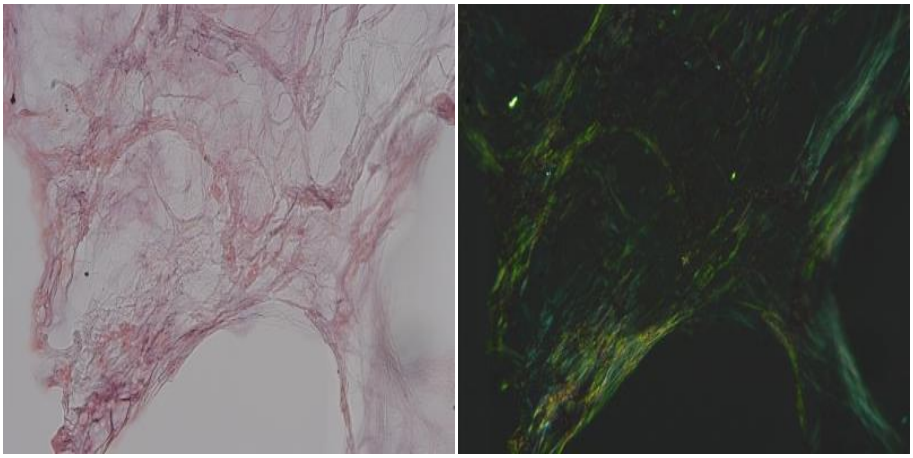


Figure 4.13: Microscopic analysis in bright field and under cross-polarized light of fat tissue before and after decellularization.

Both untreated and decellularized fat biopsies were digested by trypsin and analysed by LC-MS/MS; the proteins were identified by Mascot search. In addition, Mascot output data files were analyzed and validated by the use of Scaffold 4.6.1 tool, which provides information about:

- Protein identification probability (*P*)
- Exclusive Unique Peptide count (*EUP*): number of unique peptides associated only with each protein
- Total Spectrum count (*TS*): number of total spectra associated with each protein including those shared with other proteins

The identifications of potentially amyloidogenic proteins pre- and post-decellularization are shown in Table 4.7.

Table 4.7: Proteins identified for each sample with Mascot score probabilities and Scaffold software results: [1]Probability of protein identification, [2] Number of unique peptides and [3]Total spectrum count.

Fat aspirate	Non treated biopsy					Decellularised biopsy				
	Amyloid protein	MASCOT Score	SCAFFOLD RESULTS			Amyloid protein	MASCOT Score	SCAFFOLD RESULTS		
			[1]P	[2]EUP	[3]TS			[1]P	[2]EUP	[3]TS
Patient-1	**λ**	**809**	100%	10	28	**λ**	**2682**	100%	7	79
	κ	**560**	100%	8	15	**κ**	**427**	100%	6	11
Patient-2	**λ**	**883**	100%	12	30	**λ**	**1793**	100%	11	67
	TTR	**481**	100%	9	11	**TTR**	**359**	100%	8	9
Patient-3	**TTR**	**764**	100%	15	24	**TTR**	**377**	100%	12	16
	κ	**296**	100%	4	8	**κ**	**66**	100%	2	2
Patient-4	**λ**	**350**	100%	4	10	**λ**	**513**	100%	9	26
	κ	**60**	100%	2	3	**0**	**0**	0	0	0
Patient-5	**κ**	**505**	100%	5	15	**κ**	**614**	100%	6	17
	λ	**205**	100%	5	7	**λ**	**261**	100%	4	7
Patient-6	**λ**	**568**	100%	6	21	**λ**	**1669**	100%	11	75
	κ	**227**	100%	4	9	**κ**	**49**	0	0	0
Patient-7	**κ**	**3286**	100%	3	101	**κ**	**2965**	100%	2	94
	ApoA1	**576**	100%	13	18	**0**	**0**	0	0	0
Patient-8	**λ**	**287**	100%	5	14	**λ**	**997**	100%	5	29
	TTR	**132**	100%	2	3	**TTR**	**132**	0	0	0
Patient-9	**λ**	**553**	100%	4	26	**λ**	**914**	100%	7	34
	κ	**119**	100%	4	6	**0**	**0**	0	0	0
Patient-10	**TTR**	**578**	100%	7	13	**TTR**	**1799**	100%	15	63
	κ	**680**	100%	5	20	**κ**	**277**	100%	3	9

Starting from non-unique identification of amyloid protein, the analysis shows that after decellularization has been possible to obtain a substantial increase in Mascot score for amyloidogenic protein and a significant reduction, or in some cases total absence, of protein not part of amyloid fibrils, such as for patients 6, 7, 8 and 9.

Moreover, these results are particularly significant also for the cases 2, 3, 10, exhibiting a consistent presence of both IgG light chain and TTR protein and giving rise to ambiguous cases. In these situations, the identification of the real amyloidogenic protein is fundamental to understand amyloid type and then for the choice of the correct therapeutic approach. An example is the case number 10: in untreated sample, *k light chain* is equivalent, in term of Mascot score e total spectra, to those related to *TTR* protein. After decellularization process, Mascot score and number of peptides and spectra of k light chain decrease, while Mascot score of TTR triples and its total spectra become five times higher than before the procedure.

This result allows to clarify the amyloidogenic protein identity, in this case, TTR. This is extremely important since the application of a wrong therapy, in this case chemotherapy, used for light chain amyloidosis treatment, would harm the patient.

By reporting the natural logarithm of total spectra ratio of amyloidogenic (A) and non amyloidogenic protein (B), *ln(TsA/TsB)*, before and after decellularization (Figure 4.14), the general effect of decellularization procedure are more evident. Notably, the ratios are always higher after treatment than before, except for case number 5, suggesting that the decellularization procedure certainly improved the identification of amyloidogenic protein.

The parameters, including unique peptides, unique spectra and total spectra, provided by Scaffold, using Mascot and X!Tandem algorithms, were consistent with Mascot score.

In Figure 4.15 xy-plots show Mascot score versus total spectra of related proteins for each patient. Empty symbols refer to the results obtained though the standard procedure, while red symbols are the data obtained after decellularization. Triangle refers to the amyloidogenic protein and it is straight forward that almost in all cases the decellularization enhances the mascot score and total spectra of the culprit protein.

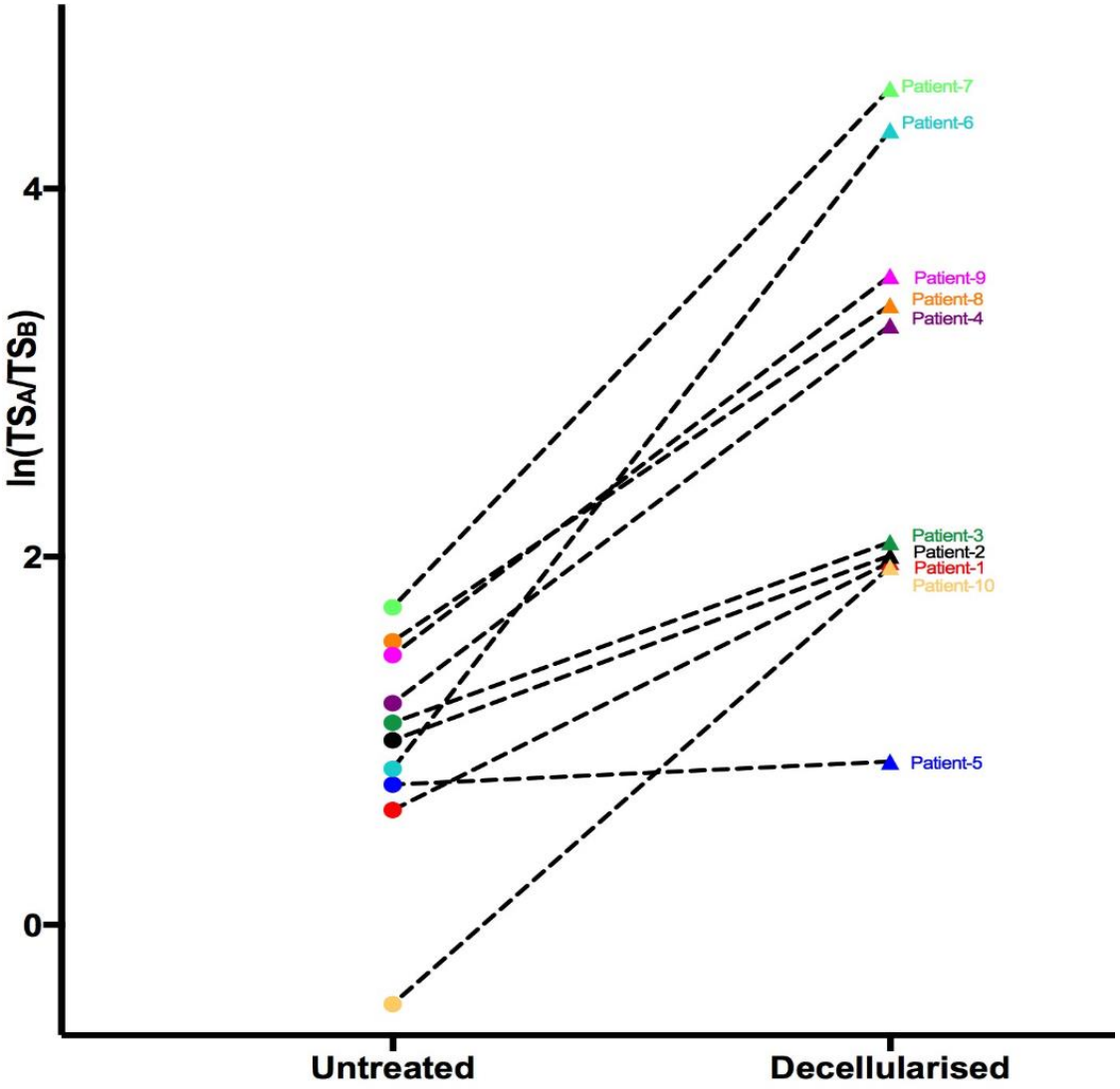
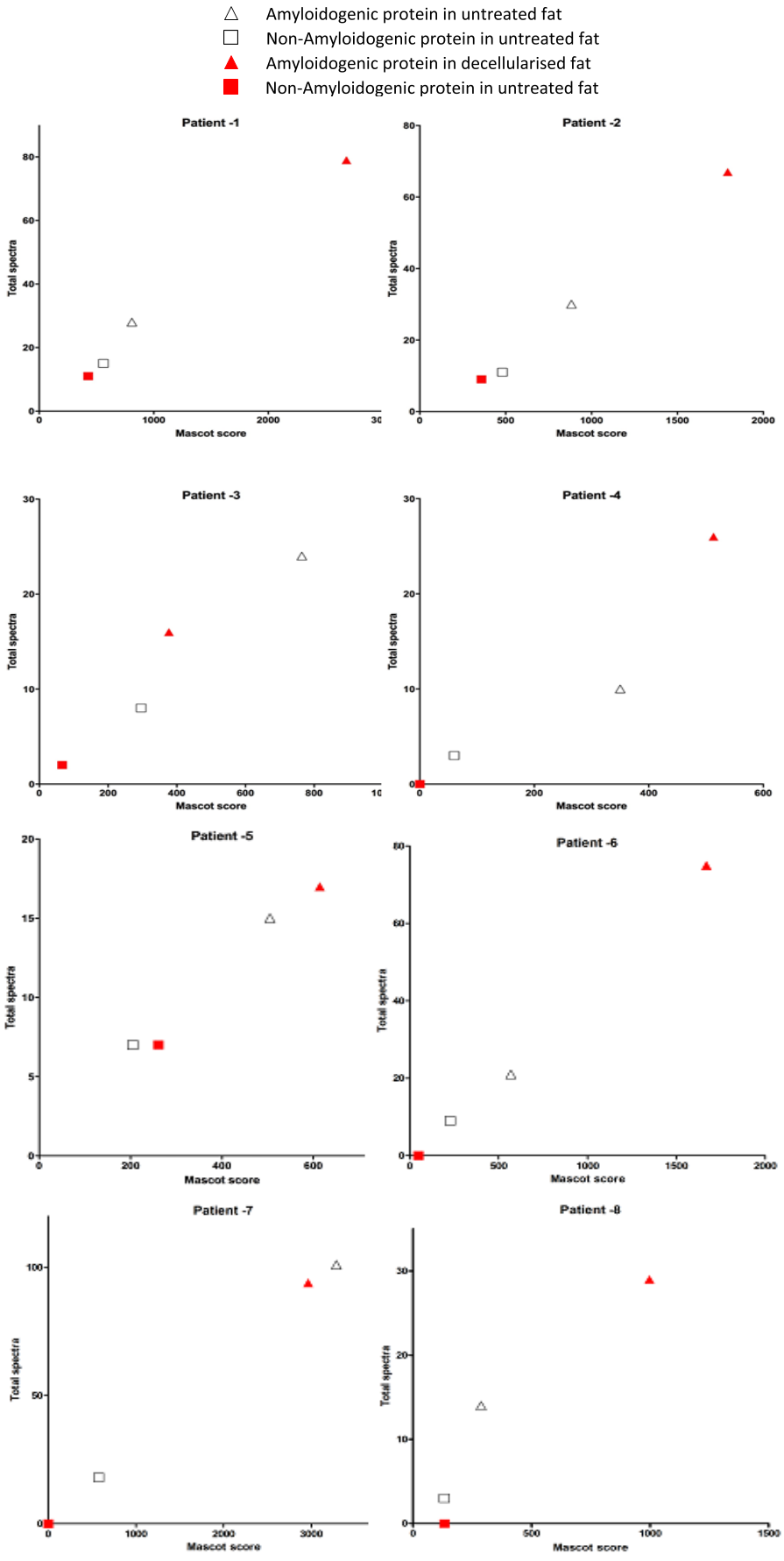


Figure 4.14: Total spectra ratio A/B, where A is amyloidogenic protein and B non amyloidogenic protein, ln(TsA/TsB), in untreated and decellularised samples.


Amyloidogenic protein in untreated fat
Non-Amyloidogenic protein in untreated fat
Amyloidogenic protein in decellularised fat
Non-Amyloidogenic protein in untreated fat
Patient -1
Patient -2
Patient -3
Patient -4
Patient -5
Patient -6
Patient -7
Patient -8
Total spectra
Mascot score

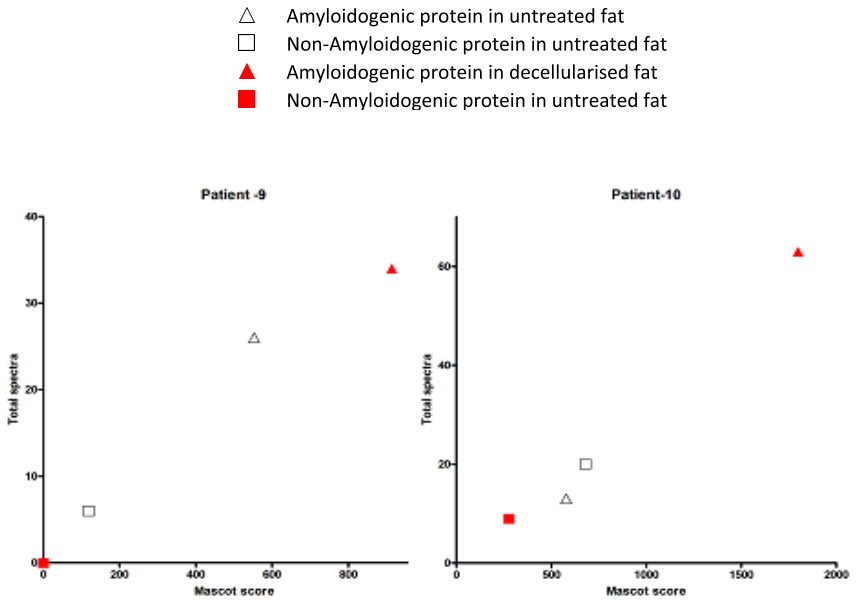


Figure 4.15: Mascot score versus Total spectra.

The amyloid proteins Serum Amyloid P component (SAP), Apolipoprotein E (ApoE) and Apolipoprotein A-IV (ApoA4), ubiquitous components of amyloid deposits, are considered a specific signature for these aggregates, independently from the protein causative of the disease. These specific signatures were monitored in all analysed samples before and after decellularization. Our results showed that at least two of three amyloid signature proteins were identified in both untreated and decellularised samples. This confirmed that the decellularization procedure does not affect natural structure and specific properties of amyloid fibrils, in agreement with a recent study [35].

In addition, the effectiveness of decellularization procedure has been demonstrated also considering total spectra ratio after and before decellularization of some proteins identified in human adipose tissue but not amyloidogenic (Figure 4.16).

Considering an average of these ratios for ten patients, it was observed that proteins belonging to cellular energetic metabolism, such as *ATP synthase subunit beta, mitochondrial* (*ATPB), Glycerol-3-phosphate dehydrogenase [NAD(+)], cytoplasmic (GPDA), Glyceraldehyde-3-phosphate dehydrogenase (G3P), Aldehyde dehydrogenase, mitochondrial (ALDH2), Fatty acid synthase (FAS), Perilipin-1(PLIN1), Perilipin-4 (PLIN4*) have a ratio TS_{AFTER}/TS_{BEFORE} lower than 1 indicating their reduction after decellularization. This is consistent with removing of cellular protein background after decellularization. Instead, some proteins, known related to amyloid and associated to extracellular matrix [36], such as *Clusterin* (*CLUS), Vitronectin (VTNC), Sushi repeat-containing protein SRPX(SRPX),* showed a ratio TS_{AFTER}/TS_{BEFORE} higher than 1.

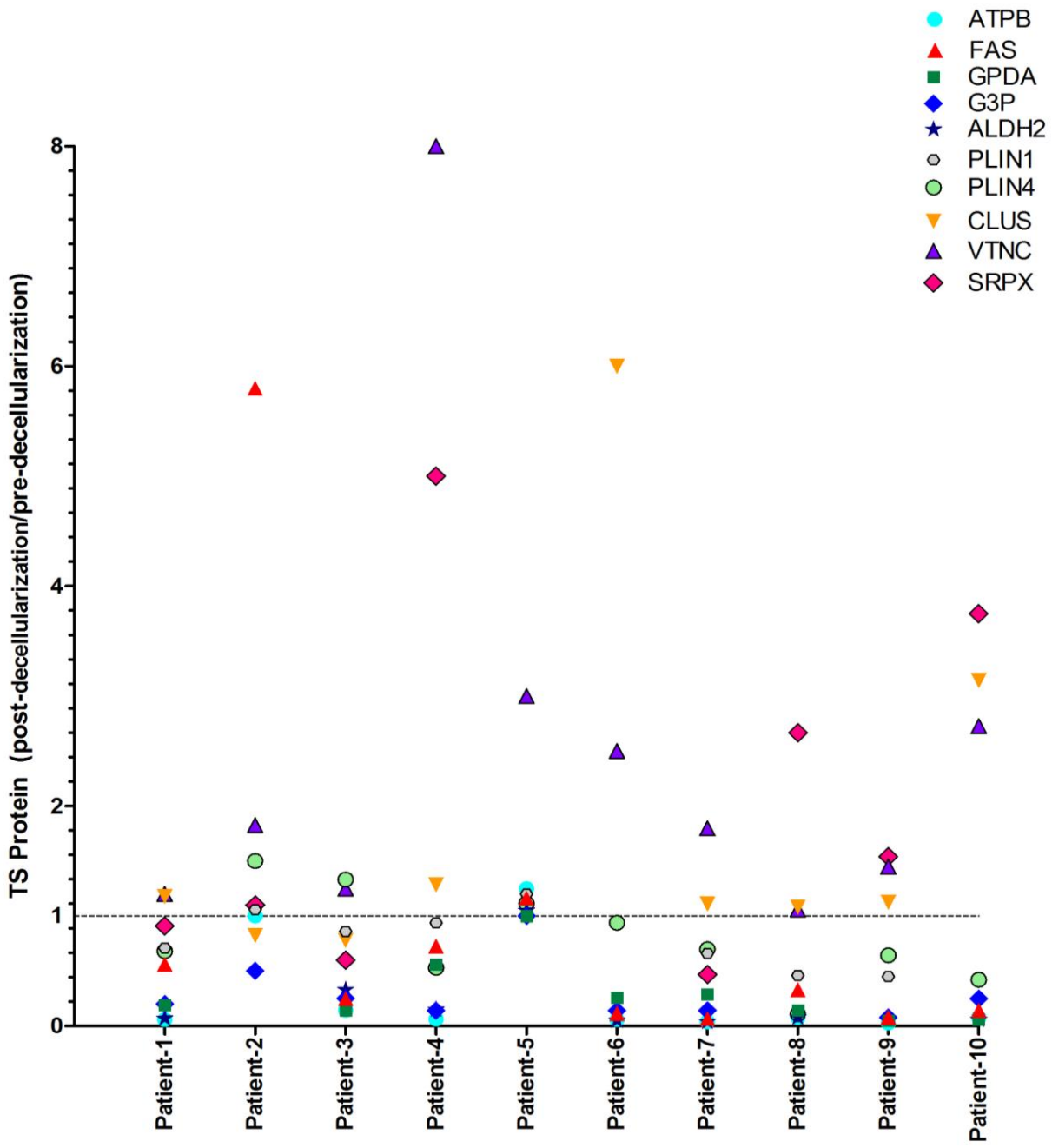


Figure 4.16: $TS_{POST\text{-}DECELLULARIZATION}/TS_{PRE\text{-}DECELLULARIZATION}$ of ATPB, GPDA, G3P, ALDH2, FAS, PLIN1, PLIN4, CLUS, VTNC, SRPX proteins identified in human adipose tissue biopsies.

This result suggests that decellularization procedure improves the identification of proteins belonging to ECM, therefore also amyloid protein, removing most of cellular protein background as demonstrated by the reduction of the amount of enzymes identified after the application of decellularization protocol.

4.3.3 Conclusions

Fine needle fat aspiration has the advantages to be a safe, simply and non-invasive technique to extract accumulated amyloid from tissue. Its disadvantage is that the proteome of the fine needle

aspiration biopsy typically contains more than one potential amyloid fibril protein precursor challenging the precise identification of the pathogenic protein.

The innovative strategy of decellularization of fresh fat aspirate specimens represents an intelligent and simple way to improve the accuracy and specificity of identification of amyloid fibril type removing most of cellular and plasma proteins background without altering properties of amyloid fibrils. Indeed, proteomic analysis revealed that in the clinical cases in which there was an ambiguous situation about amyloid protein identification, the decellularization procedure was able to improve the chemical classification of amyloid deposits.

This approach, allowing to clarify amyloidogenic protein identity, is extremely important to employ a right focused theraphy thus to increase the effectiveness and reduce side effects which would harm the patient.

4.4 References

[1] Merlini G, Bellotti V, Molecular mechanisms of amyloidosis. *The New England Journal of Medicine, Vol. 349: 583-96, 2003.*

[2] Bukau B, Weissman J, Horwich A, Molecular Chaperones and Protein Quality Control. *Cell, Vol. 125 (3): 443-51, 2006.*

[3] Stefani M, Dobson CM, Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *Journal of Molecular Medicine, Vol.81: 678–99, 2003.*

[4] Westermark P, Benson MD, Buxbaum JN, Cohen AS, Frangione B, Ikeda S, Masters CL, Merlini G, Saraiva MJ, Sipe JD, Amyloid: Toward terminology clarification Report from the Nomenclature Committee of the International Society of Amyloidosis. *Amyloid, Vol.12 (1): 1-4, 2005.*

[5] Chiti F, Dobson CM, Protein Misfolding, Functional Amyloid, and Human Disease. *Annual Reviews Biochemistry, Vol. 75: 333-66, 2006.*

[6] Pepys MB, Amyloidosis. *Annual Reviews Medcine, Vol. 57: 223–41, 2006.*

[7] Gertz MA, Immunoglobulin light chain amyloidosis: 2013 update on diagnosis, prognosis, and treatment. *American Journal of Hematology, Vol. 88 (5): 416-25, 2013.*

[8] Patel KS, Hawkins PN, Cardiac amyloidosis: where are we today?. *Journal of Internal Medicine, Vol. 278 (2): 126-44, 2015.*

[9] Dubrey S, Ackermann E, Gillmore J, The transthyretin amyloidoses: advances in therapy. *Postgradute Medical Journal, Vol. 91: 439–48, 2015.*

[10] Rye KA, Barter PJ, Formation and metabolism of prebeta-migrating, lipid-poor apolipoprotein A-I. *Arteriosclerosis, Thrombis, Vascular Biology,Vol. 24: 421–8, 2004.*

[11] Eriksson M, Schönland S, Yumlu S, Hegenbart U, von Hutten H, Gioeva Z, Lohse P, Büttner J, Schmidt H, Röcken C, Hereditary Apolipoprotein AI-Associated Amyloidosis in Surgical Pathology Specimens: identification of Three Novel Mutations in the APOA1 Gene. *Journal of Molecular Diagnostics, Vol. 11 (3): 257-62, 2009.*

[12] Obici L, Palladini G, Giorgetti S, Bellotti V, Gregorini G, Arbustini E, Verga L, Marciano S, Donadei S, Perfetti V, Calabresi L, Bergonzi C, Scolari F, Merlini G., Liver biopsy discloses a new apolipoprotein A-I hereditary amyloidosis in several unrelated Italian families. *Gastroenterology, Vol.126: 1416–22, 2004.*

[13] Mangione P, Sunde M, Giorgetti S, Stoppini M, Esposito G, Gianelli L, Obici L, Asti L, Andreola A, Viglino P, Merlini G, Bellotti V, Amyloid fibrils derived from the apolipoprotein A1 Leu174Ser variant contain elements of ordered helical structure. *Protein Science, Vol. 10: 187–99, 2001.*

[14] Rowczenio D, Dogan A, Theis JD, Vrana JA, Lachmann HJ, Wechalekar AD, Gilbertson JA, Hunt T, Gibbs SD, Sattianayagam PT, Pinney JH, Hawkins PN, Gillmore JD, Amyloidogenicity and Clinical Phenotype Associated with Five Novel Mutations in Apolipoprotein A-I. *The American Journal of Pathology, Vol. 179 (4): 1978-87, 2011.*

[15] Wechalekar AD, Gillmore JD, Hawkins PN, Systemic amyloidosis. *Lancet, Vol. 387 (10038): 2641-54, 2016.*

[16] Comenzo RL, How I treat amyloidosis. *Blood, Vol. 114 (15): 3147-57, 2009.*

[17] Mahmood S, Gilbertson JA, Rendell N, Whelan CJ, Lachmann HJ, Wechalekar AD, Hawkins PN, Gillmore JD, Two types of amyloid in a single heart. *Blood, Vol. 124 (19): 3025-7, 2014.*

[18] Maleszewski JJ, Murray DL, Dispenzieri A, Grogan M, Pereira NL, Jenkins SM, Judge DP, Caturegli P, Vrana JA, Theis JD, Dogan A, Halushka MK, Relationship between monoclonal gammopathy and cardiac amyloid type. *Cardiovascular Pathology, Vol. 22 (3): 189-94, 2013.*

[19] Gillmore JD, Maurer MS1, Falk RH, Merlini G, Damy T, Dispenzieri A, Wechalekar AD, Berk JL, Quarta CC, Grogan M, Lachmann HJ, Bokhari S, Castano A, Dorbala S, Johnson GB, Glaudemans AW, Rezk T, Fontana M, Palladini G, Milani P, Guidalotti PL, Flatman K, Lane T, Vonberg FW, Whelan CJ, Moon JC, Ruberg FL, Miller EJ, Hutt DF, Hazenberg BP, Rapezzi C, Hawkins PN, Nonbiopsy Diagnosis of Cardiac Transthyretin Amyloidosis. *Circulation, Vol. 133 (24): 2404-12, 2016.*

[20] Fielding CJ, Fielding PE, Molecular transport physiology of reverse cholesterol. *Journal of Lipid Research, Vol. 36 (2): 211-28, 1995.*

[21] Schmitz G, Grandl M, The Molecular Mechanisms of HDL and Associated Vesicular Trafficking Mechanisms to Mediate Cellular Lipid Homeostasis. *Arteriosclerosis, Thrombis, Vascular Biology, Vol. 29 (11): 1718-22, 2009.*

[22] Brouillette CG, Anantharamaiah GM, Engler JA, Borhani DW, Structural models of human apolipoprotein A-I: a critical analysis and review . *Biochimica et Biophysica Acta, Vol. 1531 (1-2): 4-46, 2001.*

[23] Bellotti V, Nuvolone M, Giorgetti S, Obici L, Palladini G, Russo P, Lavatelli F, Perfetti V, Merlini G, The workings of the amyloid diseases. *Annals of Medicine, Vol. 39 (3): 200-7, 2007.*

[24] Amarzguioui M, Mucchiano G, Häggqvist B, Westermark P, Kavlie A, Sletten K, Prydz H, Extensive Intimal Apolipoprotein A1-Derived Amyloid Deposits in a Patient with an Apolipoprotein A1 Mutation. *Biochemical and Biophysical Research Communications, Vol. 242 (3): 534-9, 1998.*

[25] Di Gaetano S, Guglielmi F, Arciello A, Mangione P, Monti M, Pagnozzi D, Raimondi S, Giorgetti S, Orrù S, Canale C, Pucci P, Dobson CM, Bellotti V, Piccoli R, Recombinant amyloidogenic domain of ApoA-I: Analysis of its fibrillogenic potential. *Biochemical and Biophysical Research Communications, Vol. 351 (1): 223-8, 2006.*

[26] Das M, Mei X, Jayaraman S, Atkinson D, Gursky O, Amyloidogenic mutations in human apolipoprotein A-I are not necessarily destabilizing - a common mechanism of apolipoprotein A-I misfolding in familial amyloidosis and atherosclerosis. *FEBS Journal; Vol. 281 (11): 2525-42, 2014.*

[27] Obici L, Franceschini G, Calabresi L, Giorgetti S, Stoppini M, Merlini G, Bellotti V, Structure, function and amyloidogenic propensity of apolipoprotein A-I. *Amyloid, Vol. 13 (4): 191-205, 2006.*

[28] Dal Piaz F, Casapullo A, Randazzo A, Riccio R, Pucci P, Marino G, Gomez-Paloma L Chembiochem *Molecular Basis of Phospholipase A2 Inhibition by Petrosaspongiolide M, Vol.3(7): 664-71, 2002*

[29] Esposito G, Michelutti R, Verdone G, Viglino P, Hernàndez H, Robinson CV, Amoresano A, Dal Piaz F, Monti M, Pucci P, Mangione P, Stoppini M, Merlini G, Ferri G, Bellotti V, Removal of the N-terminal hexapeptide from human β2-microglobulin facilitates protein aggregation and fibril formation, *Protein Science. 9 831–845, 2000*

[30] Monti M, Principe S, Giorgetti S, Mangione P, Merlini G, Clark A, Bellotti V, Amoresano A, Pucci P, *Topological investigation of amyloid fibrils obtained from β2-microglobulin, Protein Science. Vol.11: 2362–2369*, *2002*

[31] Mei X, Atkinson D, Crystal structure of C-terminal truncated apolipoproteinA-I reveals the assembly of HDL by dimerization, *Journal of Biological Chemistry, Vol 286: 38570–38582, 2011*

[32] Das M, Mei X, Jayaraman S, Atkinson D, Gursky O, Amyloidogenic mutations in human apolipoprotein A-I are not necessarily destabilizing- a common mechanism of apolipoprotein A-I misfolding in familial amyloidosis and atherosclerosis, *FEBS Journal Vol.281: 2525–2542, 2014*

[33] Del Giudice R, Arciello A, Itri F, Merlino A, Monti M, Buonanno M, Penco A, Canetti D, Petruk G, Monti SM, Relini A, Pucci P, Piccoli R, Monti DM, Protein conformational perturbations in hereditary amyloidosis: Differential impact of single point mutations in ApoAI amyloidogenic variants. *Biochimica et Biophysica Acta, Vol. 1860 (2): 434-44, 2016.*

[34] Petrlova J, Duong T, Cochran MC, Axelsson A, Mörgelin M, Roberts LM, Lagerstedt JO, The fibrillogenic L178H variant of apolipoprotein A-I forms helical fibrils. *Journal of Lipid Research, Vol 53: 390–8, 2012.*

[35] Mazza G, Simons JP, Al-Shawi R, Ellmerich S, Urbani L, Giorgetti S, Taylor GW, Gilbertson JA, Hall AR, Al-Akkad W, Dhar D, Hawkins PN, De Coppi P, Pinzani M, Bellotti V, Mangione PP, Amyloid persistence in decellularized liver: biochemical and histopathological characterization. *Amyloid, Vol. 23 (1): 1-7, 2016.*

[36] Brambilla F, Lavatelli F, Di Silvestre D, Valentini V, Palladini G, Merlini G, Mauri P, Shotgun Protein Profile of Human Adipose Tissue and Its Changes in Relation to Systemic Amyloidoses. *Journal of Proteome Research, Vol. 12 (12): 5642-55, 2013.*

PhD Course Activity Summary

Candidate: Dr. Diana Canetti **Supervisor: Prof. Maria Monti**

1) Attended Courses:

- Tecniche di estrazione solido-liquido, 8 hours, June 2014 (Prof. Naviglio)
- Spettrometria di Massa, 10 hours, July 2014 (Prof. Pucci)
- Microscopia Raman, 8 hours, December 2014 (Prof. Vergara)
- Sintesi, struttura ed applicazioni di oligonucleotide naturali e modificati, 8 hours, February 2015 (Prof. Montesarchio)
- Glicoscienza, 8 hours, July 2015 (Prof. Parrilli and Dr. Bedini)
- Produzione ricombinante di proteine naturali e mutanti, 8 hours, July 2015 (Prof. Duilio)
- Neutron scattering, 8 hours, October 2015 (Prof. Zorn)

2) Attended Seminars:

Title	Speaker	Date	Place
Studio delle basi molecolari dell'acidemia metilmalonica	Dr. Marianna Caterino	6/06/2014	CEINGE Biotecnologie Avanzate
Evidenze genetiche sulle prime espansioni umane dall'Africa	Prof. Guido Barbujani	23/10/2014	Università degli Studi di Napoli "Federico II"
Antibodies and Mass Spectrometry…collaborators or competitors?	Prof. Pietro Pucci	12/12/2014	CEINGE Biotecnologie Avanzate
Huntington: storia di un gene tra evoluzione e meccanismi di patologia	Prof. Chiara Zuccato	15/01/2015	CEINGE Biotecnologie Avanzate
Alessandro Ballio e la chimica a Napoli negli anni '60	Dr. Andrea Carpentieri	18/06/2015	Università degli Studi di Napoli Federico II
Le transglutaminasi..dalle poliammine alle bioplastiche	Prof. Raffaele Porta	30/06/2015	Università degli Studi di Napoli Federico II
Introduction to bioinformatics	Dr. Remo Sanges	18/12/2015	Università degli Studi di Napoli Federico II
Biopesticides which target voltage-gated ion channels: efficacy and biosafety	Prof. Angharad M.R.Gatehouse	14/01/2016	Università degli Studi di Napoli Federico II

Title	Speaker	Date	Place
1,1,2,2-tetraethoxybutyne-Exciting to make and work with	Prof. Leiv K. Sydnes	14/01/2016	Università degli Studi di Napoli Federico II
Functionalized and artificial enzymes:newbio-derived catalysts	Prof. Therrytron	14/01/2016	Università degli Studi di Napoli Federico II
Secondary metabolites from higher land and sea plants in Chemical ecology and Chemosystematics	Prof. Christian Zidorn	27/01/2016	Università degli Studi di Napoli Federico II
Cardiac ageing: a matter of monoamino oxidases?	Prof. Angelo Parini	18/10/2016	Royal Free Campus, University College London
Hepatology become non-invasive	Prof. Massimo Pinzani	19/10/2016	Royal Free Campus, University College London

3) Visiting periods in Institutions different from University of Naples Federico II:

Host Institution	Country	Start Date	End Date
San Raffaele Hospital of Milan	Milan (Italy)	24th November 2014	5th December 2014
University College London	London (United Kingdom)	1st April 2016	20th December 2016

4) Publications (include submitted and in preparation):

- Chesi G.; Hegde R.N.; Iacobacci S.; Concilli M.; Parashuraman S.; Festa B.P.; Polishchuk E.V.; Di Tullio G.; Carissimo A.; Montefusco S.; **Canetti D.**; Monti M.; Amoresano A.; Pucci P.; van de Sluis B.; Lutsenko S.; Luini A.; Polishchuk R.S. Identification of p38 MAPK and JNK as new targets for correction of Wilson disease-causing ATP7B mutants. *Hepatology* **2016**, *63(6)*, 1842-1859.
- Del Giudice R.; Arciello A.; Itri F.; Merlino A.; Monti M.; Buonanno M.; Penco A.; **Canetti D.**; Petruk G.; Monti S.M.; Relini A.; Pucci P.; Piccoli R.; Monti D.M. Protein conformational perturbations in hereditary amyloidosis: Differential impact of single point mutations in ApoAI amyloidogenic variants. *Biochimica et Biophysica Acta* **2016**,*1860(2)*, 434-444.
- Butturini E.; Gotte G.; Dell'Orco D.; Chiavegato G.; Marino V.; **Canetti D.**; Cozzolino F.; Monti M.; Pucci P.; Mariotto S. Intermolecular disulfide bond influences

unphosphorylated STAT3 dimerization and function. *Biochemical Journal* **2016**, *473(19)*, 3205-3219.

- Mangione P.P.; Mazza G.; Gilbertson J.A.; Rendell N.; **Canetti D.**; Giorgetti S.; Frenguelli L.; Curti M.; Rezk T.; Raimondi S.; Pepys M.B.; Hawkins P.N.; Gillmore J.D.; Taylor G.W.; Pinzani M.; Bellotti V. Increasing the accuracy of proteomic typing by decellularisation of amyloid tissue biopsies (work in preparation).
- **Canetti D.** et al. Methyl lysine formation in formalin-fixed paraffin-embedded amyloid tissue: misidentification of transthyretin and immunoglobulin variants (work in preparation).

5) Attended congresses/workshops/summer schools/contribution:

- **Canetti D.**; Monti M.; Chesi G.; Concilli M.; Polishchuk R.; Pucci P. Investigation of molecular mechanisms impaired in Wilson disease. 9th ItPA (Italian Proteomics Association) National Congress – Naples, Italy (24th – 27th June 2014). **Oral Communication.**
- Monti M.; **Canetti D.**; Porto C.; Parenti G.; Pucci P. Dissecting acid alpha glucosidase pathways in pathophysiology of Pompe disease. 9th ItPA (Italian Proteomics Association) National Congress – Naples, Italy (24th – 27th June 2014). **Poster.**
- Monti M.; **Canetti D.**; Chesi G.; Concilli M.; Polishchuk R.; Pucci P. Investigation of molecular mechanisms impaired in Wilson disease. 13th HUPO (Human Proteome Organization Word Congress) – Madrid, Spain (5th – 8th October 2014). **Poster.**
- **Canetti D.**; Monti M.; Zappa F.; De Matteis A.; Polishchuk R.; Pucci P. Investigation of CLIC-1 interactome. 10th EuPA (European Proteomics Association) Annual Congress – Milan, Italy (23rd – 28th June 2015). **Poster.**
- Landolfi A.; **Canetti D.**; Cozzolino F.; Battaglia E.; Zuccato C.; Monti M. Label free quantitation: comparative analysis of XIC and SpCs based methods. 11th ItPA (Italian Proteomics Association) National Congress – Perugia, Italy (16th -19th May 2016). **Poster.**
- Japan-UK workshop on Amyloidosis – London, United Kingdom (18th -20th September 2016). **Attendee.**
- Cozzolino F.; **Canetti D.**; Iacobucci I.; Landolfi A.; Moccia F.; Napodano E.; Pucci P.; Monti M. Functional proteomics in the elucidation of cellular mechanisms. Workshop BIO/10 -Naples, Italy (18th January 2017). **Poster.**