Romanian Academy

Institute of Biochemistry

Summary of Ph.D Thesis

Role of mannosidases in Endoplasmic Reticulum associated degradation

Scientific coordinator:

DR. ANDREI J. PETRESCU

PhD. Student:

CRISTIAN MARIAN BUTNARU

BUCHAREST

2017

Table of contents

Aims of these studies

The main aim of this thesis is to contribute to a better understanding of the folding and degradation mechanisms of conventionally secreted proteins, with emphasis on ER degradation enhancing α mannosidase like protein 3 (EDEM3). The role of this protein is not yet fully understood, as only certain aspects were investigated and the few existing reports propose alternative facts.

In order to extend the current available picture and to achieve the above-mentioned goals, we used a panel of techniques comprising: *molecular and cellular biology*, *protein biochemistry*, *mass spectrometry* (MS) *analysis*, *bioinformatics* and *molecular modeling* in an attempt to clarify the role of EDEM3 in ERAD.

The first chapter mainly focuses on proteomic aspects, having as main goal identifying the possible network of interactors for EDEM3 and how this is adjusted upon chemical inhibition of the glycosylation pathway.

The second chapter of this thesis had as primary aim to characterize EDEM3 making use of biochemical and molecular modeling methods to generate an accurate model of the domain organization of this protein.

For the third and fourth chapters of the present study, the aim was to understand the role of EDEM3 subdomains for association with client proteins and partners in the context of ERAD.

In order to establish the role of EDEM3 for ERAD substrate degradation two well-known glycoprotein models were used: tyrosinase and α1-antitrypsin. Tyrosinase is a membrane spanning glycoprotein, synthesized and folded in the ER following transport towards the melanosomes where it functions as a rate-limiting enzyme in melanin synthesis. Being heavily glycosylated and requiring a longer time to achieve its native conformation can expose this protein to possible mutations and abnormalities; thus, different pathologies (i.e. oculocutaneous albinism or melanoma) have been linked to it. Since melanoma is one of the deadliest forms of cancer and one of the proteins modified in cancer is tyrosinase, it is important to understand how this protein is processed in melanoma cells. As ERAD substrates we used soluble tyrosinase constructs that were previously characterized in our lab; and alpha 1-antitrypsin, a well-studied model for soluble glycoprotein secretion and its mutant NHK, employed in research studies as typical ERAD substrates. Mutations of this gene produce liver cirrhosis and lung emphysema, due to aggregate accumulation in hepatocytes and lack of transport and inhibitory function to the lungs.

By the approaches mentioned above in tackling the EDEM3 functioning problems I was able to identify and confirm a number of EDEM3 interactors by coupling immunoprecipitation with advanced MS interactomics.

We were further able to assess the changes induced in the network of interactors and their abundance by altering the ER glycosylation pathway, presumed to affect ERAD, with two glycosylation inhibitors.

By using bioinformatics, structure prediction and molecular modeling techniques, we also obtained structural insights onto the EDEM3 organization.

Based on modeling results I generated and characterized truncated forms of this protein and experimentally characterized them.

Finally, yet importantly, I investigated the requirement of the deleted domains in EDEM3 for association with ERAD substrates and partner proteins using cell biology and biochemistry methods as well as siRNA-mediated knockdown.

Introduction

Protein folding is a crucial process in ensuring proper cell functioning and implicitly tissue homeostasis. One-third of the eukaryotic cell proteins are synthesized on ER bound ribosomes, translocated and glycosylated co-translationally, and folded inside the ER with the help of molecular chaperones and enzymes. However, as any cellular process, protein folding is not 100% efficient, therefore, the requirement of a quality control checkpoint emerged. Protein quality control implies monitoring the native folding state of proteins from the ER and if the checkpoint is passed the protein is packed into secretory vesicles and exported to the secretory pathway. However if the protein fails to attain native conformation it is targeted to ERAD (ER-associated degradation pathway) for degradation in the cytosol and recycling of basic components-amino acids (Figure 1).

It is important to understand the mechanism regarding the degradation of folding incompetent proteins due to the fact that rapid disposal of these polypeptides is necessary to ensure the cellular homeostasis (Meusser et al., 2005). If the necessary degradation rate is not achieved, then accumulation of aberrant proteins may impede the ER capacity to support the maturation of newly synthetized polypeptides leading to cell malfunction and eventually death. N-glycan trimming was proposed to be the signal for selection of misfolded proteins to ERAD due to the finding that inhibiting the removal of mannose residues will protect the non-native proteins from degradation, thus proposing a mannose timer concept. This notion suggested a progressive protein de-mannosylation that will end the maturation stage and start a cascade of events to clear the ER and translocate to the cytosol the terminally misfolded polypeptides for degradation (Tamura et al., 2008).

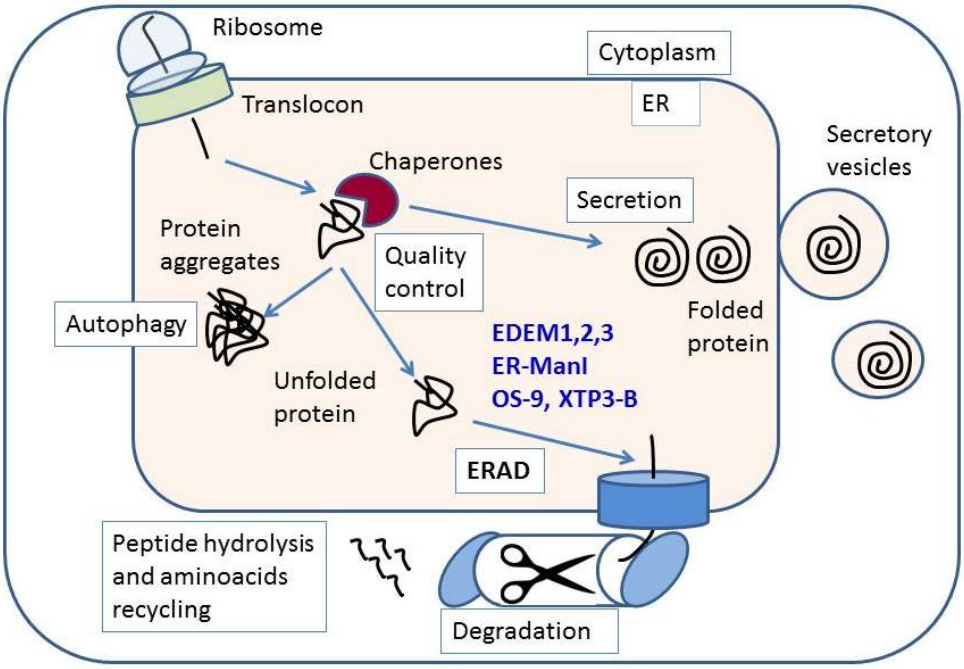

Figure 1. Schematic representation of the protein folding and quality control in the ER

In order for these events to happen, a set of specialized proteins are necessary to ensure the correct functioning of the process, some of the most important being: EDEM 1, EDEM 2, EDEM 3, OS-9, XTP3-B, ER ManI, SEL1L and HRD1.

Over the past decade there had been intense debates upon the role of the three mammalian EDEM genes (EDEM1, 2 and 3), members of the Glycosyl Hydrolase 47 (GH47) family along with the endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase (ERManI) that presents a 40% similarity in sequence and conserve their mannosidase-like domain. Although the ERAD pathway is well characterized, especially in yeast, in the mammalian system it is yet unclear the role of each protein. In agreement with the mannose timer hypothesis, this process requires a two-step mannose trimming process and it is believed the first one is made by ERManI converting Man9GlcNAc2 (M9) to Man8GlcNAc2 (M8B) allowing one of the three EDEMs to trim from M8B the α1,6 mannose bond on the B arm of the glycan. The puzzle consists on which one of them is the second rate-limiting enzyme.

In 2003, Hosokawa and others proposed that overexpressing EDEM1 and EDEM3, but not EDEM2, stimulates mannose trimming at various steps (Hosokawa et al., 2003). This helped reinforcing the idea that ERManI is the first enzyme followed by EDEM1 and EDEM3 and that the mannosidase domain of EDEM2 is inactive. These results were in contradiction with the previous article of the same lab in which they show EDEM1 lacks the α1,2 mannosidase activity and the involvement of ERManI in the formation of Man7-5GlcNAc2 based on knockdown experiments and biochemistry. Additionally, by 2009 and 2010 the controversy got deeper as Cormier et al, Tamura et

al and our lab showed that EDEM1 is able to recruit misfolded non-glycoproteins and deliver them for destruction through the SEL1-HRD1 complex and questioning whether it functions as a mannosidase or a lectin (Cormier et al., 2009, Tamura et al., 2010, Marin et al., 2012). By 2014 and 2015, another point of view was added to the puzzle by Ninagawa et al., showing through transcription activator-like effector nuclease-mediated gene knock-out of EDEM1, 2 and 3, and ERMan1, that EDEM2 has mannosidase activity and is the first step in trimming from M9 to M8 followed by EDEM3, which was responsible for converting M8 to M7-6. Furthermore, the authors proposed the idea that knockdown of all three EDEMs proteins activates a non-glycoprotein degradation pathway in order to maintain the cellular homeostasis (Ninagawa et al., 2015, Ninagawa et al., 2014).

As mentioned above, the N-glycans appear to serve as tag that provide not only the necessary information about protein maturation, but also act as binding site for mannose-6-phosphate receptor homology (MRH) domain. Trimming of outermost α1,2-linked mannose on C-arm of high-mannose-type glycan and binding of processed α1,6-linked mannosyl residues by the MRH domain are critical steps in guiding misfolded glycoproteins to enter ERAD; this process is thought to be fulfilled by OS-9 and XTP3-B. According to Christianson et al., the MRH domain of these proteins is required for interaction with SEL1L but not with ERAD substrates, and the lectin activity of OS-9 is not required for binding to misfolded glycoproteins (Christianson et al., 2008). In contrast to these results, other groups reported the specificity of the MRH domains for the M1-6 glycans and its importance in guiding the substrates towards degradation.

Because hydrophobic patches are exposed on misfolded proteins, they tend to aggregate. In order for the ER to maintain normal conditions, ERAD substrates need to be disposed. This process has been shown to require them soluble, unfolded and without disulfide bonds. One of the most intense studied ERAD pathways is the one mediated by the protein Sel1 homolog 1 (SEL1L) and the E3 ubiquitin- protein ligase synoviolin (HRD1). This complex is proposed to be the main route through which the misfolded substrates take to the proteasome due to adaptor role of SEL1L in the substrate recognition and linking to the HRD1. The importance of SEL1L also lies in its ability to interact with Derlin-1 and 2, VIMP, UBXD8, AUP1 and HERP, which helps recruiting VCP/p97 complex to drive the dislocation of the substrate across the ER membrane.

As emphasized above, the cell is able to regulate its function through well-defined processes some of them well understood and others not elucidated yet. In physiological conditions, in addition to a tight control of protein folding, to maintain the specific concentration of proteins in the cell, constant protein degradation is encountered, process generally called protein turnover. The primary process regulating ER homeostasis is ER associated degradation (ERAD) through which proteins from the endoplasmic reticulum are transported to the cytosol and degraded in the proteasome.

Currently, the most likely model for ERAD is an adaptive one, where dynamic network of interacting functional complexes facilitates the recognition, recruitment, dislocation, extraction, ubiquitination and degradation of the diverse classes of secretory proteins.

Model proteins employed as ERAD substrates

Tyrosinase (figure 2) is constitutively expressed in melanocytes and melanoma cells where it is synthesized in the ER and transported through the secretory pathway to melanosomes for melanin synthesis. It is a glycoprotein with six or seven glycans exposed to the lumen of melanosomes and requires the coordination of two cooper ions to fulfil its enzymatic function (Olivares et al., 2003).

Mutations in tyrosinase, leading to enzyme inactivation and N-glycosylation abnormalities, were identified in oculocutaneous albinism IA (OCA IA). In melanoma cells tyrosinase is either hyperactivated (pigmented melanoma), down-regulated at RNA level or synthesized, but not active, due to deficient transport to melanosomes (amelanotic melanoma). In most cases tyrosinase mutations lead to ER retention and transport through ERAD for degradation, thus this protein is a good model to study glycoprotein degradation associated to the ER (Popescu et al., 2005, Branza-Nichita et al., 2000).

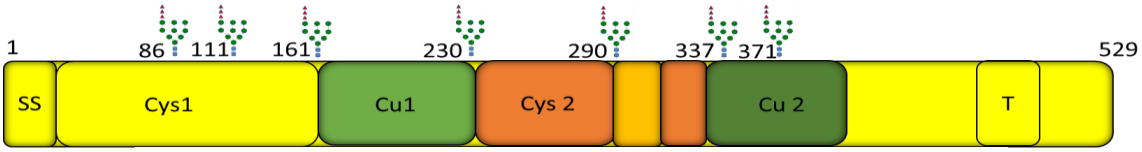

Figure 2. Schematic representation of tyrosinase structure

Alpha 1-antitrypsin (AT) is a monomeric glycoprotein predominantly synthesized in liver cells that functions as protease inhibitor in the lungs. The degradation of AT from the ER was found to be ERAD dependent and considering the high number of studies conducted on alpha 1-antitrypsin it was proposed as a good model for studying the role of N-glycans in folding and degradation of glycoproteins associated to the endoplasmic reticulum (Termine et al., 2005).

Two of the numerous mutations identified for AT were found to be secretion incompetent or secretion impaired, NHK (truncated AT) and ATZ (point mutation) respectively (figure 3). Most NHK forms are terminally misfolded, implicitly retained in the endoplasmic reticulum and subjected to proteasomal degradation. However, ATZ expression favors polymer formation and hinders the secretion of monomeric protein and is discarded both by proteasomal or lysosomal degradation, involving canonical autophagy (Graham et al., 1990).

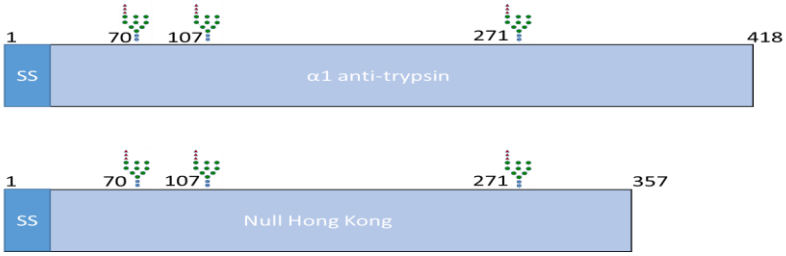

Figure 3. Schematic representation of alfa 1-antitrypsin and its truncated mutant NHK.

Results

Proteomics. Identifying the possible interactors of EDEM3.

A major limitation of the protein-protein interaction (PPI) network data is their static representation, which neglects the temporal and spatial organization of protein dynamics and as well as the effect of posttranslational modifications (PTM). Furthermore, this PPI may differ due to perturbations given by diseases or other external factors. The nature of protein–protein interaction is thus an inherently dynamic process that changes with time, environment, cell cycle or factors and will require further investigation to unravel a complete view. To cope with all the variable parameters that define the proteome different strategies, methods and techniques were combined to enlarge the field of view and better understand how these networks adapt. Due to the last technological improvements, we can now assess more than just one parameter of the proteome (ex: abundance, PTM) using a wide range of biochemical assays and instruments. A method that has emerged in recent decades and managed to seize a well-earned position is mass spectrometry. Mass spectrometry based proteomic approaches specifically aim to combine improvements in instrumentation and analytical procedures with experimental designs, focusing on annotating the proteome with multi-dimensional biological information, thus in turn, allowing the user to distinguish pools of proteins that behave differently and examine their properties (Yates et al., 2009).

In this study, we aimed to employ a systematic, multi-layered approach that integrates bioinformatics tools, mass spectrometry and biochemical assays in order to elucidate the interconnectivity of EDEM3 with ERAD partners.

Identifying the possible association clients of EDEM3 through immunoprecipitation coupled with mass spectrometry requires optimization of each step in order to efficiently and correctly pinpoint the interactors. Briefly, the cells were grown to a confluence of 90%, the lysis buffer included Digitonin 1% (it preserves better the protein complexes) and as elution buffer was used 150mM ammonium hydroxide pH 12, due to a better compatibility with mass spectrometry and its volatility that allowed the sample to be concentrated (Figure 4).

As the conditions were optimized for efficient extraction and conservation of protein-protein interaction, we proceeded in determining the interaction network that EDEM3 follows and how it is modified upon chemical treatment.

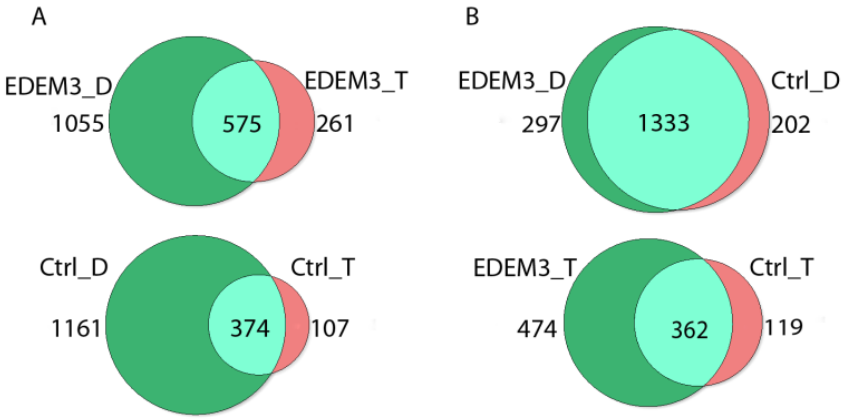

Figure 4. Venn diagram depicting the common and specific proteins identified by different cell lysing conditions. Digitonin favors the preservation of protein complexes, while Triton provides a better extraction for proteins.

In order to assess statistical significance, the experiments were made as biological triplicates. Concisely, A375 cell stably overexpressing EDEM3 or A375 control cells, were treated overnight with: a) kifunensine, a specific inhibitor of glycan de-mannosylation that blocks the activity of ERmannosidases; and b) N-butyl deoxynojirimycin (NB-DNJ), an iminosugar that blocks the activity of α-glucosidases I and II and prevent glucose trimming. An important effect of the inhibition of the N-glycosylation process is the increased interaction of immature polypeptides with the folding chaperones calnexin/calreticulin in the presence of kifunensine, versus the abolishment of these interactions by NB-DNJ (Hering et al., 2005, Petrescu et al., 1997).

Due to the fact that a large number of proteins are identified by MS, a statistical analysis of the acquired spectrums was made using MaxQuant. Briefly, the workflow consists in setting up the conditions in which the experiment was made, namely the enzyme used, missed cleavage sites, mass fragment tolerance, and posttranslational modifications as well as the database type. The quantification method used was labeling free (LFQ), and was made based on the peak intensities. The obtained LFQ values were further analyzed using Perseus, where the obtained list of possible interactors was annotated by adding the gene, protein name and subcellular localization.

The obtained data was hierarchically clustered by rows and columns using the following parameters: *distance*, a Euclidean function that describes how near two pairs are; and the *linkage*, which describes how the distance between two clusters is defined, based on the distances between single items. As shown in figure 5, we found that for the column clustering, the overall profile of the identified proteins is different upon treatment with inhibitors compared to the control samples. As for the row clustering, EDEM3 measured intensity was higher in the untreated cells overexpressing the protein; while for the chemically treated cells other proteins were more intense than our target protein, namely: GRP78, GRP94, CNX, CRT, PDI, PDIA4. The treatment with kifunensine or NBDNJ lowered the EDEM3 intensity compared to untreated cells, whilst the NBDNJ treatment induced the overexpression of ERAD and ERQC components HSPA8, BiP, GRP94, ERP60, PDIA3 and PDI.

Interestingly, aside these expected ER partners, figure 5 indicates that also some cytosolic proteins related to degradation such as UBA1 and UBA2 were pulled down in both the presence/absence of inhibitors suggesting possible functional association with EDEM3. Based on these analyses we can conclude that the interaction network of EDEM3 is changed upon chemical inhibition of different pathways as different proteins are having their expression modified in these cases. Furthermore, by using a statistical approach a possible list of association candidates could be formed, allowing us to further validate these hypotheses through complementary techniques.

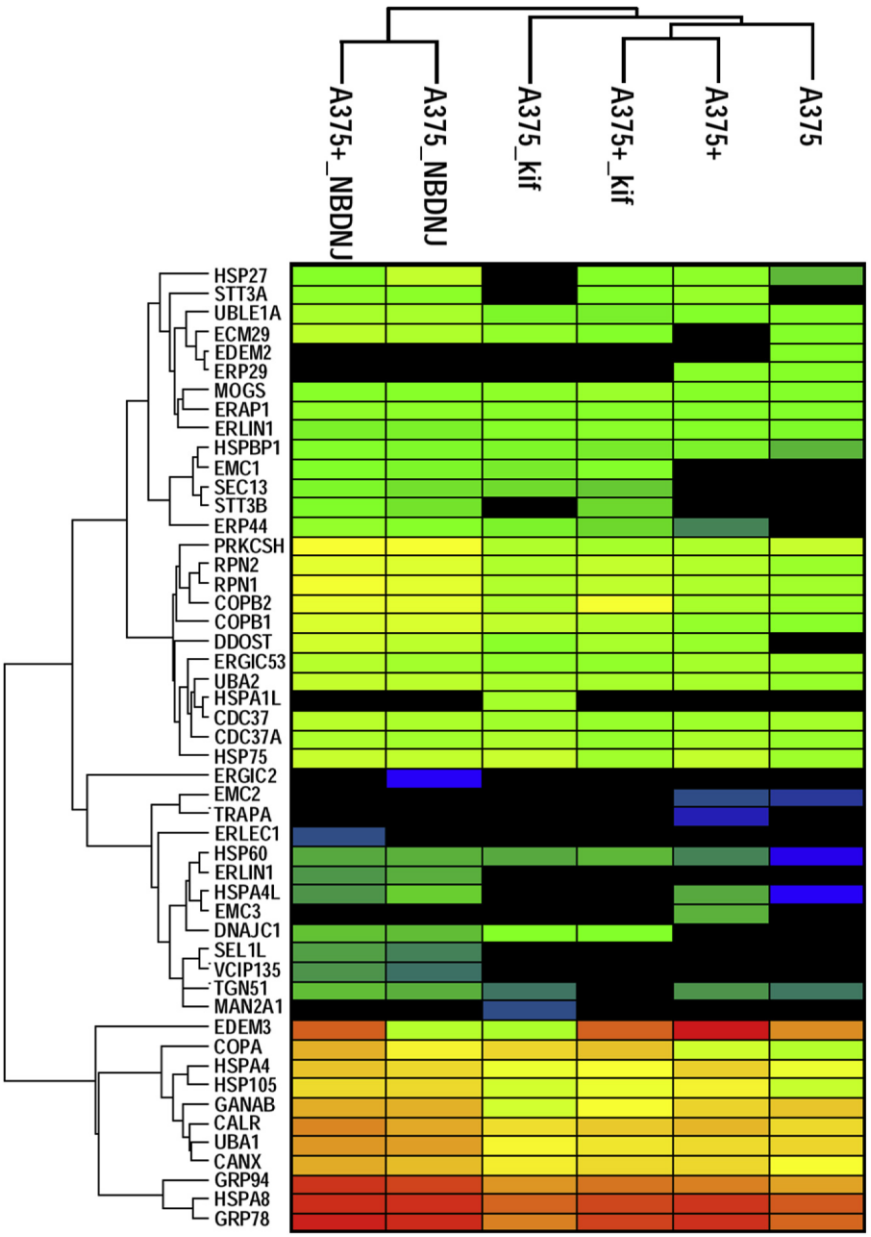

Fig. 5 Hierarchical clustering of EDEM3 interactors in the absence/presence of N-glycan processing inhibitors. MS based quantification of EDEM3 interactors in A375 cells in the absence (A375) or presence of kifunensin (A375kif) or NB-DNJ (A375-NBDNJ); and in A375 cells stably expressing EDEM3 in the absence (A375þ) or presence of kifunensin (A375 þ kif) or NB-DNJ (A375 þ NBDNJ). The rainbow code is based on ⟨log2(LFQ)⟩3 values. These run from blue (~15e17), to green(~17e22), to yellow (~22e28) and red (~28e31). Black means absence from IP. Species present also in the control were not included.

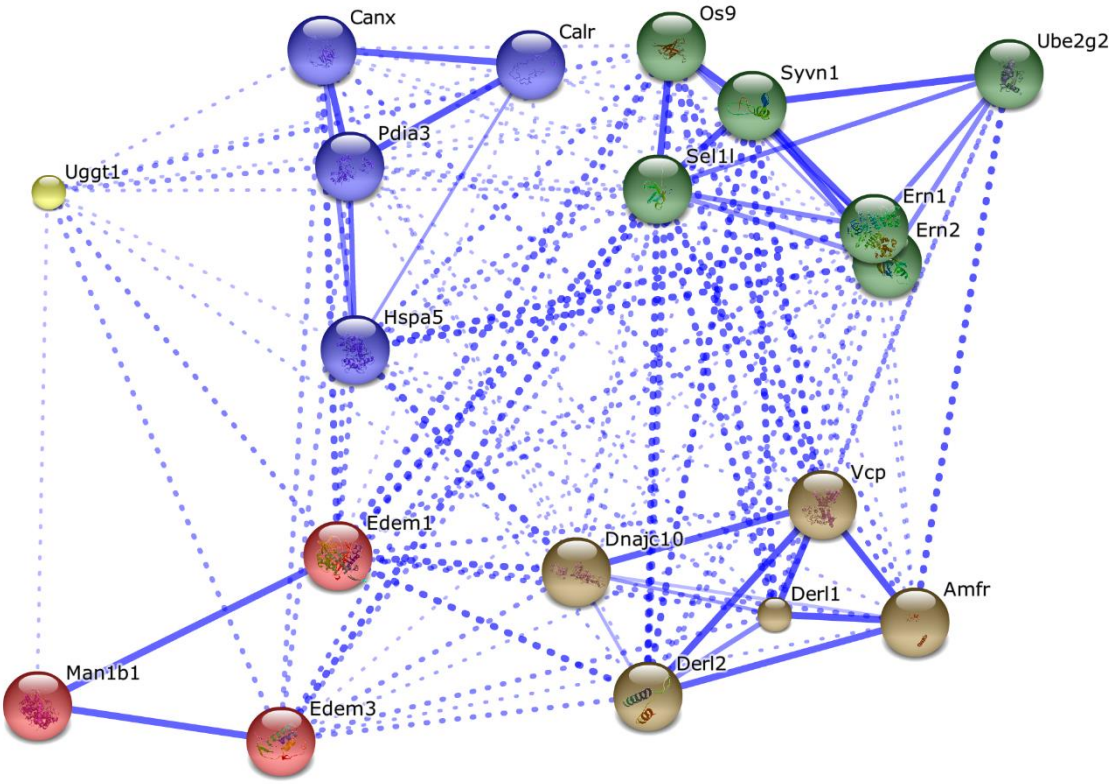

Figure 6. Predicted proteins association map. Shown here is a representation of the possible clusters that could be formed with the predicted proteins. The dashed lines represent inter-cluster edges

Based on hits provided by the statistical analysis of MS/MS data and the documented role in the ER, the following residents were selected as a first group of inputs for network mapping using STRING (figure 6): EDEM3, SEL1L, Dnajc10, CNX, CRT and GRP94. In addition, STRING was allowed to enrich the network by adding up to 10 direct and 5 indirect putative interactors sourced from: scientific literature, documented protein-protein interactions and curated databases. This simulation ascribed Man1B1, SEL1L, Uggt1 and OS-9 as main EDEM3 interactors. Further, map refinement was performed based on Markov Cluster algorithm (MCL). STRING global score from the three selected sources was used to obtain a distance matrix that was inputted into clustering procedure. By this, the contenders were divided into three clusters colour coded. Here the dashed lines show group interconnectivities while thickness describes the interaction strength.

As observed above, the number of potential interactors of EDEM3 determined by mass spectrometry analysis is high and in certain cases false positives may arise, thus it requires complementary methods to validate some of the targets. In this case, we chose to validate some of the proteins identified by mass spectrometry through immunoprecipitation and Western blotting. Therefore, we first selected a number of proteins considering the difference observed between the

overexpressed protein and the endogenous one, potential interaction already described in the literature, involvement in ERAD or ER quality control, or proteasomal degradation.

As depicted in figure 7A, EDEM3 was co-immunoprecipitated with calnexin, suggesting that EDEM3 could be establishing a direct interaction with calnexin or it may have functional association via the substrate proteins that either are bound to calnexin for productive folding or extracts proteins exiting this cycle. Moreover, we also tested SEL1L as an EDEM3 interactor (Figure 7B). SEL1L was already reported to associate with EDEM3, and we confirmed it by both mass spectrometry and IP coupled with WB, despite its low difference between the overexpressed and endogenous level of EDEM3.

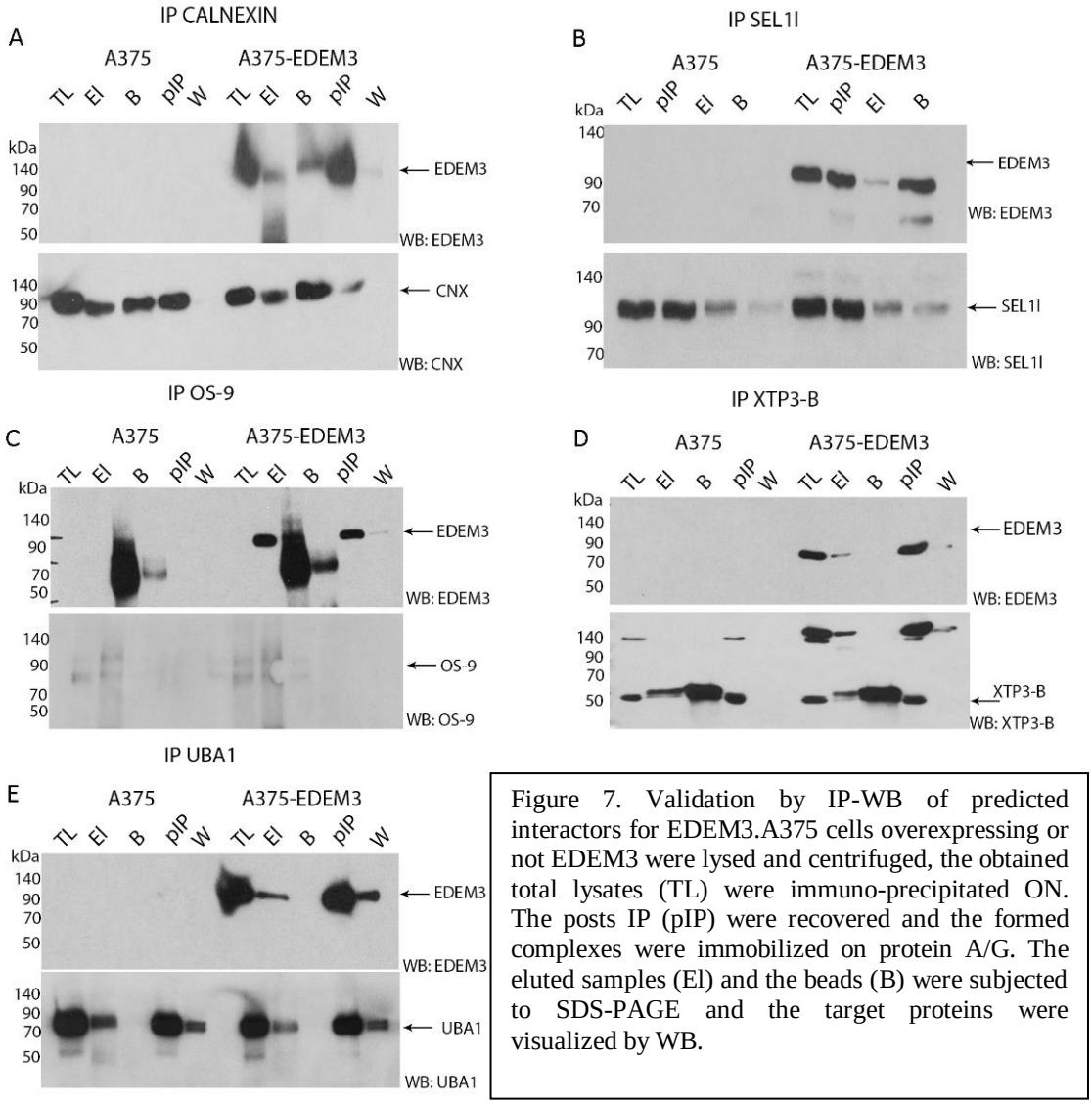

Figure 7. Validation by IP-WB of predicted interactors for EDEM3.A375 cells overexpressing or not EDEM3 were lysed and centrifuged, the obtained total lysates (TL) were immuno-precipitated ON. The posts IP (pIP) were recovered and the formed complexes were immobilized on protein A/G. The eluted samples (El) and the beads (B) were subjected to SDS-PAGE and the target proteins were visualized by WB.

Two proteins, OS-9 and XTP3-B, were proposed to function as dedicated lectins in ERAD, recognizing and binding oligomannosidic glycans via their mannose 6 phosphate receptor homology (MRH) domains. We hypothesized that the presence of these domains that allow them to interact with

misfolded proteins and thus they may associate to EDEM3 or substrates bound to EDEM3. As observed in figure 7C and D, weak association with EDEM3 was detected for both OS-9 and XTP3-B, although only the association with OS-9 was predicted by STRING, and both are detected more abundantly for the overexpressed EDEM3, most probably due to the higher amount of protein.

Another protein we identified by mass spectrometry was UBA1 (Ubiquitin like modifier activating enzyme 1), reported to catalyze the first step in ubiquitin conjugation in order to tag misfolded protein for degradation through the ubiquitin-proteasome system. In comparison to the other two enzymes involved in ubiquitin tagging the ubiquitin conjugating enzymes (~35 proteins) and ubiquitin ligases (>1000 proteins) in humans, ubiquitin activating enzymes are encoded by 9 genes; amongst them the one related to ERAD is UBA1. Thus, we investigated whether this association was real by confirming it through IP and WB. As shown in figure 7E, the detected interaction appeared to be strong and stable in comparison to the other proteins tested, therefore leading us to hypothesize that EDEM3 might function as a mediator between the ERQC and ERAD.

In addition, the herein results indicate that interference with N-glycan processing modulates the interactome composition. Thus, EDEM3 interacts with components of the ERAD dislocation machinery, such as SEL1L, HRD1, and ERLEC1 when the interaction of calnexin with glycoproteins is impaired. In contrast, there is a constant association of EDEM3 with calnexin irrespective of the glycosylation status of the ERAD clients. These data may indicate that EDEM3 associates with the dislocon components to deliver the misfolded proteins unassisted by the calnexin cycle, which is impaired when glycan trimming is blocked. Whilst a clear association of EDEM3 with calnexin occurs, this interaction is not required for the actual EDEM3 associated dislocation of the ERAD substrates. Hence, EDEM3 interaction network is inherently dynamic in nature and changes with stress induced by altering the normal N-glycan processing. Other factors affecting it, such as cell cycle, environment or other stress conditions, require further investigation to unravel a complete picture of the EDEM3 pathway.

Building up 3D models of EDEM3 domains

EDEM3 is a member of the GH47 family of proteins predicted to have a mannosidase-like domain, a protease associated domain and an ER retrieval sequence (KDEL). Here we aimed to dissect structural aspects of this protein by using bioinformatics tools for structure prediction, domain identification and molecular dynamics.

Using the methods mentioned above we were able to predict three major domains for EDEM3: a mannosidase domain, a protease associated (PA) domain and an intrinsically disordered (ID) domain, connected with each other through two linkers. Due to the length and complexity of the protein, each domain was modeled separately by homology and remote homology, as further described.

The mannosidase- like domain (E3MAN) of EDEM3 is the most extended comprising amino acids 50 to 500 from EDEM3 structure and thus was the first domain we wanted to investigate. A homology model was built based on a crystal structure template (PDB code - 1NXC) (Tempel et al., 2004) belonging to the alfa 1-2-mannosidase IA from *Mus musculus*. Sequence alignment between the target and the template revealed a ~30% identity and ~50% similarity for the EDEM3 mannosidase-like domain with the crystal structure. The secondary structure consensus shows a very good fit with the secondary structure of the template and disorder predictions indicate that the mannosidase domain is highly ordered, excepting the region comprising aa 83-96 at the N-terminus. The structural model of the mannosidase domain was raised against this template with specific refinements and is presented in figure 8.

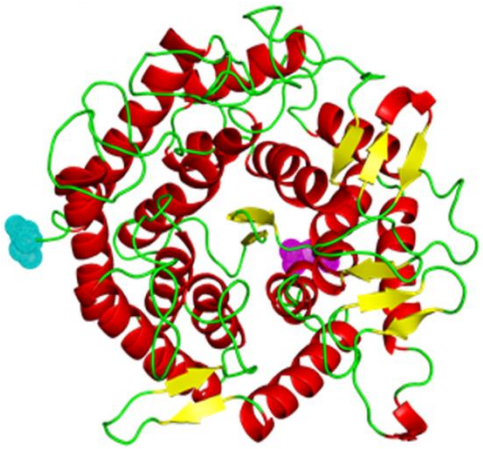

Figure 8. Homology model of the mannosidase domain of EDEM3 based on the template 1NXC. Red represents alpha helix, yellow depicts beta sheet structures and in green are represented the linker regions

EDEM3 sequence revealed the existence of a protease associated (PA) domain (aa 679-779 – E3PA) that was modelled using two templates: an *Escherichia coli* aminopeptidase (PDB code - 2EK9) and a *Spodoptera frugiperda* viral protein (PDB code - 3KAS) (Abraham et al., 2010). The resemblance with the templates shows a ~17.9% identity & ~33.3% similarity with the aminopeptidase, and ~14% identity & ~26% similarity with the viral protein respectively (figure 9).

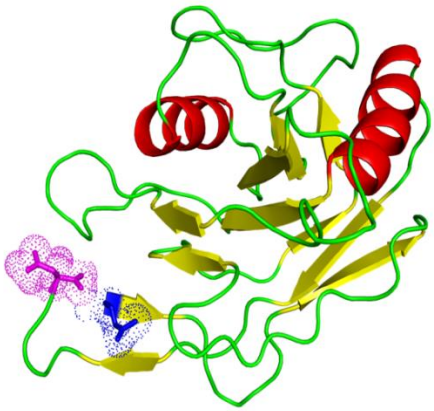

Figure 9. Remote homology model of the protease associated domain based on the templates 2EK9 and 3KAS. Red represents alpha helix, yellow depicts beta sheet structures and in green are represented the linker regions

All the models were generated using INSIGHT II from Accelrys and were refined by rounds of implicit solvent Generalized Born and explicit solvent molecular dynamics (MD) experiments using NAMD (Phillips et al., 2005) on a 14 x HP BL280c G6 high performance computing cluster.

The simulation used the CHARMM36 force field and harmonic position restraints on the backbone of the protein in regions of secondary structure, whereas the loops were left to move freely to eliminate steric conflicts and bring the model to a lower energy minimum. The validation of the model accuracy used QA-RecombineIT method (Pawlowski et al., 2013).

The overall models were brought to (GDT_TS = 72.3; RMSD (model deviation form an optimal Cα pathway) = 2.16Å) for the mannosidase domain while for the PA domain (GDT_TS=61.34 and RMSD=3.03 Å), both scores corresponding to very good models according to the QA-RecombineIT validation model.

Design and characterization of EDEM3 truncated mutants

Correlating the obtained information from the bioinformatics approach, two mutants were generated in order to identify the potential role of each domain for EDEM3 in ERAD. The first mutant generated was the ΔIDD-EDEM3 by removing the intrinsically disordered domain, the second one was the ΔPA-EDEM3 mutants lacking the protease-associated domain.

The next step was to test the expression of the domains in HEK293T cells through transient transfection. The theoretical mass for each domain is: ΔIDD-EDEM3= 85 kDa, ΔPA-EDEM3= 87.35 kDa, and a band at similar molecular weight was detected for each mutant by Western blotting.

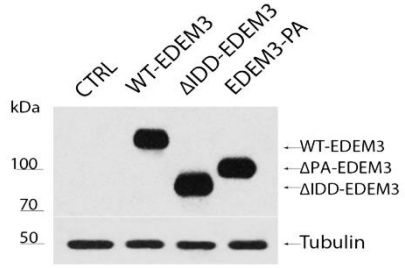

Figure 10 Expression of EDEM3 mutants transfected in HEK 293T cells was determined by Western blotting using mouse anti-HA antibodies.

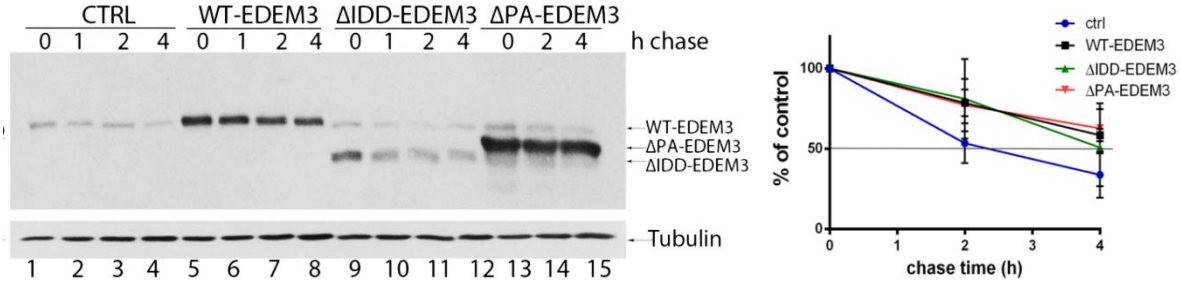

Figure 11. Half-life of EDEM3 truncates. To determine the rate of degradation of EDEM3 mutants, HEK293T cells overexpressing them were treated with cycloheximide and harvested at 0, 1, 2, and 4 hours. The graph shows the average percent EDEM3- mutants remaining after treatment for the specified times. Band intensity quantification was made using Image J software and the represented results are the mean of three independent experiments ±SEM

In addition, similar expression levels (figure 10) and half-life, determined by cycloheximide chase were observed (figure 11).

In this line, we were interested to investigate whether the intracellular localization of the truncated EDEM3 was similar to the wild type protein. Thus, an immunofluorescence experiment was performed to assess the co-localization of the EDEM3 truncates with calnexin, which was used as ER marker (Figure 12).

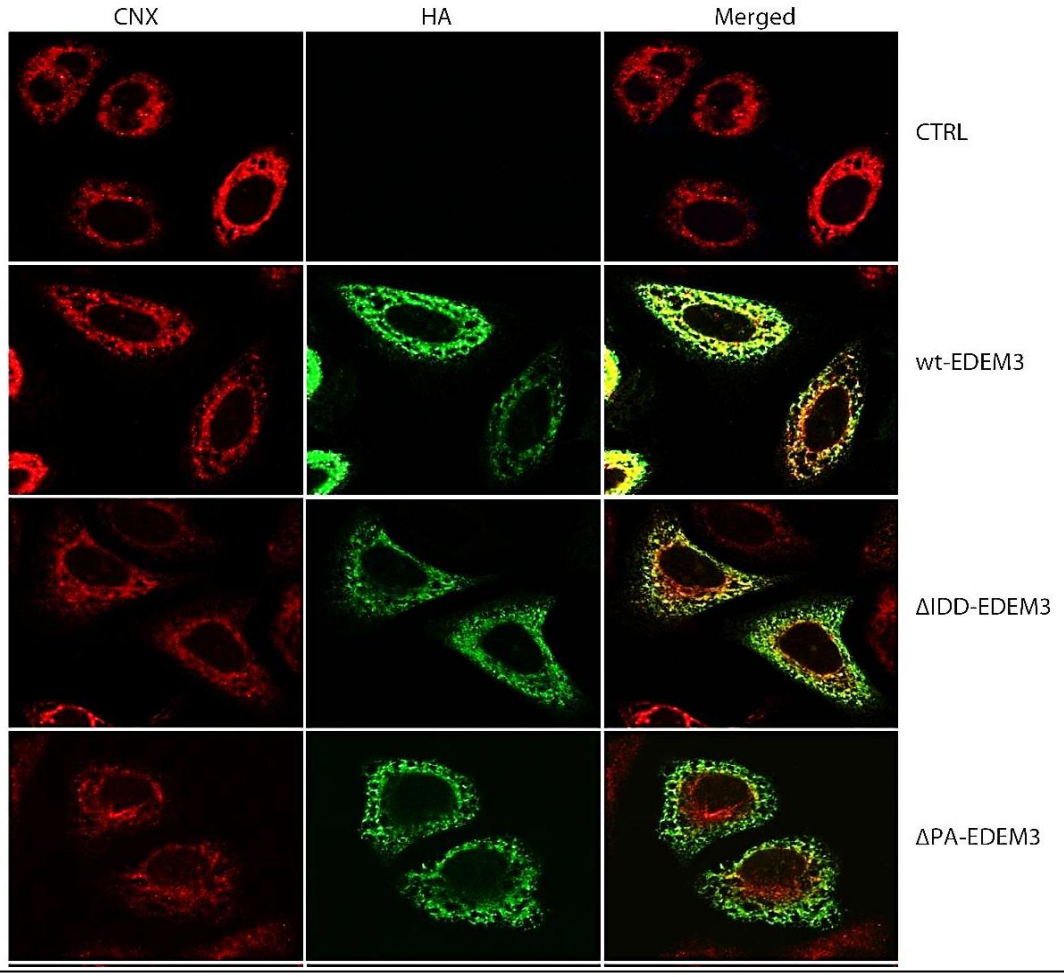

Figure 12. Co-localization of EDEM3 mutant domains with calnexin was assessed by transfection of target protein in HeLa cells followed by PFA fixation and permeabilization with Triton 0.2%. Next, the samples were probed with antibodies against HA to detect EDEM3 mutants and antibodies against calnexin, followed by incubation with fluorophore-coupled secondary antibodies.

Next, the EDEM3 mutants were used for co-expression experiments with model ERAD substrates to assess the effect of deleting different domains over the degradation of client proteins. For this purpose, tyrosinase and alpha 1-antitrypsin were employed as well-characterized substrates for proteasomal degradation. Moreover, the association of these mutants with the ERAD substrates was also studied. We observed here that deleting different domains of EDEM3 did not abolish the interaction of EDEM3 with ERAD substrates, implying that the association with client proteins is dependent on more than one of the EDEM3 domains. However, my results suggest that deletion of the ID domain enhances the degradation of the model ERAD substrates as determined by co-expression

and western blotting, as well as turnover rates monitored by cycloheximide chase. This might infer that the IDD of EDEM3 may function as a negative regulator and limits the activity of the WT-EDEM3, while in the absence of this domain the degradation activity induced by EDEM3 is increased.

Conclusions

By the above presented studies, we were able to gain insights into the structural organization of EDEM3 and assign the potential role of this protein in the ER processing machinery line-up.

Employing molecular and cellular biology, biochemistry techniques, advanced MS analysis, bioinformatics and molecular modelling we were able to predict, characterize the structural domains of EDEM3, and investigate their role in EDEM3 networking in ERAD.

The first approach I employed was to analyze EDEM3 interactors through MS analysis of immuno-isolated EDEM3 combined with *in silico* prediction of the network of interactors of EDEM3 to generate lists of potential interactors. These experiments were performed using normal cells and cells overexpressing EDEM3 in order to assess both endogenous and overexpressed EDEM3 interactors. This allowed us to assess interactor network modifications, stability and reorganization according to the amount of available EDEM3.

Using different inhibitors for protein glycosylation, the dynamics of the EDEM3 interactors network detected by mass spectrometry and quantification using MaxQuant software, changes probably as response to the stress factors. Moreover, I also confirmed the association of EDEM3 by IP and WB with calnexin, SEL1L, OS-9, XTP3-B and Uba1, proteins involved in ERAD that were identified as potential interactors by mass spectrometry.

To better understand how EDEM3 functions in ERAD I first attempted a short characterization of this protein that will add to the little information that is known about it in the literature. For this reason, I used chemical compounds that are used as specific inhibitors for different pathways to identify the potential degradation pathway for EDEM3. The experiments presented here suggest that the turnover and degradation of EDEM3 is modulated in a glycan dependent manner since it showed sensitivity to all glycosylation inhibitors and its half-life was extended in presence of kifunensine.

The following aim was to use secondary structure predictors, charge profile, and posttranslational modifications software to analyze and have a rough prediction for the structural organization of EDEM3 based on the amino acid sequence. All the information obtained from these predictions was combined with target to template modelling to identify the best-known structure that would fit the structure of the designated domains. Thus for EDEM3 structure predictions identified three major domains: a mannosidase-like domain, a protease-associated domain and intrinsically

disordered domain. I generated the 3D models of the first to domains with good confidence and the models were raised against the crystal structures of known domains.

To analyze the role of each domain predicted from the structure of EDEM3 I generated truncated mutants by deleting alternatively the predicted domains using advanced molecular biology methods. The mutants were first characterized in terms of expression and cellular localization and as presented here they are expressed and localize to the ER, as expected.

Next, I investigated the behavior of these mutants concerning substrate degradation and association with the client proteins. The results presented in this study suggest that the IDD of EDEM3 may be required for a stronger interaction with the substrates since deleting this domain weakens the association of EDEM3 with the substrate as demonstrated by immunoprecipitation and Western blotting complemented with pulse –chase and immunoprecipitation experiments. Additionally, the experiments of WB and cycloheximide chase suggest that the IDD of EDEM3 modulates its activity to accelerate the degradation of ERAD substrates. Moreover, when overexpressed EDEM3 lacking the PA domain does not induce an increased degradation of the ERAD substrates compared to the IDD suggesting this domain might be important for activity and functions as negative regulator in this context.

To investigate whether the deletion of these domains is important for association with SEL1L, a known interactor of EDEM3, I tested the association of these mutants with the endogenously expressed SEL1L. It seems that the association with SEL1L is not abolished by the deletion of any of the two domains, but it is weaker when the IDD domain is missing.

Considering the above-mentioned observations, I also tested whether the EDEM3 induced degradation of the substrates is SEL1L dependent or it can bypass this protein. Thus, the endogenous SEL1L was silenced and the effect that EDEM3 mutants had over the degradation of typical ERAD substrates and an important observation made here is that EDEM3 mutants can induce degradation of the substrates independent of the level of SEL1L. This could be explained by a hyperactivity of the remaining SEL1L to export misfolded proteins or another alternative pathway for ER export and degradation is activated when the classical one is impaired, the latter being more reliable.

In summary, in my attempt to understand the role of EDEM3 in ERAD several aspects concerning these proteins were established summarized as follows:

- I optimized a protocol for sample preparation from human serum to identify biomarkers using mass spectrometry
- The potential interactions of EDEM3 were identified by *in silico* simulation and mass spectrometry analysis and validated through biochemical methods
- Half-life and degradation of both endogenous and overexpressed EDEM3 is affected by glycosylation inhibitors as observed from WB and cycloheximide chase experiments

- Kifunensine impairs the EDEM3 induced degradation of the substrates, while NB-DNJ does not interfere with this process, as observed by co-expression in HEK293T cells and western blotting
- I generated the 3D model for two of the EDEM3 domains by homology modelling
- Based on the bioinformatics predictions I designed and generated truncated mutants of EDEM3 and characterized them
- Investigating the interaction of the generated mutants with selected ERAD substrates showed that the IDD of EDEM3 seems to modulate this association
- The EDEM3 mutant lacking the PA domain is less efficient in accelerating the degradation of the substrates
- Deletion of the upper mentioned domains of EDEM3 does not abolish the association with SEL1L, however lack of the ID domain weakens the interaction
- EDEM3 accelerates the degradation of ERAD substrate independent of the level of SEL1L possibly by activating an alternative pathway for dislocation from the ER.

Acknowledgements:

First, I would like to acknowledge Dr. Andrei J. Petrescu and Dr. Stefana M. Petrescu, for support and scientific guidance during all these years and to express my gratitude for trusting me with the research project based on which this thesis was conceived.

I owe a special appreciation to Dr. Marioara Chiritoiu for her guidance and discussions in designing the experiments presented in this thesis.

As well, I express my gratitude to Drd. Marius Surleac for input and advice, with whom I collaborated to generate the two models for the EDEM3 domains and for the critical discussions during the development of this project.

In addition, I would like to express my thankfulness to Dr. Cristian V.A. Munteanu for his help and discussions related to the results generated by mass spectrometry.

Many thanks to Dr. Gabriela Chiritoiu for generating the stable cell lines used in this study and the support and discussions for some of the presented experiments.

I owe special thanks to Dr. Simona Ghenea who helped me with all the molecular biology experiments performed during this project, for our scientific talks and her valuable advice.

Also, I have to thank all my colleagues from the Molecular Cell Biology and Bioinformatics & Structural Biochemistry departments for fruitful discussions and suggestions for my work.

Finally, I would like to thank my family for moral support, encouragement and understanding throughout this period.

List of publications:

1. **Butnaru CM,** Chiritoiu MB, Chiritoiu GN, Petrescu SM, Petrescu AJ, "*Inhibition of N-glycan processing modulates the network of EDEM3 interactors*", ***Biochemical and Biophysical Research Communications***, **486(4)**, 978-984, (2017);

   AI=0.7; IF=2.29

2. **Butnaru C. M.,** Munteanu C. V.A., , Chiritoiu G. N., Ghenea S., Petrescu A.J., Petrescu S. M., "Comparison of protein extraction conditions for EDEM3 interactors in melanoma cells", ***Romanian Journal of Biochemistry***, **52(1)** 2015, ISSN 1582-3318

   AI < 0.1; IF = 0

List of contributions to national & international conferences

1. September 2015: **Butnaru C. M.,** Chiritoiu M.B., Surleac M., Petrescu A.J., Petrescu S.M., "Investigating and assigning functions to EDEM3 domains" *Annual International Meeting of RSBMB, Bucharest, Romania*
2. June 2015: Ionescu A.I., Menţel M, Munteanu C. V.A., **Butnaru C. M.,** Badea R.A., Iancu I., Szedlacsek S.E. "Mass spectrometric investigations regarding phosphorylation of human EYA3 by SRC kinase", *Europhosphatase, Turku, Finland*.
3. June 2014: **Butnaru C. M.,** Chiritoiu M.B., Surleac M., Petrescu AJ, Petrescu S.M., "Role of EDEM3 in ERAD"-*Annual International Meeting of RSBMB, Oradea, Romania*
4. June 2014: Munteanu C. V.A., **Butnaru C. M.,** Jiglaru C., Ganea E., "Advanced mass spectrometrical analysis of glycated BSA, and comparative proteomic screening of temporal and hyppocampal old brain tissue" -*Annual International Meeting of RSBMB, Oradea, Romania*

List of papers in preparation

1. Chiritoiu M. B., Chiritoiu G.N., **Butnaru C. M.,** Munteanu C. V.A., Pastrama F., Ivessa N.E., Petrescu S. M.: "An intrinsically disordered region of EDEM1 selects ERAD clients for degradation" –***manuscript in preparation***
2. Cârciumaru A., **Butnaru C.M.,** Petrescu S.M., "Identification of new biomarkers for rheumathoid pathologies using mass spectrometry"-***manuscript in preparation***
3. **Butnaru C. M.,** Chiritoiu M. B., Chiritoiu G. N., Surleac M., Ghenea S., Petrescu S. M., Petrescu A.J., "Modelling and characterization of EDEM3 domains and investigating their role in ERAD substrates degradation"- ***manuscript in preparation***

Selected references:

ABRAHAM, J., CORBETT, K. D., FARZAN, M., CHOE, H. & HARRISON, S. C. (2010) Structural basis for receptor recognition by New World hemorrhagic fever arenaviruses. *Nat Struct Mol Biol,* 17**,** 438-44.

BRANZA-NICHITA, N., NEGROIU, G., PETRESCU, A. J., GARMAN, E. F., PLATT, F. M., WORMALD, M. R., DWEK, R. A. & PETRESCU, S. M. (2000) Mutations at critical N-glycosylation sites reduce tyrosinase activity by altering folding and quality control. *J Biol Chem,* 275**,** 8169-75.

CHRISTIANSON, J. C., SHALER, T. A., TYLER, R. E. & KOPITO, R. R. (2008) OS-9 and GRP94 deliver mutant alpha1-antitrypsin to the Hrd1-SEL1L ubiquitin ligase complex for ERAD. *Nat Cell Biol,* 10**,** 272-82.

CORMIER, J. H., TAMURA, T., SUNRYD, J. C. & HEBERT, D. N. (2009) EDEM1 recognition and delivery of misfolded proteins to the SEL1L-containing ERAD complex. *Mol Cell,* 34**,** 627-33.

GRAHAM, A., HAYES, K., WEIDINGER, S., NEWTON, C. R., MARKHAM, A. F. & KALSHEKER, N. A. (1990) Characterisation of the alpha-1-antitrypsin M3 gene, a normal variant. *Hum Genet,* 85**,** 381-2.

HERING, K. W., KARAVEG, K., MOREMEN, K. W. & PEARSON, W. H. (2005) A practical synthesis of kifunensine analogues as inhibitors of endoplasmic reticulum alpha-mannosidase I. *J Org Chem,* 70**,** 9892-904.

HOSOKAWA, N., TREMBLAY, L. O., YOU, Z., HERSCOVICS, A., WADA, I. & NAGATA, K. (2003) Enhancement of endoplasmic reticulum (ER) degradation of misfolded Null Hong Kong alpha1-antitrypsin by human ER mannosidase I. *J Biol Chem,* 278**,** 26287-94.

MARIN, M. B., GHENEA, S., SPIRIDON, L. N., CHIRITOIU, G. N., PETRESCU, A. J. & PETRESCU, S. M. (2012) Tyrosinase degradation is prevented when EDEM1 lacks the intrinsically disordered region. *PLoS One,* 7**,** e42998.

MEUSSER, B., HIRSCH, C., JAROSCH, E. & SOMMER, T. (2005) ERAD: the long road to destruction. *Nat Cell Biol,* 7**,** 766-72.

NINAGAWA, S., OKADA, T., SUMITOMO, Y., HORIMOTO, S., SUGIMOTO, T., ISHIKAWA, T., TAKEDA, S., YAMAMOTO, T., SUZUKI, T., KAMIYA, Y., KATO, K. & MORI, K. (2015) Forcible destruction of severely misfolded mammalian glycoproteins by the non-glycoprotein ERAD pathway. *J Cell Biol,* 211**,** 775-84.

NINAGAWA, S., OKADA, T., SUMITOMO, Y., KAMIYA, Y., KATO, K., HORIMOTO, S., ISHIKAWA, T., TAKEDA, S., SAKUMA, T., YAMAMOTO, T. & MORI, K. (2014) EDEM2 initiates mammalian glycoprotein ERAD by catalyzing the first mannose trimming step. *J Cell Biol,* 206**,** 347-56.

OLIVARES, C., SOLANO, F. & GARCIA-BORRON, J. C. (2003) Conformation-dependent post-translational glycosylation of tyrosinase. Requirement of a specific interaction involving the CuB metal binding site. *J Biol Chem,* 278**,** 15735-43.

PAWLOWSKI, M., BOGDANOWICZ, A. & BUJNICKI, J. M. (2013) QA-RecombineIt: a server for quality assessment and recombination of protein models. *Nucleic Acids Res,* 41**,** W389-97.

PETRESCU, S. M., PETRESCU, A. J., TITU, H. N., DWEK, R. A. & PLATT, F. M. (1997) Inhibition of N-glycan processing in B16 melanoma cells results in inactivation of tyrosinase but does not prevent its transport to the melanosome. *J Biol Chem,* 272**,** 15796-803.

PHILLIPS, J. C., BRAUN, R., WANG, W., GUMBART, J., TAJKHORSHID, E., VILLA, E., CHIPOT, C., SKEEL, R. D., KALE, L. & SCHULTEN, K. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem,* 26**,** 1781-802.

POPESCU, C. I., PADURARU, C., DWEK, R. A. & PETRESCU, S. M. (2005) Soluble tyrosinase is an endoplasmic reticulum (ER)-associated degradation substrate retained in the ER by calreticulin and BiP/GRP78 and not calnexin. *J Biol Chem,* 280**,** 13833-40.

TAMURA, T., CORMIER, J. H. & HEBERT, D. N. (2008) Sweet bays of ERAD. *Trends Biochem Sci,* 33**,** 298-300.

TAMURA, T., SUNRYD, J. C. & HEBERT, D. N. (2010) Sorting things out through endoplasmic reticulum quality control. *Mol Membr Biol,* 27**,** 412-27.

TEMPEL, W., KARAVEG, K., LIU, Z. J., ROSE, J., WANG, B. C. & MOREMEN, K. W. (2004) Structure of mouse Golgi alpha-mannosidase IA reveals the molecular basis for substrate specificity among class 1 (family 47 glycosylhydrolase) alpha1,2-mannosidases. *J Biol Chem,* 279**,** 29774-86.

TERMINE, D., WU, Y., LIU, Y. & SIFERS, R. N. (2005) Alpha1-antitrypsin as model to assess glycan function in endoplasmic reticulum. *Methods,* 35**,** 348-53.

YATES, J. R., RUSE, C. I. & NAKORCHEVSKY, A. (2009) Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng,* 11**,** 49-79.