

Molecular Characterization of Calreticulin Mutants Implicated in Sudden Unexplained Death

by

Hector Vega

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Biochemistry

University of Alberta

© Hector Vega, 2018

## Abstract

Sudden unexplained death (SUD) is a term used when comprehensive medical examination and autopsy fail to find a conclusive cause of sudden death. In cases of SUD, post-mortem molecular and genetic evaluation methods have helped identify ion channel abnormalities leading to arrhythmias as cause. However, many cases of SUD remain unexplained by ion channel mutation evaluation.

Using whole exome sequencing (WES) analysis, followed by strategic variant filtration, several mutations in the calreticulin gene (*CALR*) have been identified in cases of SUD. One of these mutations leads to an ultra-rare frame-shift in *CALR*, encoding a truncated mutant protein, *CALR*<sup>376fs</sup>. *CALR* is an endoplasmic reticulum (ER) Ca<sup>2+</sup>-binding protein and molecular chaperone for proper folding, assembly, and retention of secreted and membrane proteins, among several other functions. The involvement of *CALR* mutants in the SUD phenotype is puzzling and the mechanism is unknown.

Molecular characterization of the *CALR* mutants compared to wild-type protein was performed, with focus on *CALR*<sup>376fs</sup>. It was found that the mutant proteins were folded different compared to wild-type *CALR*. Further analysis of *CALR*<sup>376fs</sup> indicated that it was also highly degraded and had impaired chaperone function. Traffic to the plasma membrane and activity of a Ca<sup>2+</sup> ion channel linked to SUD has recently been found to be higher in presence of *CALR*<sup>376fs</sup> compared to wild-type protein. The present suggests that altered structure, decrease in abundance, and impaired chaperone function of *CALR*<sup>376fs</sup> lead to a defect in ER quality control and higher activity of a Ca<sup>2+</sup> ion channel likely to cause cardiac arrhythmia and SUD.

## Acknowledgements

The present work was made possible thanks to the Michalak lab and former lab member Dr. Elzbieta Dudek, who started the project, provided tools, designed and ran experiments, and was always available to share information and knowledge for continuation of the project. This work was also possible thanks to Dr. Michael Ackerman's research group, which identified the mutants, and shared their manuscript and data. Thanks as well to Dr. Evaldas Čiplys and his lab for purification of high-quality proteins and for help with the methods section.

I am thankful and in debt to my awesome labmates for all their help and support in and out of the lab. I learned a lot with them, and they created a warm, fun, and positive atmosphere. Special thanks to Dr. Jody Groenendyk for that and also providing constant invaluable expertise with experiments and theory. I am forever grateful to our supervisor Dr. Marek Michalak for mentorship and guidance, for the support and challenges, for teachings and training, for great jokes, and for sharing his wisdom about science and life to help us grow as curious individuals.

Many thanks to Dr. Nicolas Touret and Dr. Todd Alexander for their expertise, input, evaluation, and time dedicated to my supervisory committee. Lastly, thanks to the Department of Biochemistry for all the training, and for providing a great setting and a stimulating environment.

## Table of contents

### CHAPTER ONE: Introduction

Sudden unexplained death.....	2
Molecular autopsy.....	3
Cardiac channelopathies.....	4
Long QT syndrome.....	4
Short QT syndrome.....	5
Brugada syndrome.....	5
Catecholaminergic polymorphic ventricular tachycardia.....	6
Whole exome sequencing.....	7
Novel mutants identified and case index.....	8
CALR.....	12
N-domain.....	13
P-domain.....	14
C-domain.....	14
Additional functions of CALR domains.....	14
The ER.....	15
Ca <sup>2+</sup> in the ER.....	16
Protein folding.....	16
PDIA3.....	18
ER stress and the unfolded protein response.....	18
CALR outside the ER.....	20
Calr knockout and cardiac-specific overexpression of Calr in mice.....	22
CALR in disease.....	23
CALR mutants identified in SUD.....	24
Hypothesis.....	26
Objective.....	26

### CHAPTER TWO: Materials and Methods

Cell culture.....	28
Immunoblotting and immunofluorescence microscopy.....	28
Ca <sup>2+</sup> flux measurements.....	29
Protein expression, purification, and proteolytic digestion.....	30
Aggregation assay.....	30
Q-PCR analysis.....	31
Plasmid DNA and site-specific mutagenesis, expression and protein purification.....	31
Microscale thermophoresis.....	34

Surface plasmon resonance analysis .....	34
Statistical analysis .....	34
Equipment and software.....	35
<b>CHAPTER THREE: Characterization of CALR Mutants</b>	
Identification of specific mutants along CALR structure .....	37
Cellular localization of CALR .....	39
Ca <sup>2+</sup> binding .....	45
Binding to oxidoreductase PDIA3 .....	51
Folding of CALR .....	57
<b>CHAPTER FOUR: Characterization of CALR<sup>376fs</sup> Mutant</b>	
CALR <sup>WT</sup> vs CALR <sup>376fs</sup> .....	66
Cellular localization of the CALR <sup>376fs</sup> mutant.....	69
Ca <sup>2+</sup> binding affinity to the CALR <sup>376fs</sup> mutant .....	71
Binding of oxidoreductase PDIA3 to the CALR <sup>376fs</sup> mutant.....	72
Ca <sup>2+</sup> dynamics of the CALR <sup>376fs</sup> mutant.....	73
Processing of the CALR <sup>376fs</sup> mutant.....	75
Induction of UPR in cells expressing the CALR <sup>376fs</sup> mutant.....	77
Protein folding of the CALR <sup>376fs</sup> mutant.....	78
Chaperone function of the CALR <sup>376fs</sup> mutant.....	80
<b>CHAPTER FIVE: Discussion</b>	
CALR mutants localize to the ER.....	83
CALR mutants bind Ca <sup>2+</sup> and PDIA3.....	83
Cells expressing CALR <sup>376fs</sup> have normal Ca <sup>2+</sup> dynamics.....	84
The CALR <sup>376fs</sup> mutant does not induce higher UPR in cells.....	85
The CALR <sup>376fs</sup> mutant protein is highly degraded.....	86
CALR mutants have a folding defect.....	86
The CALR <sup>376fs</sup> mutant does not efficiently suppress aggregation of MDH .....	88
Increased activity of L-type Ca <sup>2+</sup> channel in presence of the CALR <sup>376fs</sup> mutant .....	89
Future studies .....	89
Conclusions.....	90
<b>REFERENCES .....</b>	<b>94</b>

## List of Tables

Table 1. CALR mutants identified in cases of SUDY and SIDS.....	9
Table 2. List of primers used for mutagenesis.....	32
Table 3. Summary of Ca <sup>2+</sup> binding affinities of CALR mutants.....	50
Table 4. Summary of PDIA3 binding affinities of CALR mutants.....	56
Table 5. Comparison of basic properties of CALR <sup>WT</sup> and CALR <sup>376fs</sup> .....	69
Table 6. Available data of the CALR mutants.....	90

## List of Figures

Figure 1. Nucleotide change identified and pedigree of CALR <sup>376fs</sup> case of SUDY .....	11
Figure 2. Linear representation of CALR domains and three-dimensional model .....	13
Figure 3. Mapping of CALR mutants to CALR specialized regions .....	38
Figure 4. Cellular localization of CALR <sup>WT</sup> and CALR mutants .....	40
Figure 4 a. Cellular localization of CALR <sup>P245L</sup> .....	40
Figure 4 b. Cellular localization of CALR <sup>D295N</sup> .....	41
Figure 4 c. Cellular localization of CALR <sup>D302N</sup> .....	41
Figure 4 d. Cellular localization of CALR <sup>S323C</sup> .....	42
Figure 4 e. Cellular localization of CALR <sup>K355 del</sup> .....	42
Figure 4 f. Cellular localization of CALR <sup>E380G</sup> .....	43
Figure 4 g. Cellular localization of CALR <sup>E381A</sup> .....	43
Figure 5. Cellular localization of CALR <sup>D302N</sup> and CALR <sup>S323C</sup> and overlap with ER marker .....	44
Figure 6. Ca <sup>2+</sup> binding affinity of CALR <sup>WT</sup> and CALR mutants .....	46
Figure 6 a. Ca <sup>2+</sup> binding affinity of CALR <sup>P245L</sup> .....	47
Figure 6 b. Ca <sup>2+</sup> binding affinity of CALR <sup>D295N</sup> .....	47
Figure 6 c. Ca <sup>2+</sup> binding affinity of CALR <sup>D302N</sup> .....	48
Figure 6 d. Ca <sup>2+</sup> binding affinity of CALR <sup>S323C</sup> .....	48
Figure 6 e. Ca <sup>2+</sup> binding affinity of CALR <sup>355K del</sup> .....	49
Figure 6 f. Ca <sup>2+</sup> binding affinity of CALR <sup>E380G</sup> .....	49
Figure 6 g. Ca <sup>2+</sup> binding affinity of CALR <sup>E381A</sup> .....	50
Figure 7. Surface plasmon resonance (BIAcore) analysis of the interaction of PDIA3 with CALR <sup>WT</sup> and CALR mutants .....	52
Figure 7 a. BIAcore analysis of the interaction of PDIA3 with CALR <sup>P245L</sup> .....	53
Figure 7 b. BIAcore analysis of the interaction of PDIA3 with CALR <sup>D295N</sup> .....	53
Figure 7 c. BIAcore analysis of the interaction of PDIA3 with CALR <sup>D302N</sup> .....	54
Figure 7 d. BIAcore analysis of the interaction of PDIA3 with CALR <sup>S323C</sup> .....	54
Figure 7 e. BIAcore analysis of the interaction of PDIA3 with CALR <sup>K355 del</sup> .....	55
Figure 7 f. BIAcore analysis of the interaction of PDIA3 with CALR <sup>E380G</sup> .....	55
Figure 7 g. BIAcore analysis of the interaction of PDIA3 with CALR <sup>E381A</sup> .....	56
Figure 8. Proteolytic digestion of CALR <sup>WT</sup> and CALR <sup>P245L</sup> .....	58
Figure 8 a. Proteolytic digestion of CALR <sup>D295N</sup> .....	59

Figure 8 b. Proteolytic digestion of CALR <sup>D302N</sup> .....	60
Figure 8 c. Proteolytic digestion of CALR <sup>S323C</sup> .....	61
Figure 8 d. Proteolytic digestion of CALR <sup>K355del</sup> .....	62
Figure 8 e. Proteolytic digestion of CALR <sup>E380G</sup> .....	63
Figure 8 f. Proteolytic digestion of CALR <sup>E381A</sup> .....	64
Figure 9. CALR <sup>WT</sup> and CALR <sup>376fs</sup> structure and sequence comparison .....	67
Figure 10. Cellular localization of CALR <sup>376fs</sup> .....	70
Figure 11. Calcium binding affinity of CALR <sup>376fs</sup> .....	71
Figure 12. Surface plasmon resonance (BIAcore) analysis of the interaction between PDIA3 and CALR <sup>376fs</sup> .....	72
Figure 13. Ca <sup>2+</sup> dynamics of cells expressing the CALR <sup>376fs</sup> mutant .....	74
Figure 14. Processing of CALR <sup>376fs</sup> in cells .....	76
Figure 15. Induction of UPR by CALR <sup>376fs</sup> .....	77
Figure 16. Proteolytic digestion of the CALR <sup>376fs</sup> mutant .....	79
Figure 17. Aggregation suppression activity of CALR <sup>376fs</sup> on malate dehydrogenase (MDH).....	81

## List of Abbreviations Used

A.U.	arbitrary units
ANOVA	one-way analysis of variance
ATF6	activating transcription factor 6
ATP	adenosine triphosphate
AV	atrioventricular
BiP	immunoglobulin binding protein
BrS	Brugada syndrome
C1q	complement component 1q
Ca <sup>2+</sup>	calcium ion
CALM	calmodulin
CALR	calreticulin
CALR <sup>376fs</sup>	calreticulin 376 frame-shift
CALR <sup>WT</sup>	calreticulin wild-type
CANX	calnexin
CASQ	calsequestrin
Cav1.2	voltage-gated calcium channel 1.2
cDNA	complementary DNA
CFP	cyan fluorescent protein
CPVT	catecholaminergic polymorphism ventricular tachycardia
CTL	cytotoxic T-lymphocytes
Cx40	connexin40
Cx43	connexin43
del	deletion
DMEM	Dulbecco's Modified Eagle's Medium
DNA	deoxyribonucleic acid
ECG	electrocardiogram
EGTA	ethylene glycol tetraacetic acid
eIF2 $\alpha$	eukaryotic translational initiation factor 2 $\alpha$
ER	endoplasmic reticulum
ERAD	ER-associated degradation
ERSE	ER stress element
EV	empty vector

FBS	fetal bovine serum
Fnorm	normalized fluorescence
fs	frame-shift
fura2-AM	fura2-acetoxymethyl ester
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GFP	green fluorescent protein
GRP94	glucose-related protein 94
HA	hemagglutinin
HEK	human embryonic kidney
HRP	horseradish peroxidase
HSP47	heat-shock protein 47
ICD	immunogenic cell death
IP3R	inositol phosphate receptor 3
IRE1 $\alpha$	inositol-requiring enzyme 1
JNK	janus kinase
K <sup>+</sup>	potassium
KCNH2	potassium voltage-gated channel subfamily H member 2
KCNQ1	potassium voltage-gated channel subfamily Q member 1
$K_D$	dissociation constant
kDa	kilo Dalton
LDL	low-density lipid
LOF	loss-of-function
LPS	lipopolysaccharide
LQTS	long QT syndrome
LRP	LDL-receptor related protein
LTCC	L-type calcium channel
MALDI-TOF MS	matrix assisted laser desorption ionization-time of flight mass spectrometry
MDH	malate dehydrogenase
MEF	mouse embryonic fibroblast
MHC I	major histocompatibility complex class I
MoRF	molecular recognition feature
MPN	myeloproliferative neoplasm
mRNA	messenger RNA

MST	microscale thermophoresis
NA	not available
Na <sup>+</sup>	sodium
NFAT	nuclear factor of activated T-cells
NMR	nuclear magnetic resonance
ORAI1	Calcium release-activated calcium channel protein 1
PBS	phosphate buffered saline
PDI	protein disulfide isomerase
PDIA3	protein disulfide isomerase family A member 3
PDT	photodynamic therapy
PERK	double-stranded RNA-activated protein kinase-like ER kinase
pI	isoelectric point
PLC	peptide-loading complex
PLI	probability of LOF intolerance
PS	phosphatidylserine
PTI	photon technology international
PVC	premature ventricular contraction
Q-PCR	quantitative polymerase chain reaction
QTc	corrected QT interval
RFP	red fluorescent protein
RNA	ribonucleic acid
ROI	region of interest
ROS	reactive oxygen species
Ro-SSA	ribonucleoprotein complex Sjögren's-syndrome-related antigen A
RU	response units
RYR2	ryanodine receptor 2
S1P	site 1 protease
S2P	site 2 protease
SAXS	small-angle X-ray scattering
SCD	sudden cardiac death
SCN5A	sodium voltage-gated channel alpha subunit 5
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SERCA	sarco/endoplasmic reticulum Ca <sup>2+</sup> ATPase
SFV	Semliki forest virus

SIDS	sudden infant death syndrome
SLE	systemic lupus erythematosus
SOCE	store-operated calcium entry
SQTS	short QT syndrome
STAT3	signal transducer and activator of transcription 3
STIM1	stromal interaction molecule 1
SUD	sudden unexplained death
SUDY	sudden unexplained death in the young
TG	thapsigargin
TGF $\beta$ 3	tumor growth factor $\beta$ 3
TpoR/MPL	thrombopoietin receptor
TRDN	triadin
TSP-1	thrombospondin-1
UPR	unfolded protein response
V2R	vomer nasal receptor
VSN	vomer nasal sensory neuron
VT	ventricular tachycardia
VUS	variant of unknown significance
WES	whole exome sequencing
WT	wild type
XBP1s	spliced XBP1
YFP	yellow fluorescent protein

# CHAPTER ONE:

## Introduction

---

## **Sudden unexplained death**

Sudden death is the natural and unexpected death that occurs within an hour after the onset of symptoms, or when death occurs unwitnessed within 24 hrs of the deceased being seen alive and in a normal state of health [1-3]. Approximately 85% of all sudden deaths are of cardiac origin, otherwise known as sudden cardiac death (SCD) [4]. SCD is one of the most common causes of death in developed countries, with at least 3 million people dying suddenly each year [5, 6]. In the United States, an estimated 1,000 individuals die of SCD each day, with most of cases occurring in the elderly and often secondary to coronary artery disease [6].

Meanwhile, sudden death is relatively uncommon in the 1-40 years age group, with an estimated incidence rate of 1.3 to 8.5 cases per 100,000 person-years [7]. Structural cardiac abnormalities such as hypertrophic cardiomyopathy, arrhythmogenic right ventricular cardiomyopathy, congenital coronary artery anomalies, or myocarditis can be identified at autopsy in the majority of cases [8]. However, in up to 30% of cases no structural abnormality is evident during autopsy and no cause of death is identified at post-mortem examination [6, 7, 9].

The term sudden unexplained death (SUD) is used in cases where exhaustive post-mortem examination and autopsy fail to determine a conclusive cause of death [2, 10]. For cases of SUD in children, teenagers or young adults (individuals in the 1-40 years age group) sudden unexplained death in the young (SUDY) is the common terminology.

On the other hand, sudden death in infants (individuals <1 year of age) can be attributed to infection, cardiovascular anomalies, child abuse/negligence, accidents, homicide or metabolic/genetic disorders after a thorough medico-legal investigation including autopsy. However, in as many as 70-80% cases of sudden death in infancy no identifiable cause is found,

and the cases are classified as sudden infant death syndrome (SIDS) [11-13]. SIDS has an estimated incidence of 53:100,000 live births in the United States and 40:100,000 in the United Kingdom [14, 15], and it is the leading cause of infant mortality in developed countries [16]. The high incidence of death in infants for no obvious reason and without warning is poorly understood, but it is believed to be due to the complex interaction of multiple factors. The convergence of three overlapping risk factors or ‘triple risk hypothesis’ has been proposed for the pathogenesis of SIDS: a critical developmental period, an exogenous homeostatic stressor, and an underlying vulnerability [17-19].

### **Molecular autopsy**

A genetic evaluation or “molecular autopsy” consisting of sequence analysis of DNA extracted from blood or tissue of the victim can help identify genetic mutations associated with the disease [7]. In the absence of explanations at post-mortem investigation, including lack of cardiac structural abnormalities during autopsy, cardiac arrhythmias are considered a likely cause of sudden unexplained death. Potentially lethal cardiac ion channel diseases, or channelopathies, leave no evidence to be found during a comprehensive autopsy and can trigger fatal arrhythmias [11]. Therefore, a “molecular autopsy” for analysis of cardiac ion channels mutations in the deceased has become an important tool in elucidating causes of SUDY and SIDS [20-25].

In this manner, mutations in four genes encoding ion channels have been identified in the past in cases of SUD [6, 10]:

- *KCNQ1* (K<sup>+</sup> voltage-gated channel subfamily Q member 1).
- *KCNH2* (K<sup>+</sup> voltage-gated channel subfamily H member 2).
- *RYR2* (ryanodine receptor 2).

- *SCN5A* (Na<sup>+</sup> voltage-gated channel alpha subunit 5).

### **Cardiac channelopathies**

Mutations in K<sup>+</sup>, Na<sup>+</sup>, and Ca<sup>2+</sup> channels lead to ionic imbalances and arrhythmia-causing diseases [23, 26]. These cardiac ion channels associated diseases (channelopathies) are identified by abnormalities in a 12-lead electrocardiogram (ECG) [2, 27]. The four types of cardiac channelopathies are:

- Long QT syndrome (LQTS).
- Short QT syndrome (SQTS).
- Brugada syndrome (BrS).
- Catecholaminergic polymorphic ventricular tachycardia (CPVT).

### **Long QT syndrome**

Long QT syndrome (LQTS) is an inherited arrhythmia condition characterized by prolongation of the QT interval in electrocardiogram (ECG) due to prolonged ventricular repolarization. There is an increased risk for ventricular tachycardia (“*torsades de pointes*”), syncope, seizures, and sudden death with a structurally normal heart, which may be the presenting event in 5 to 10% of LQTS cases [6, 28]. Clinical diagnosis is based on a comprehensive evaluation based on personal and family history, and relies on electrocardiographic presence of a prolonged corrected QT interval (QTc)  $\geq 460$  ms, as well as other ECG findings such as *torsades de pointes* or abnormal T wave morphologies [4, 28]. The prevalence of LQTS is 1:2,000 live births [29].

Several different LQTS-susceptibility genes have been identified, indicating the genetic heterogeneity of the disease, and leading to its classification into 16 different subtypes (LQTS1 to

LQTS16) [4]. The major susceptibility genes encode  $\alpha$  subunits of ion channels, and are gain-of-function mutations in *KCNQ1* (LQT1) and *KCNH2* (LQTS2), as well as loss-of-function mutations in *SCN5A* (LQT3), which account for about 75% of cases of LQTS. The other minor genes encode channel regulatory subunits and accessory proteins, and account for about 5% of cases, thus rendering roughly 20% of cases undetermined by genetic evaluation [6].

### **Short QT syndrome**

Short QT syndrome (SQTS) is a rare inherited disease characterized by a short-corrected QT interval (QTc) of  $\leq 360$  ms. Symptoms include syncope and palpitations, however SQTS confers considerable predisposition to atrial and ventricular fibrillation, and cardiac arrest and sudden death are the most common manifestations of the disease. The estimated prevalence is of less than 1:10,000 [30, 31].

Six genes have been linked to SQTS: gain-of-function mutations in *KCNH2*, *KCNQ1*, *KCNJ2* encoding  $K^+$  channels, and loss-of-function mutations in *CACNA1C*, *CACNB2b*, and *CACNA2D1* encoding L-type  $Ca^{2+}$  channel subunits [4]. Even though familial association is found in the majority of cases of SQTS, the yield of genetic screening has been low, varying between 15% and 40%, thus high rates of between 60% to 85% of cases have an undetermined cause at present [2].

### **Brugada syndrome**

Brugada syndrome (BrS) is a hereditary disease characterized by right ventricular conduction delay and  $\geq 2$  mm coved ST-segment elevation in the anterior right precordial leads (V1-V3) in an ECG. One of the main manifestations is syncope, and ventricular arrhythmias typically occur at rest or during sleep in association to high vagal tone and fever. The prevalence

of BrS is estimated at 1:500-2,000 individuals and is highest in Southeast Asia, however these estimates are not based in robust evidence [17].

Pathogenic variants have been identified in 24 genes encoding Na<sup>+</sup>, K<sup>+</sup>, and Ca<sup>2+</sup> channels or associated proteins. However, loss-of-function mutations in *SCN5A* account for approximately 20-25% of cases, while loss-of-function mutations in genes encoding L-type Ca<sup>2+</sup> channel subunits (*CACNA1C*, *CACNB2b*, and *CACNA2D1*) account for approximately 5% of cases. Variants in other susceptibility genes are rare, and therefore as many as 70% of diagnosed cases remain unresolved [2, 27].

### **Catecholaminergic polymorphic ventricular tachycardia**

Catecholaminergic polymorphic ventricular tachycardia (CPVT) is a pathological condition where intense physical exercise or acute emotional stress may trigger ventricular tachycardia/arrhythmia. Typical manifestations include dizziness and syncope; however, arrhythmias may degenerate into rapid polymorphic ventricular tachycardia and fibrillation, leading to sudden death. Diagnosis of individuals with CPVT is difficult since they have normal electrocardiograms at rest. The diagnostic characteristics become evident during exercise ECG, where ventricular arrhythmias begin at 100-120 beats/min. The prevalence is about 1:1,000 although this estimation is based on minimal evidence [4, 17].

The main CPVT-susceptibility gene identified is *RYR2*, encoding the sarcoplasmic reticulum Ca<sup>2+</sup> channel ryanodine receptor 2 (RYR2), which may account for more than 50% of cases. Another important susceptibility gene is *CASQ2* encoding calsequestrin, a ryanodine receptor interacting and Ca<sup>2+</sup>-handling protein. Other minor rare susceptibility genes include *CALM2* and *CALM3* encoding the RYR2-interacting and Ca<sup>2+</sup>-handling protein calmodulin, and

gene *TRDN* encoding the RYR2 and CASQ2-scaffolding protein triadin. Altogether, these susceptibility genes may account for up to 65% of CPVT cases, with as many as 35% remaining unknown [2, 27].

There is a high percentage of undetermined causes for these channelopathies to this date, and it is similarly high for unresolved cases of SUD. Of note, most of the susceptibility genes identified thus far encode ion channels or interacting proteins at the membrane site [2]. More research is required for better understanding of processes related to ion channel genesis, processing, trafficking, and recycling, as well as of intracellular handling of ions in relation to arrhythmias with no cardiac structural defects [32]. This may help elucidate alternative mechanisms contributing to channelopathies, and decrease the prevailing high uncertainty around SUD events.

### **Whole exome sequencing**

More advanced, robust, and increasingly accessible genetic evaluations such as whole exome analysis (WES) by next generation sequencing, allow for examination of mutations in the whole protein-coding regions of the genome [33, 34]. The disadvantage of this is a vast amount of data generated and uncertainty on how to proceed with it [35]. The number of mutations identified could likely be high, so it must be processed and narrowed down to possibly pathogenic mutations. Another problem consists of obtaining false positives, for which the mutations need to be confirmed. Lastly, the contribution of the identified genetic mutations to disease may not be evident at the time, and thus should be classified as variants of unknown significance (VUS) [36].

A VUS requires confirmation in the case index, analysis of family history, observation of phenotype, identification of variant or gene in other case studies, and finally molecular and

functional characterization, in order to be categorized as pathogenic variant. Such careful analysis is carried out to avoid the problematic confusion of a non-pathogenic variant for disease causing, and help properly clarify and validate pathogenic genes and variants as contributors to the disease phenotype [36]. This is of high importance for future reference, as well as for adequate genetic testing and treatment/counseling of patients and relatives.

### **Novel mutants identified and case index**

Post-mortem genetic evaluation performed in victims of SUDY and SIDS by WES analysis identified several mutations in the *CALR* gene (Table 1), elucidating a novel gene and variants that confer susceptibility to the sudden unexplained death phenotype.

**Table 1. CALR mutants identified in cases of SUDY and SIDS**

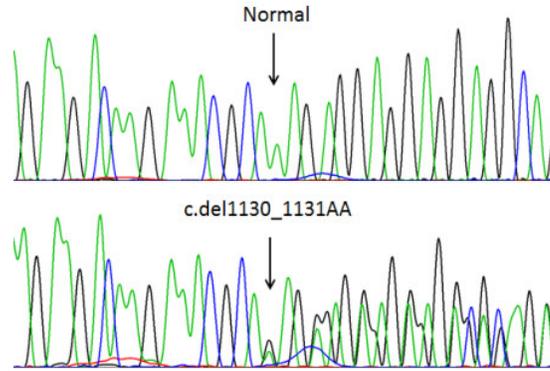
Age	Gender	Ethnicity	Additional Notes	cDNA Nucleotide Change	Mutation
28	Female	Caucasian	Died in sleep.	31 C>T	L11F
3 months	Female	Caucasian	-	170 A>G	Y57C
4 months	Female	Caucasian	-	341 C>G	P114R
1.75 months	Female	Caucasian	-	566 G>C	S189T
2 months	Male	Caucasian	-	734 C>T	P245L
19	Male	Caucasian	Died walking across room.	883 G>A	D295N
18	Male	Caucasian	Died during exercise (basketball).	904 G>A	D302N
2	Female	African American	Died in sleep.	968 C>G	S323C
6	Male	-	Possible LQTS.	del 1062_1064 GAA	355 K Deletion
21	Male	Caucasian	Died in sleep. Family history of abnormal ECG.	del 1130_1131 AA	376fs
20	Male	Caucasian	Died in an unwitnessed collapse. Complained of chest pain.	1139 A>G	E380G
16	Male	Asian	Died in sleep.	1142 A>C	E381A
16	Male	Hispanic	Died in unwitnessed collapse. Family history of syncope.	1142 A>C	E381A
2 months	Female	Caucasian	-	1142 A>C	E381A
3 months	Male	Caucasian	-	1142 A>C	E381A
3 months	Male	Caucasian	-	1142 A>C	E381A
2 months	Female	Caucasian	-	1142 A>C	E381A
8 months	Male	Caucasian	-	1142 A>C	E381A

6 months	Male	Caucasian	-	1142 A>C	E381A
2 months	Female	Caucasian	-	1142 A>C	E381A
20	Male	Caucasian	Drowning.	del 1191_1199 TGAGAAGAA	398-400 EED Deletion
1 month	Female	African American	-	del 1191_1199 TGAGAAGAA	398-400 EED Deletion
3 days	Female	Caucasian	-	del 1191_1199 TGAGAAGAA	398-400 EED Deletion

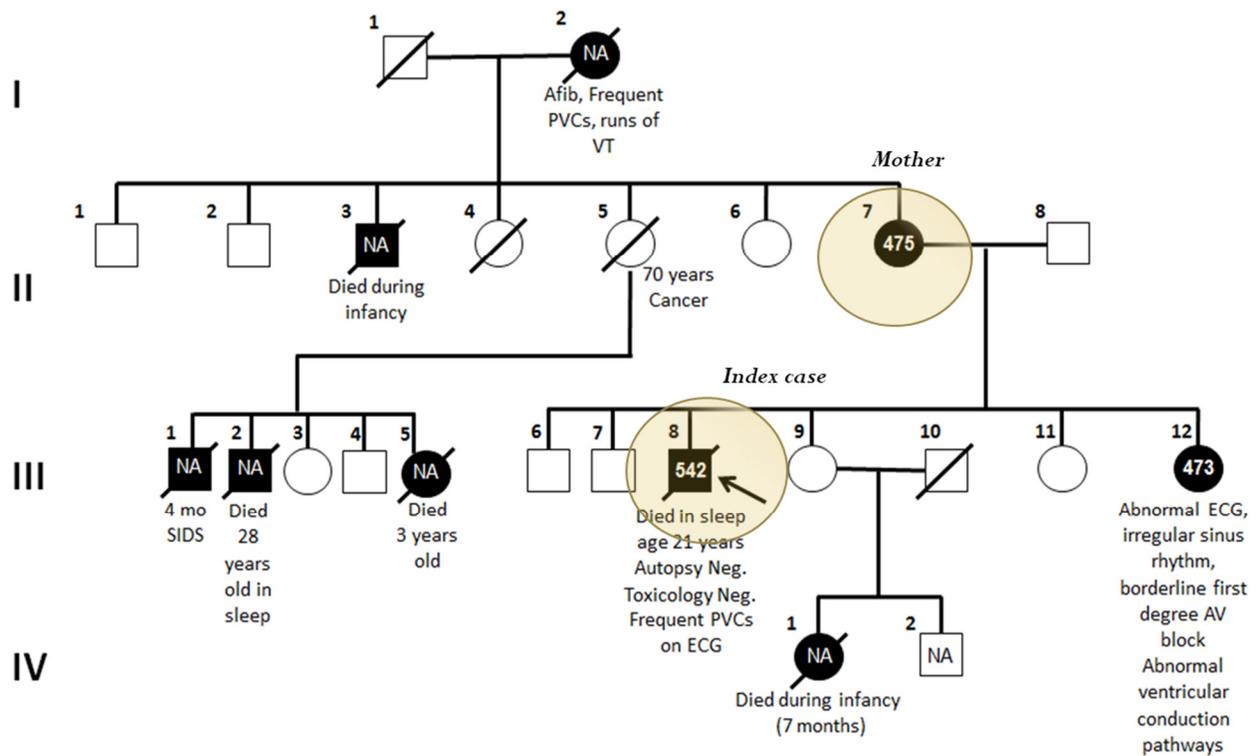
Information about the victims of SUD is presented. The nucleotide changes in the *CALR* gene identified by WES and corresponding amino acid changes in CALR protein are shown (courtesy of Ackerman *et al.*, Mayo Clinic, Rochester, MN, USA).

One of the elucidated mutations in the *CALR* gene, initially found by WES in a case of SUDY, narrowed down by strategic variant filtration analysis and identified as most probable cause of the phenotype, is an ultra-rare (minor allele frequency [MAF] < 0.005% in Allele Frequency Community [AFC, n=13,000][37], Exome Aggregation Consortium [ExAC, n=60,706][38], and the National Heart, Lung and Blood Institute Grand Opportunity Exome Sequencing Project [ESP, n=6,503][39] databases) frame-shift mutation *CALR* c.del1130\_1131AA, encoding a truncated CALR protein at amino acid 376 (CALR<sup>376fs</sup>). The genetic mutation was confirmed in the victim of SUDY (Figure 1A) and in his mother, indicating an autosomal dominant maternal inheritance pattern. Available electrocardiogram (ECG) of the deceased indicated long QT syndrome (LQTS), a cardiac channelopathy leading to arrhythmia, as likely cause of sudden death (QTc= 542 ms). The mother of the deceased also showed prolongation of the QT interval (QTc= 475 ms). Furthermore, there was a history of multiple cases of SIDS and SUDY in the family pedigree (Figure 1B).

A.



B.



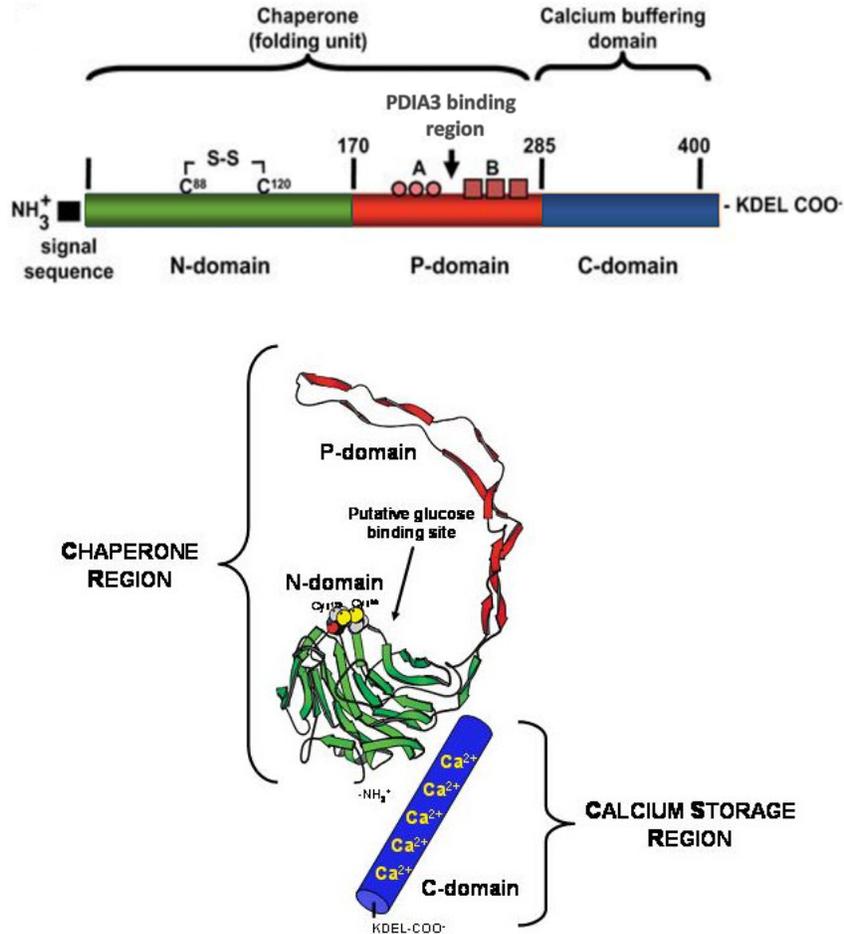
**Figure 1. Nucleotide change identified and pedigree of CALR<sup>376fs</sup> case of SUDY**

A. Sanger sequence confirmation of the mutation in DNA obtained from the deceased. B. The CALR 376 frame-shift (CALR<sup>376fs</sup>) mutation was identified in the index case (arrow, circled) and in the mother of the deceased (circled). Numbers in squares or circles indicate heart-corrected QT interval in milliseconds for family members with electrocardiogram (ECG) available. NA = ECG not available. (Courtesy of Ackerman *et al.*, manuscript in preparation).

*CALR* loss-of-function (LOF) mutations such as frame-shift, nonsense, or splice error, are extremely rare in genome databases, with the *CALR*<sup>376fs</sup> – coding mutation absent in >140,000 individuals. Furthermore, LOF variants in the *CALR* gene are highly intolerable, with a probability of LOF intolerance (PLI) of 0.94 in a scale of 0-1 (Ackerman *et al.* manuscript in preparation). The genes most likely to influence disease are those that are the most intolerant of functional variation in the human population. A high PLI score such as that observed for *CALR*, indicates extreme selective constraint for a gene, where defects might lead to lethal disease, or heterozygous loss-of-function mutations might confer non-trivial survival or reproductive disadvantages [38, 40].

## **CALR**

The human *CALR* gene is localized to chromosome 19, consists of 9 exons and 8 introns, and spans approximately 3.6 kb of genomic DNA [41, 42]. *CALR* encodes a 46-kDa protein, with 417 amino acids in length, containing an N-terminal cleavable amino acid signal sequence and a C-terminal KDEL ER retrieval signal, important for targeting and retaining the protein in the ER lumen [43]. The *CALR* protein consists of three structural and functional domains, which carry out two main activities: Ca<sup>2+</sup> binding and protein folding. The N-domain and P-domain are important for chaperone activities, and the C-domain is highly acidic and binds Ca<sup>2+</sup> with high capacity (Figure 2). The major functions of *CALR* are as an ER Ca<sup>2+</sup> buffer and molecular chaperone, but it has many other functions and binding activities as well.



**Figure 2. Linear representation of CALR domains and three-dimensional model**

The CALR protein consists of an N-terminal signal sequence, N-domain, P-domain, C-domain, and a C-terminal KDEL sequence. Sites of disulfide bond formation and repeats A (IXDPXADXKPEDWDX) and B (GXWXPPXIXNPXYX) are shown. PDIA3 binding region, glucose binding region, chaperone and  $\text{Ca}^{2+}$  buffering regions are indicated. (Modified from [43]).

### **N-domain**

X-ray crystallography studies identified the N-domain as composed of eight antiparallel  $\beta$ -sheets, making a stable globular core resistant to proteolysis in presence of  $\text{Ca}^{2+}$ . The N-domain includes polypeptide and carbohydrate binding sites, a zinc-binding site, and a disulfide linkage site [43, 44].

## **P-domain**

The middle portion of CALR, the P-domain, is proline-rich forming an extended flexible arm. It contains three sets of amino acid repeat A (IXDPXADXXKPEDWDX) and of repeat B (GXWXPPXIXNPXYX), forming a structural backbone that may be involved in carbohydrate binding activities. The P-domain contains a region stabilized by three antiparallel  $\beta$ -sheets that interact with protein disulfide isomerase family A member 3 (PDIA3, also known as ERp57), for synergistic folding of proteins. This region binds  $\text{Ca}^{2+}$  with high affinity ( $K_D = 1 \mu\text{M}$ ) but low capacity (1 mol of  $\text{Ca}^{2+}$  per mol of protein). The P-domain also contains polypeptide and oligosaccharide binding regions, and in conjunction with the N-domain might form a folding unit responsible for chaperone function [43].

## **C-domain**

The C-tail of CALR is highly acidic, containing a large number of negatively charged residues, which bind  $\text{Ca}^{2+}$  with low affinity ( $K_D = 2 \text{ mM}$ ) and high capacity (25 mol of  $\text{Ca}^{2+}$  per mol of protein). It is responsible for the  $\text{Ca}^{2+}$  buffering activity of the protein, which binds about 50% of ER  $\text{Ca}^{2+}$  [45]. The binding of  $\text{Ca}^{2+}$  can affect overall structure, function, and binding activities of the protein, thus the C-domain playing a regulatory role [46-50].

## **Additional functions of CALR domains**

Besides the main activities in  $\text{Ca}^{2+}$  binding and chaperone function, other functions have been identified for the domains of CALR [41]. The N-domain has been shown to interact with the DNA-binding domain of the glucocorticoid receptor, with rubella virus RNA,  $\alpha$ -integrin, protein disulfide isomerase (PDI), and has been shown to inhibit proliferation of endothelial cells and to suppress angiogenesis, affecting tumor growth and metastasis. The P-domain has been shown to

interact with PDI and perforin, a component of cytotoxic T-cell granules. The C-domain has been shown to bind to blood-clotting factors and to inhibit injury-induced restenosis.

The main functions of CALR as  $\text{Ca}^{2+}$  buffer and molecular chaperone are well documented in the ER, however many of the functions outlined here and several others have been attributed to CALR at cellular locations other than the ER [51].

## **The ER**

The ER is a dynamic cellular organelle that forms an extensive network of tubules and cisternae, representing the largest membrane system of animal cells [52]. It is a key component of intracellular  $\text{Ca}^{2+}$  signaling as main site of  $\text{Ca}^{2+}$  storage, and plays a crucial role in the synthesis, correct folding, and processing of membrane and secreted proteins. The ER is also the site of synthesis of phospholipids, cholesterol, and steroids, among other functions [41].

The ER membrane is not a homogenous and steady machinery for protein and lipid synthesis, but rather the flexible structure formed by tubules and cisternae allows for contact and communication with other cellular compartments, such as nucleus, cytosol, mitochondria, Golgi apparatus, and plasma membrane [53-56]. The ER is a spatially and functionally heterogeneous system that forms compartmentalized microdomains through non-uniform distribution of resident proteins and ions, allowing for discrete and temporal signaling [57]. It is therefore a multifunctional organelle able to detect and integrate incoming signals, modulate its own luminal dynamics, and generate output in response to environmental changes [43]. In this way, the ER affects many aspects of cell physiology [41].

## **Ca<sup>2+</sup> in the ER**

Ca<sup>2+</sup> is a fundamental signaling molecule, important in regulation of a variety of critical processes, ranging from cell “birth” (oocyte fertilization), to muscle contraction, gene transcription, and apoptosis [58]. Intracellular Ca<sup>2+</sup> homeostasis is maintained mainly by the ER, which has the highest Ca<sup>2+</sup> concentration and its total is estimated to be about 1-3 mM [59]. Estimates of free ER Ca<sup>2+</sup> range between 60-400 μM, however this concentration lowers to 1-50 μM upon agonist-dependent store depletion [60]. Most ER Ca<sup>2+</sup> is buffered by abundant luminal Ca<sup>2+</sup>-binding proteins, with CALR binding about 50% of it.

Loss of ER Ca<sup>2+</sup> requires refilling of the organelle store, which is mediated by a process known as store-operated Ca<sup>2+</sup> entry (SOCE) [61]. The ER transmembrane protein stromal interaction molecule 1 (STIM1) senses a state of depletion of ER Ca<sup>2+</sup> levels, leading to its mobilization to the subplasmalemmal punctae and coming into contact with the plasma membrane protein Ca<sup>2+</sup> release-activated Ca<sup>2+</sup> channel protein 1 (ORAI1), which induces Ca<sup>2+</sup> entry from the extracellular space [55, 61]. Ca<sup>2+</sup> enters the ER through sarco/endoplasmic reticulum Ca<sup>2+</sup> ATPase (SERCA) and is released by inositol phosphate receptor 3 (IP3R) [62]. SERCA and IP3R therefore regulate Ca<sup>2+</sup> dynamic fluxes in and out of the ER. CALR as a luminal Ca<sup>2+</sup> buffer and modulator of ER Ca<sup>2+</sup> content [45] has been shown to impact IP3R and SERCA activities, suggested to be in part by directly binding to SERCA [63].

## **Protein folding**

Approximately one-third of total cellular proteins are synthesized in the ER [64], which must adopt a correct three-dimensional conformation to perform their biological function. This is supported by specialized proteins in the ER such as folding enzymes and molecular chaperones,

which retain and facilitate the correct folding, maturation, assembly and disassembly of proteins before their entry in the secretory pathway [65]. Misfolded proteins that do not attain their proper conformation are sent for ER-associated degradation (ERAD) [56].

The ER contains a set of  $\text{Ca}^{2+}$ -binding molecular chaperones for the large amount of proteins being synthesized and exported. Some of these chaperones are immunoglobulin binding protein (BiP), glucose-regulated protein 94 (GRP94), the oxidoreductase PDI-like family of proteins, heat-shock protein 47 (HSP47), calnexin (CANX), and CALR [56, 66]. Molecular chaperones are critical for development, and deletion of genes encoding CALR, BiP, GRP94, and PDIA3 are embryonic lethal [67-70].

CALR and CANX are lectin-like chaperones involved in a key quality control process known as the CALR/CANX cycle, which retains monoglucosylated proteins for their proper folding and assembly [41, 43, 65]. Sequences of glucose trimming and addition to proteins (by glucosidases and the enzyme UDP-glucose glycoprotein glucosyl transferase [UGGT], respectively) regulate the monoglucosylated proteins that enter the CALR/CANX cycle. CALR and CANX are also required components that stabilize the peptide-loading complex (PLC) of major histocompatibility complex class I (MHC I) molecules. Both proteins bind  $\text{Ca}^{2+}$  and show high structural similarity and sequence homology, particularly of the P-arm, however CANX has a shorter transmembrane C-tail which has a cytosolic portion. While CANX is an ER membrane molecular chaperone, CALR is soluble and works from within the ER lumen.

CALR has lectin-like chaperone activity by binding to monoglucosylated intermediates, and also traditional chaperone activity by binding exposed hydrophobic peptide portions [71]. CALR has been shown to be a specific chaperone to certain proteins such as influenza

hemagglutinin (HA), Semliki forest virus (SFV) [72], vomeronasal type 2 pheromone receptor (V2R) [73], tyrosinase [74], and bradykinin receptor [45] for example, and to have a delay effect in the trafficking of peptides through retention, which is suggested to help stabilize proteins for folding and to suppress formation of non-productive aggregates. Indeed, in Calr-deficient cells protein folding is accelerated and quality control is compromised, leading to accumulation of unfolded proteins [75].

The N-and-P-domains of CALR are critical for chaperone activity, however such activity is partially  $\text{Ca}^{2+}$ -dependent through effects on the flexible C-tail conformation [47, 76], which might affect the structure of the protein and accessibility of peptide-binding portions.

### **PDIA3**

One important component of the CALR/CANX cycle is the oxidoreductase PDIA3, which binds and works in synergy with each of these chaperones individually for proper folding of substrates, by catalysing disulfide bond formation and isomerization [77].

PDIA3 is also a required component of the PLC of MHC I molecules, it has been shown to affect signal transducer and activator of transcription 3 (STAT3) signaling [70], and to modulate SERCA [78] and STIM1 [79] function, with subsequent effects on  $\text{Ca}^{2+}$  fluxes.

### **ER stress and the unfolded protein response**

As a large multifunctional organelle that is in dynamic communication with other cellular compartments, ER homeostasis can be affected by a wide range of cellular disturbances, such as viral infection, environmental toxins, oxidative stress, reactive oxygen species (ROS) production, heat, changes in pH, drugs, inflammatory cytokines, lipotoxicity, metabolic starvation,  $\text{Ca}^{2+}$

depletion, impaired protein quality control system, and protein mutations [80]. Disruption of ER homeostasis compromises protein folding efficiency, leading to accumulation of misfolded, unfolded, or aggregated proteins, which is termed ER stress. Activation of ER stress leads to initiation of adaptive coping responses such as the unfolded protein response (UPR) [64]. The UPR is a fundamental signaling process that functions to restore ER homeostasis and protein folding capacity, mainly through stimulation of transcriptional activation of ER chaperones genes such as *CALR*, attenuation of general protein translation, and degradation of misfolded proteins. If ER stress remains unresolved, continuous UPR activation induces cell death by apoptosis [81].

The UPR is a complex signal transduction pathway that transmits information about the protein folding status of the ER to nucleus and cytosol to restore ER homeostasis. ER stress activates three ER transmembrane proteins to initiate the UPR: double-stranded RNA-activated protein kinase-like ER kinase (PERK), activating transcription factor 6 (ATF6), and inositol-requiring enzyme 1 alpha (IRE1 $\alpha$ ). Normally, these transmembrane proteins are bound and inactivated by the molecular chaperone BiP. However, accumulation of misfolded proteins in the ER lumen engages BiP, leading to dissociation from PERK, ATF6 and IRE1 $\alpha$ , and mobilization of UPR signaling [64, 80, 81].

Upon dissociation from BiP, PERK phosphorylates eukaryotic translational initiation factor 2 $\alpha$  (eIF2 $\alpha$ ), which stops translation of proteins. ATF6 translocates to the Golgi complex, where it is cleaved by site 1 protease (S1P) and site 2 protease (S2P), releasing a DNA-binding domain that functions as transcription factor of ER stress elements (ERSE) that upregulate expression of molecular chaperones and ERAD proteins. IRE1 $\alpha$  through endoribonuclease activity splices mRNA encoding transcription factor XBP1. The spliced XBP1 (XBP1s) when translated relocates to the nucleus to induce activation of genes encoding proteins involved in folding,

transport, and degradation of proteins. Apoptosis can be initiated during prolonged UPR in part by preferential translation of transcription factor ATF4 after PERK phosphorylation of eIF2 $\alpha$ , and through activation of Janus kinase (JNK) pathways by IRE1 $\alpha$  endonuclease activity.

The ER is an important intracellular organelle that allows cells to adjust to a wide variety of conditions through UPR activation. However, abnormal and sustained ER stress and UPR are involved in many pathologies including neurodegeneration, diverse forms of cancer, diabetes, pro-inflammatory conditions, and cardiovascular disease [81]. There are therefore adaptive and maladaptive components of UPR activation in physiology and disease [80].

### **CALR outside the ER**

Many functions, properties, and effects have been attributed to CALR at other cellular locations than the ER [51]. For example, CALR in the extracellular matrix is associated with the healing process of cutaneous wounds. CALR supports wound healing in part by upregulating tumor growth factor  $\beta$ 3 (TGF $\beta$ 3) which increases proliferation and migration of keratinocytes and fibroblasts, and possibly by interaction with components of the extracellular matrix such as collagen, laminins, and metalloproteinases required for remodeling [82, 83].

Cell surface CALR also causes apoptotic and cancer cells to undergo immunogenic cell death (ICD) by phagocytosis, suggested to be in part due to binding to pro-phagocytic signal phosphatidylserine (PS), complement component 1q (C1q), and low-density lipid (LDL) receptor-related protein (LRP) on phagocytes. Furthermore, CALR affects adaptive immunity by interaction with thrombospondin-1 (TSP-1) mediating focal adhesion of T-cells, and by interaction with perforin in secretory granules of cytotoxic T-lymphocytes (CTLs) for recognition of target cells and lytic activity [51, 82].

CALR has been implicated in activities in cytoplasm and nucleus, by binding to  $\alpha$ -integrin, rubella virus RNA, and glucocorticoid receptor for example, with proposed regulation of cell adhesion, translation and gene expression [41, 83].

It is unclear how CALR can mobilize out of the ER to the cell surface or the cytosol. Cell-surface CALR seems to be induced by various stressors that trigger the unfolded protein response (UPR) such as hypoxia, glucose fluctuation, viral infections, and changes in ER  $\text{Ca}^{2+}$  homeostasis [51, 83]. UPR activation increases CALR expression and abundance, which is suggested to allow escape from ER retention and entry to the secretory pathway, ending with exocytosis in a SNARE-dependent pathway. Anthracycline antibiotics, and photodynamic therapy (PDT) have also been shown to induce cell-surface CALR [82]. Meanwhile, cytosolic CALR has been proposed to be due to retro-translocation from the ER, with modulation of a C-terminal signal sequence and post-translational modifications such as glycosylation and arginylation involved in the process [83].

However, the retro-translocation of CALR to the cytosol has not been documented, and translocation to the cell surface is a matter of debate [82]. The presence of CALR on the cell surface could be partially due to a number of reasons from the experimental procedure, such as cell necrosis and release of CALR to neighboring cells, CALR antibody specificity issues, imprecision in technique used like flow cytometry, addition of extracellular CALR rather than intracellular translocation, quality of the recombinant CALR used, or contamination with lipopolysaccharide (LPS) which can affect observed outcomes such as immune response [82].

A better understanding of mechanisms involved in CALR localization outside of the ER is required for clarification. It has been proposed that CALR could modulate activities such as cell adhesion or gene expression by having effects on signaling networks from within the ER lumen

[41, 84]. This modulation could result from CALR's role as a  $\text{Ca}^{2+}$  buffer, molecular chaperone, and component of the unfolded protein response [43, 82]. Such broad effects of CALR exemplifies the central role that the ER can play in diverse cellular functions.

Regardless of precise mechanism, the vast variety of binding partners and effects of CALR in events such as wound healing, cell adhesion, angiogenesis, gene expression, cancer, and ICD, highlights the multifunctional nature of the protein and identifies it as an attractive therapeutic target [51, 82, 85].

### **Calr knockout and cardiac-specific overexpression of Calr in mice**

The Calr-deficient mouse, created by homologous gene recombination technique, displayed embryonic lethality due to defects in cardiac development [67]. Histological analysis showed major lesions in the heart and marked decrease in ventricular wall thickness. Expression of activated calcineurin in Calr-deficient mice led to activation of nuclear factor of activated T-cells (NFAT) and rescued embryonic lethality [86], indicating that Calr involvement in  $\text{Ca}^{2+}$  signaling is important for embryonic development.

Meanwhile, Calr overexpression in adult mice hearts causes bradychardia, complete heart block, and sudden death [87, 88]. Pathological changes in function of the sinus and AV nodes were observed, with Calr suggested to play a role in the cardiac conductive system. Histological analysis showed dilated atria and ventricular chambers, thinner ventricular walls, and disarray of cardiomyocytes. The cardiomyocytes displayed decreased current density of L-type  $\text{Ca}^{2+}$  channels (LTCC) and significantly reduced expression of gap junction proteins connexin43 and connexin40 (Cx43 and Cx40), with suggestion of a possible Calr role in the folding, assembly and/or traffic of these channel and gap junction proteins [87].

Calr is a ubiquitous protein that is highly expressed in the developing heart, but its expression is downregulated in late embryonic stages and abundance in the adult heart is low [67, 89]. In adult cardiomyocytes, calsequestrin (Casq2) is the main  $\text{Ca}^{2+}$ -buffering protein in the sarcoplasmic reticulum, important for demanding cardiac muscle contraction requirements [90]. The structural and spatial separation of sarcoplasmic reticulum and ER compartments in cardiomyocytes is not well understood [66, 91], but while the sarcoplasmic reticulum might be important for the contractile machinery through  $\text{Ca}^{2+}$  storage and release, the ER might be required for housekeeping functions, including regulation of protein quality control and other aspects of  $\text{Ca}^{2+}$  signaling, where Calr is relevant [88].

The studies in Calr-deficient mice embryos and overexpression in adult mice hearts indicate that defects in Calr expression lead to severe cardiac pathology, and thus regulation of Calr activity is critical for proper heart function.

### **CALR in disease**

High abundance of intracellular CALR has been found in different types of cancer, being positively correlated with tumor size in breast cancer and with metastasis in gastric, pancreatic, prostate, and ovarian cancers, for example. Curiously, higher CALR in the ER is associated with cancer progression, while CALR at the cell surface is detrimental for cancer cells [82, 85].

CALR also plays a role in the pathophysiology of autoimmune diseases, as it has been identified as an antigen in sera from patients suffering from systemic lupus erythematosus (SLE), coeliac disease, polychondritis, and rheumatic disease. CALR has been found to associate with ribonucleoprotein complex Sjögren's-syndrome-related antigen A (Ro-SSA), an auto-antigen in patients with SLE disorders. CALR may be an auto-antigen or may influence the pathology of

autoimmune disease by association with other complexes, such as Ro-SSA or components of innate and adaptive immunity as outlined previously [41, 51, 82, 83].

Current research on CALR has been accelerated due to the discovery of somatic *CALR* mutations in humans with myeloproliferative neoplasms (MPNs) by two independent groups [92, 93]. The mutations consist primarily of a 52 base-pair insertion or a 5 base-pair deletion, leading to frame-shifts in exon 9 and a truncated CALR protein. These mutations in CALR account for as many as 40% of cases of certain subtypes of MPNs, known as essential thrombocythemia and primary myelofibrosis, cancers associated with increased proliferation of megakaryocytes. One of the proposed mechanisms is that CALR mutants can induce increased activity of the thrombopoietin receptor (TpoR/MPL) through direct interaction [94-96], with a subsequent JAK/STAT signaling characteristic of MPNs. However, research is still ongoing on this matter.

### **CALR mutants identified in SUD**

CALR is a multifunctional protein that has been primarily identified in the ER, but with activities suggested at other cellular locations. CALR deficiency is embryonic lethal in mice due to impaired heart development, and cardiac-specific overexpression in adult mice leads to heart defects. CALR has been implicated in many cellular processes and linked to diseases such as cancer, autoimmunity, and with specificity to MPNs in recent years. The discovery of *CALR* mutations in victims of SUD adds additional weight to the relevance of CALR to human physiology. The identified CALR mutants provide an exciting opportunity to explore the involvement of the protein in proper cellular function, with the observed contribution to the pathology of disease when defective: arrhythmia and the SUD phenotype.

In order to do so, first the main known functions of CALR must be systematically studied and validated in the mutants, with focus on the truncated CALR<sup>376fs</sup>, by analysis of Ca<sup>2+</sup> and PDIA3 binding, cellular Ca<sup>2+</sup> homeostasis, cellular localization and processing, induction of unfolded protein response (UPR), folding of the protein, and chaperone function.

## **Hypothesis**

The CALR mutants have defects in structure and function that contribute to arrhythmia and SUD.

## **Objective**

To identify such defects by performing protein characterization through cellular, biochemical, and biophysical studies of the most relevant CALR activities.

# CHAPTER TWO:

## Materials and Methods

---

## Cell culture

Wild-type and *Calr*<sup>-/-</sup> mouse embryonic fibroblasts (MEFs) were previously described [87]. Wild-type (female origin), *Calr*<sup>-/-</sup> (male origin), HeLa (female origin), and HEK293 (female origin) cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Sigma-Aldrich) with 10% fetal bovine serum (FBS) (Sigma-Aldrich) and incubated at 37°C and 5% CO<sub>2</sub>. Wild-type and *Calr*<sup>-/-</sup> MEFs were transfected with YFP-CALR<sup>WT</sup> or YFP-CALR<sup>376fs</sup> expression vectors using the Neon transfection system (Invitrogen) as per the manufacturer's protocol. HeLa cells were transfected with YFP-CALR<sup>WT</sup>, or YFP-CALR<sup>376fs</sup>, or with each of the other YFP-CALR mutants expression vectors by using Turbofect transfection reagent (Thermo Fisher Scientific, R0531). HEK293 cells were transfected with YFP-CALR<sup>376fs</sup>, or CALR<sup>D302N</sup>-GFP, or CALR<sup>S323C</sup>-GFP expression vectors using Turbofect transfection reagent.

HEK293 cells were treated with tunicamycin (10 µg/ml), an inhibitor of protein glycosylation (Sigma-Aldrich, T7765), Brefeldin A (5 µg/ml), an ER-to-Golgi transport inhibitor (Sigma-Aldrich, B7651), or 1 µM MG132, proteosomal inhibitor (Sigma-Aldrich, C2211) for 16 hrs. For induction of UPR cells were incubated for 5 hrs with 0.5 µM thapsigargin, an ER Ca<sup>2+</sup> uptake inhibitor (Sigma-Aldrich, T9033).

## Immunoblotting and immunofluorescence microscopy

Wild-type, *Calr*<sup>-/-</sup> and HEK293 cells were harvested 24-48 hours post-transfection into RIPA lysis buffer containing 50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EGTA, 1 mM EDTA, 1% Triton X-100, 0.1% SDS, and 0.5% sodium deoxycholate. Lysates were prepared in sample buffer containing 1% β-mercaptoethanol, 10% glycerol, 2% SDS, 65 mM Tris pH 6.8, and 0.005 mg/ml bromophenol blue and separated on SDS-PAGE (10% acrylamide), followed by immunoblotting.

Immunoreactive protein bands were detected using peroxidase-conjugated secondary antibodies, followed by standard enhanced chemiluminescence reaction. Blots were probed with the following antibodies: homemade goat anti-CALR (1:300) and rabbit anti-PDIA3 (1:1000); and with rabbit anti-GAPDH (1:2000 Abcam, ab9485). Secondary antibodies were HRP-conjugated: rabbit anti-goat and goat anti-rabbit (1:5000, Jackson ImmunoResearch). Immunoblot images were scanned and quantified by densitometry using Image Studio Lite Ver 5.0 (Li-Cor).

Immunofluorescence analyses were performed by Dr. Elzbieta Dudek. HeLa cells were transfected with YFP-CALR<sup>WT</sup>, YFP-CALR<sup>376fs</sup>, or all other YFP-CALR mutant expression vectors, while HEK293 cells were transfected with YFP-CALR<sup>376fs</sup> and CFP-ER, CALR<sup>D302N</sup>-GFP and RFP-ER, or CALR<sup>S323C</sup>-GFP and RFP-ER expression vectors. Fluorescent imaging was carried out 24-48 hrs post-transfection on a Leica TCS SP5 confocal microscope using a 60x oil immersion lens. Overlap of YFP-CALR<sup>376fs</sup> and CFP-ER signal, or GFP-tagged mutants and RFP-ER marker were analyzed using ImageJ (<https://imagej.nih.gov/ij/download.html>). A straight line was drawn across each cell and identified as a region of interest (ROI). The signal intensity of each channel (green for YFP-CALR<sup>376fs</sup>, blue for CFP-ER, green for CALR<sup>D302N</sup>-GFP and CALR<sup>S323C</sup>-GFP, and red for RFP-ER) was obtained using the corresponding ROI and the values were plotted along the same X axis coordinates to identify regions of overlap.

### **Ca<sup>2+</sup> flux measurements**

Measurement of cytoplasmic Ca<sup>2+</sup> concentrations were carried out using a Ca<sup>2+</sup>-free buffer, composed of 143 mM NaCl, 6 mM KCl, 1 mM MgSO<sub>4</sub>, 20 mM Hepes pH 7.4, 0.1% glucose, and 1 mM EGTA. Ca<sup>2+</sup> measurements on wild-type and *Calr*<sup>-/-</sup> cells in suspension were carried out using Fura2-acetoxymethyl ester (Fura2-AM, F1201) [87]. Measurements were carried out using

a Photon Technology International instrument (PTI) QuantaMaster spectrofluorometer system with excitation wavelengths of 340 nm and 380 nm and emission wavelength of 510 nm [45].  $\text{Ca}^{2+}$  release from ER stores was induced by treatment with 2  $\mu\text{M}$  thapsigargin, and store-operated  $\text{Ca}^{2+}$  entry was induced with 4 mM  $\text{CaCl}_2$ .  $\text{Ca}^{2+}$  concentrations were calculated via two calibration steps: (1) a maximum value obtained by addition of 10 mM  $\text{CaCl}_2$  and 10  $\mu\text{M}$  ionomycin and (2) a minimum value obtained by addition of 33 mM EGTA, 25 mM Tris, and 0.42% Triton X-100.

### **Protein expression, purification, and proteolytic digestion**

Proteins were purified at the Department of Eukaryote Gene Engineering, Institute of Biotechnology at Vilnius University, Lithuania, as specified under the “plasmid DNA and site-specific mutagenesis, expression and protein purification” section. Ten  $\mu\text{g}$  of purified  $\text{CALR}^{\text{WT}}$  or  $\text{CALR}^{376\text{fs}}$  or all other CALR mutants were incubated at 37 °C with trypsin at 1:100 (trypsin/protein; w/w), with a final trypsin concentration of 10  $\mu\text{g}/\text{ml}$ . Aliquots were taken at the indicated time points and the reaction was stopped by addition of sample buffer containing 1%  $\beta$ -mercaptoethanol, 10% glycerol, 2% SDS, 65 mM Tris pH 6.8, and 0.005 mg/ml bromophenol blue. The proteins were separated in SDS-PAGE (12% acrylamide) and stained with Coomassie blue.

### **Aggregation assay**

One  $\mu\text{M}$  malate dehydrogenase (MDH, Sigma-Aldrich, M2634) was mixed with 0.2  $\mu\text{M}$   $\text{CALR}^{\text{WT}}$ , with 0.2  $\mu\text{M}$   $\text{CALR}^{376\text{fs}}$  mutant, or with 0.1  $\mu\text{M}$   $\text{CALR}^{\text{WT}}$  plus 0.1  $\mu\text{M}$   $\text{CALR}^{376\text{fs}}$  mutant. Samples were incubated at 45 °C in 50 mM sodium phosphate, pH 7.5, in a quartz cuvette at total volume of 1.2 ml. Light scattering was monitored for 60 minutes, using a PTI QuantaMaster

spectrofluorometer system equipped with a temperature-controlled cell holder. Excitation and emission wavelengths were set to 320 and 360 nm, respectively [97-99].

### **Q-PCR analysis**

Total RNA was isolated from cells with the RNeasy Mini Kit (QIAGEN), and first strand cDNA synthesis was performed with iSCRIPT Reverse Transcription Supermix for Q-PCR (BIO-RAD). For the real-time Q-PCR reaction, PerfeCTa SYBR Green FastMix (Quanta Biosciences) was used, and quantification was performed on Rotor-Gene Q (QIAGEN). The final quantitation of amount of target (Ct value) was converted to the amount of transcript and normalized by glyceraldehyde 3-phosphate dehydrogenase (GAPDH). The following primers were used: XBP1s Forward: 5'-CCGCAGCAGGTGCAGG-3'; XBP1s Reverse: 5'-GAGTCAATACCGCCAGAATCCA-3'; GAPDH Forward: 5'-AGGGCTGCTTTTAACTCTGGT-3'; GAPDH Reverse: 5'-CCCCACTTGATTTTGGAGGGA-3'.

### **Plasmid DNA and site-specific mutagenesis, expression and protein purification**

Plasmid DNA and mutagenesis experiments were previously done by Dr. Elzbieta Dudek. cDNA encoding human CALR<sup>WT</sup> was inserted in a pEYFP vector containing an N-terminal signal sequence (SS-pEYFP) to generate YFP- CALR<sup>WT</sup>. CALR<sup>WT</sup> was cut out from the vector with restriction enzymes EcoRI and XbaI, to use as template for generation of the CALR<sup>376fs</sup> mutant. CALR<sup>376fs</sup> was generated by PCR-driven amplification of the CALR<sup>WT</sup> cDNA template by using a forward primer 5'- CTGAGGAGTTTGGCAACGAGACG-3' and reverse primer 5'- TCTAGATCCTCCTTGTCTCTGCCTCC-3'. The PCR product was ligated into a SS-pEYFP plasmid for generation of YFP- CALR<sup>376fs</sup>.

Other CALR mutants were generated using QuickChange site-directed mutagenesis kit.

The primers used for this are summarized in Table 2.

**Table 2. List of primers used for mutagenesis**

<b>Mutant</b>	<b>Forward primer</b>	<b>Reverse primer</b>
<b>CALR<sup>P245L</sup></b>	5'- GAGCATATCCCTGACCTTGATGCTAAGAAGCCCG -3'	5'- CGGGCTTCTTAGCATCAAGGTCAGGGATATGCTC-3'
<b>CALR<sup>D295N</sup></b>	5'- GATCCACCCAGAAATTAACAACCCCGAGTATTCT CCC-3'	5'- GGGAGAATACTCGGGGTTGTTAATTTCTGGGTGGAT C-3'
<b>CALR<sup>D302N</sup></b>	5'- CCCCGAGTATTCTCCAATCCCAGTATCTATGCC - 3'	5'- GGCATAGATACTGGGATTGGGAGAATACTCGGGG - 3'
<b>CALR<sup>S323C</sup></b>	5'- CCTCTGGCAGGTCAAGTGTGGCACCATCTTTG- 3'	5'-CAAAGATGGTGCCACACTTGACCTGCCAGAGG - 3'
<b>CALR<sup>K355 del</sup></b>	5'-CAAAGGCAGCAGAACAAATGAAGGAC -3'	GTCCTTCATTGTCTGCTGCCTTTG
<b>CALR<sup>E380G</sup></b>	5'- CAAGAAACGCAAAGAGGAGGGGGAGGCAGAGG -3'	5'-CCTCTGCCTCCCCCTCCTCTTTGCGTTTCTTG -3'
<b>CALR<sup>E381A</sup></b>	5'- GAAACGCAAAGAGGAGGAGGCGGCAGAGGACA AG -3'	5'-CTTGCCTCTGCCGCTCCTCCTCTTTGCGTTTC - 3'

CALR<sup>WT</sup>-GFP, CALR<sup>376fs</sup>-GFP, and the rest of the mutants were generated by ligation of the previously obtained cDNAs into an SS-pEGFP vector using EcoRI and XbaI restriction enzymes. For CALR<sup>WT</sup>-GFP, forward primer 5'-CCGGGTACCATGGCCTCCGACGTGCTA-3' and a reverse primer 5'-CCGGAATTCTTAGAGATCCTCCTGTGC-3' were used. For

CALR<sup>376fs</sup>-GFP and other mutants were the same primers as above. The plasmids were sent to Dr. Evaldas Čiplys at the Department of Eukaryote Gene Engineering, Institute of Biotechnology at Vilnius University, Lithuania, for protein purification.

CALR<sup>WT</sup>-GFP and CALR<sup>376fs</sup>-GFP mutant were purified from *Saccharomyces cerevisiae*, transformed with pFDC- CALR<sup>WT</sup> and pFDC- CALR<sup>376fs</sup> plasmids (forward primer 5'-CCGGGTACCATGGCCTCCGACGTGCTA-3' and a reverse primer 5'-CCGGAATTCTTAGAGATCCTCCTGTGC-3' were used for amplification of CALR<sup>WT</sup> gene, and forward primer 5'-TCTAGAATGCTGCTATCCGTGCCGCTGCTGC-3' and reverse primer 5'-TCTAGATCATCCTCCTTGTCTCTGCCTCC-3' were used for amplification of CALR<sup>376fs</sup> gene, which were restriction-cloned into the pFDC vector using XbaI restriction enzymes), with growth medium as described previously [100]. Culture medium was microfiltered through 1.6 μM, 0.45 μM and 0.2 μM filters using a Pressure Filter Holder and a vacuum pump. After microfiltration, proteins from the culture medium were concentrated and transferred into the binding buffer (20 mM L-histidine, 100 mM NaCl, pH 5.5) through tangential ultrafiltration using cassettes with 100 kDa cut-off membranes and a peristaltic pump. Proteins were loaded onto the column packed with Q Sepharose FastFlow resin equilibrated in the same buffer. The column was washed with binding buffer and bound proteins were eluted with a NaCl concentration gradient (100–500 mM). CALR proteins were eluted in a single peak with approximately 250 mM NaCl. Elution fractions containing purified recombinant protein were pooled and buffer was exchanged to CALR storage buffer (20 mM Tris-HCl, 150 mM NaCl, 3 mM CaCl<sub>2</sub>, pH 7.5) using Sephadex G25 column. Proteins were stored frozen at -70 °C. Human PDIA3 (ERp57) was also purified from *S. cerevisiae*, transformed with pFDC-PDIA3 plasmid, growth medium as described previously [101].

### **Microscale thermophoresis**

MST assays were carried out by Dr. Elzbieta Dudek, as described in [102]. Serial dilutions of  $\text{Ca}^{2+}$  were mixed with 300 nM of CALR protein in PBS buffer (25 mM phosphate pH 7.4, 150 mM NaCl, 0.05% Tween-20) and incubated for 20 minutes. About 10  $\mu\text{l}$  of sample were loaded into standard monolith NT capillaries provided by the manufacturer and assays were performed in a Nanotemper Monolith NT.015T. By plotting the concentrations of unlabeled binding partner with the changes in fluorescent thermophoresis signal,  $K_D$  values were determined using the Nanotemper analysis software. The overlay diagram was made by normalizing the data to fraction bound.

### **Surface plasmon resonance analysis**

Surface plasmon resonance measurements were performed by Dr. Elzbieta Dudek using BIAcore. Sensor chips were coupled with purified PDIA3 in 10 mM sodium acetate buffer, pH 4 at a flow rate of 5 ml/min, followed by blocking with 1 M ethanolamine, pH 8.5, for 7 mins. The binding of aliquots of purified CALR was measured in a buffer containing 20 mM Tris, pH 7.0, 135 mM KCl, 0.05% Tween-20, 200  $\mu\text{M}$  phenylmethylsulfonyl fluoride, 100  $\mu\text{M}$  benzimidazole, and protease inhibitors. Binding was carried out at 20 °C at a flow rate of 30 ml/min.

### **Statistical analysis**

Data are presented as mean  $\pm$  standard error of mean. Statistical analysis was performed using GraphPad Prism with one-way analysis of variance (ANOVA). Statistical significance was accepted at a  $p < 0.05$  value.

## **Equipment and software**

The Neon transfection system was purchased from Invitrogen. Leica TCS SP5 confocal microscope was from Leica Microsystems. The cell culture cabinet class II type A2 (NU-543-600) was from Labgard (Leduc, Canada), and the cell incubator from NUAIRE US autoflow. Centrifuge (5702) was obtained from Eppendorf. Rotor-Gene Q was from QIAGEN. The QuantaMaster spectrofluorometer was from Photon Technology International (PTI). The microscale thermophoresis equipment Monolith NT.015T was purchased from NanoTemper Technologies. Surface plasmon resonance instrument BIAcore T200 system was from GE Healthcare Life Sciences. ImageJ was obtained from (<https://imagej.nih.gov/ij/download.html>). Statistical analysis were performed using GraphPad Prism 6 by GraphPad Software Company (La Jolla, USA)

# CHAPTER THREE:

## Characterization of CALR Mutants

---

## Identification of specific mutants along CALR structure

The classical nomenclature of CALR separates the protein into structural and functional domains (N-domain, P-domain, and C-domain) based on the continuous amino acid sequence. However, continuing research on structure and function of the protein suggest that the domains of CALR can be further viewed as specialized regions: a globular core region composed of N-domain and the N-terminus portion of C-domain, a flexible P-arm, and an acidic C-terminus region (Figure 3A). This is based on the observed folding of the protein by X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and small-angle X-ray scattering (SAXS) studies, where the N-domain forms a globular core, the P-arm extends as a flexible loop, and the beginning of the C-domain integrates into the N-domain forming a  $\beta$ -sheet-based globular region, ending with an  $\alpha$ -helix that leads to the flexible and highly disordered acidic C-tail [44, 50, 103-105].

The length and high intrinsic disorder of the C-tail have precluded determination or modelling of the structure of this region, while the highly flexible P-arm has not been crystallized either, but rather has been separately solved by NMR. Overall, CALR is characterized by a highly dynamic structure containing a globular core decorated with long flexible regions, but such flexible regions (P-arm and C-tail) have prevented structural determination of the CALR protein.

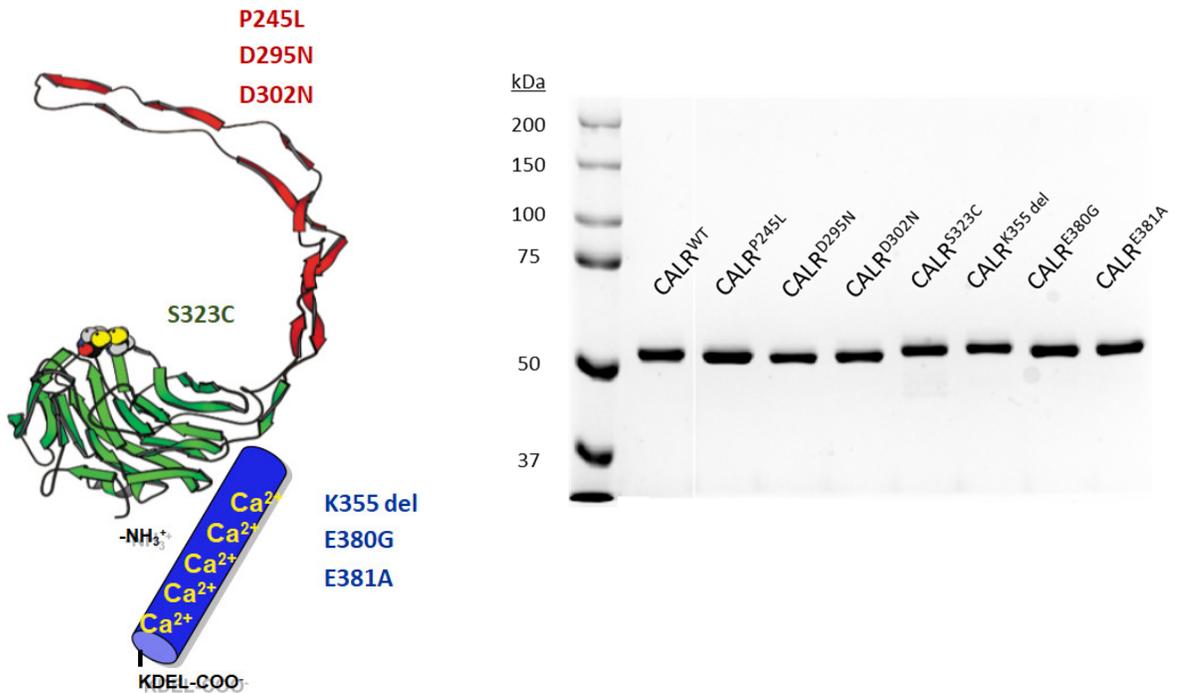
Based on the current understanding of the structure of CALR and its arrangement into separate regions, the mutants identified in SUD can be mapped to either core globular region, P-arm, or C-tail (Figure 3B). Mutants CALR<sup>P245L</sup>, CALR<sup>D295N</sup>, CALR<sup>D302N</sup>, CALR<sup>S323C</sup>, CALR<sup>K355 del</sup>, CALR<sup>E380G</sup>, and CALR<sup>E381A</sup> (the case index CALR<sup>376fs</sup> is presented in a separate chapter) were partially characterized here by analysis of cellular localization, binding to Ca<sup>2+</sup>, binding to PDIA3, and folding of the protein. The purified CALR mutants and wild-type protein (CALR<sup>WT</sup>) used in

this study were run in SDS-PAGE and stained with Coomassie blue, showing a single band with mobility at similar molecular weights of 55-kDa (Figure 3B).

A.



B.



**Figure 3. Mapping of CALR mutants to CALR specialized regions**

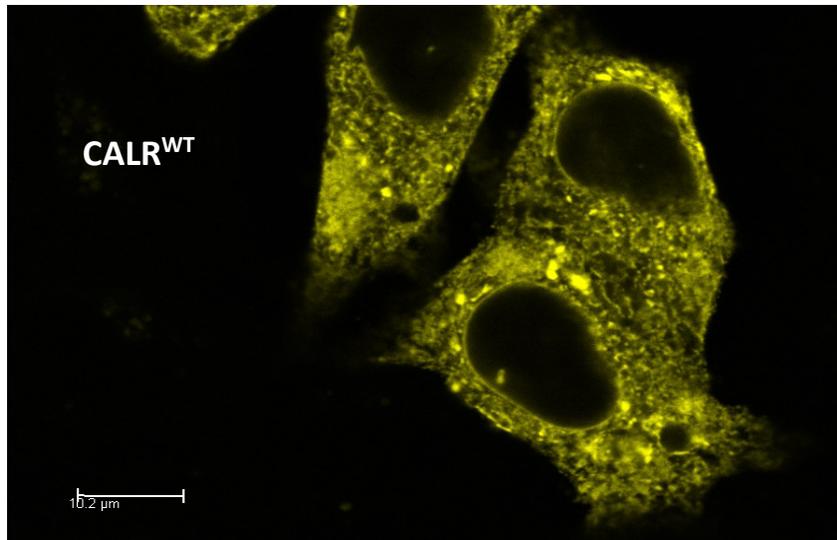
A. Suggested separate regions of CALR formed by its domains, based on structure and folding studies (modified from [50]). B. Location of CALR mutants identified in SUD to such regions in a three-dimensional model according to amino acid number (left). Mobility in SDS-PAGE of purified CALR mutants evaluated by Coomassie blue staining (right).

## Cellular localization of CALR

CALR is an ER lumen resident protein, containing a cleavable amino acid signal sequence for targeting to the ER and a KDEL sequence for ER retrieval. However, CALR has been identified as a multifunctional protein, with activities suggested at other cellular compartments including the cytosol, nucleus, and most importantly at the cell surface [51, 82, 106]. Localization at other cellular compartments might depend upon post-translational modifications of the protein, or certain environmental conditions, such as in stressed or apoptotic cells. To investigate whether the CALR mutants identified in SUD undergo alternative distribution compared to wild-type protein that could affect function, the preferential cellular localization of CALR and mutants was evaluated.

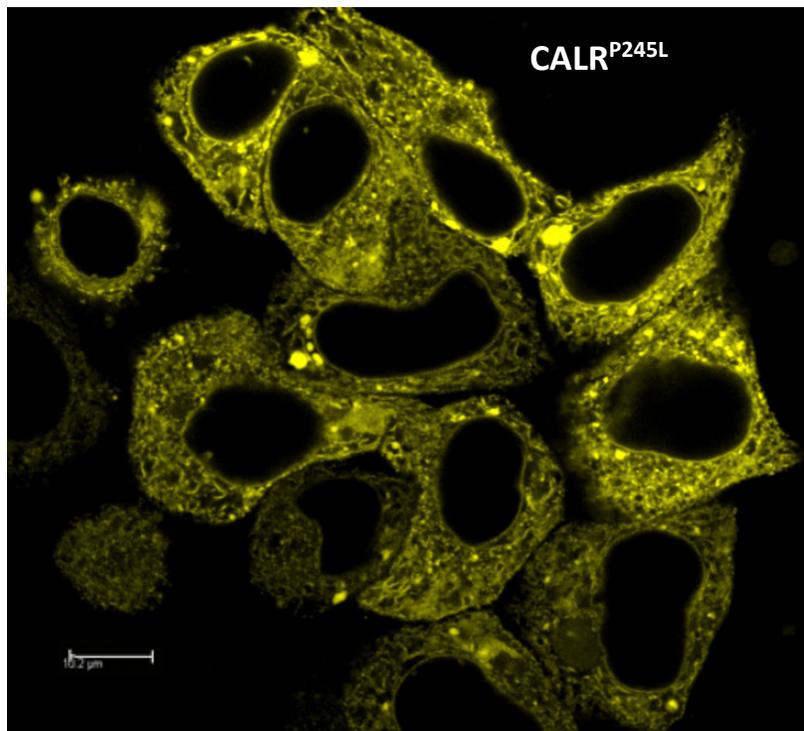
YFP-tagged mutants or wild-type CALR were expressed in HeLa cells and fluorescence analysis was performed by confocal microscopy [Figure 4 (a-g)]. Two of the mutants, CALR<sup>D302N</sup> and CALR<sup>S323C</sup> were also GFP-tagged, expressed in HEK293 cells, and analyzed against an RFP-ER marker by fluorescence overlap comparison (Figure 5).

All the mutants displayed an ER distribution pattern similar to wild-type protein in intact cells. There was no visible lining of the plasma membrane by any of the mutants or wild-type CALR, but evident lining of nucleus contact sites. GFP-tagged CALR<sup>S323C</sup> overlapped with the ER marker, while CALR<sup>D302N</sup> also showed high overlap, however with some few localized clusters of non-overlap.

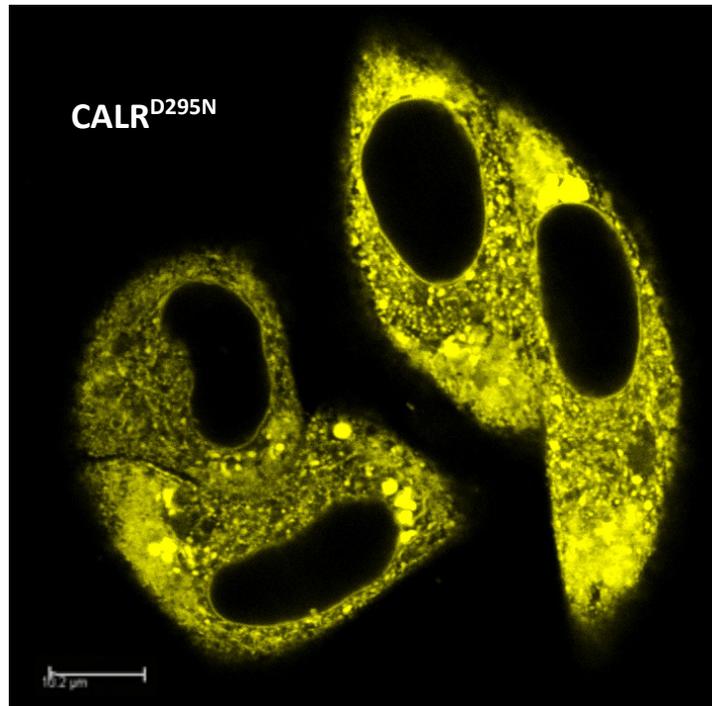


**Figure 4. Cellular localization of CALR<sup>WT</sup> and CALR mutants**

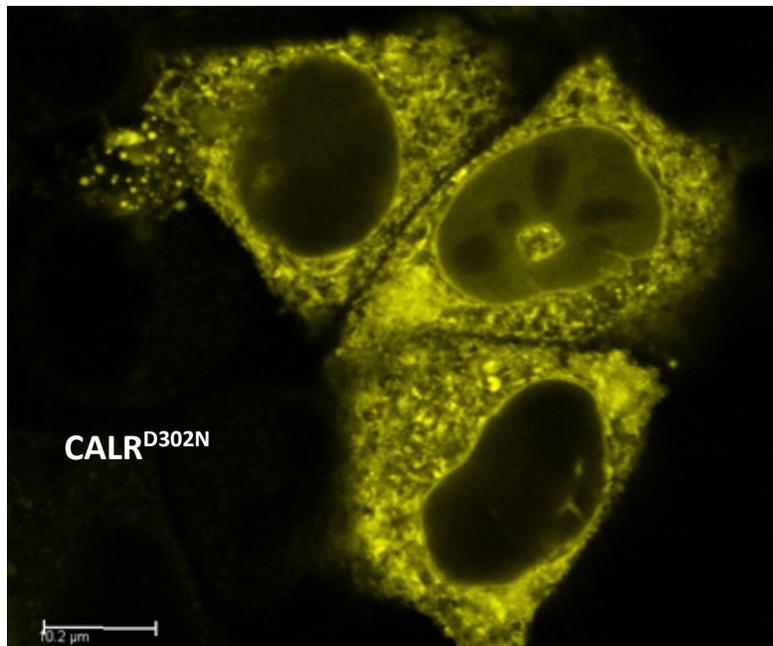
Intracellular distribution of YFP-CALR<sup>WT</sup> and YFP-CALR mutants in HeLa cells visualized by fluorescence microscopy. Scale bar: 10.2 μM. (Performed by Dr. Elzbieta Dudek).



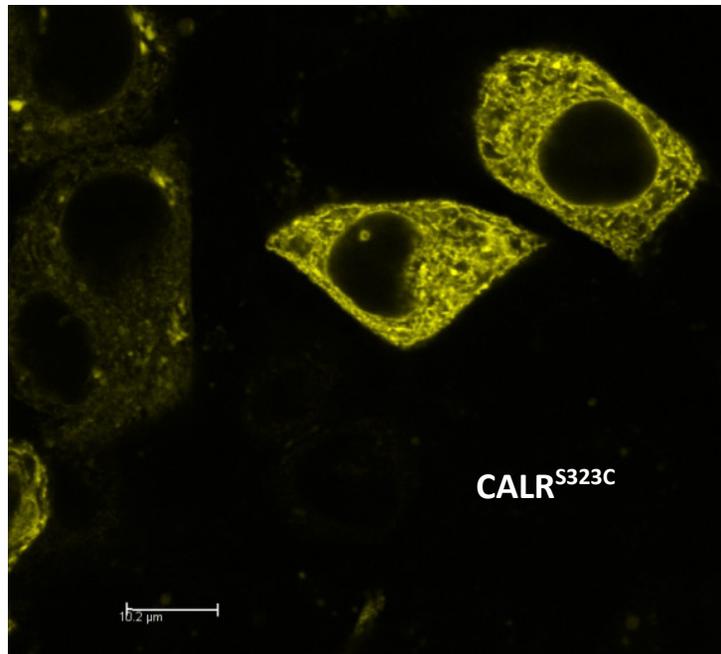
**Figure 4 a. Cellular localization of CALR<sup>P245L</sup>**



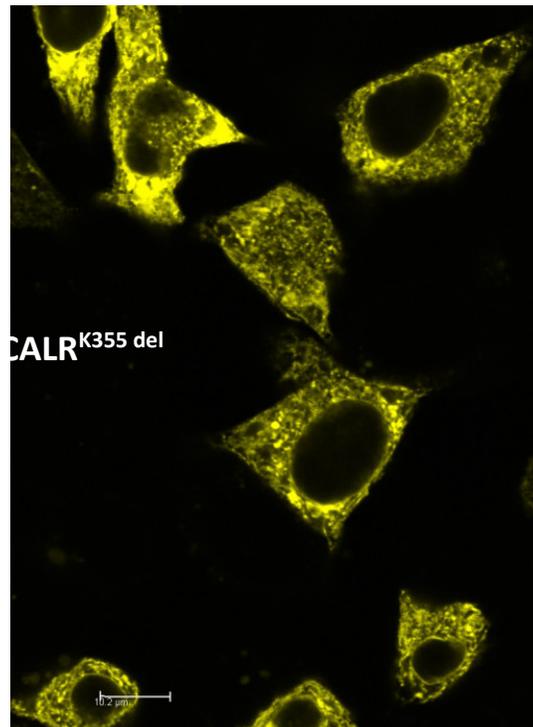
**Figure 4 b. Cellular localization of CALR<sup>D295N</sup>**



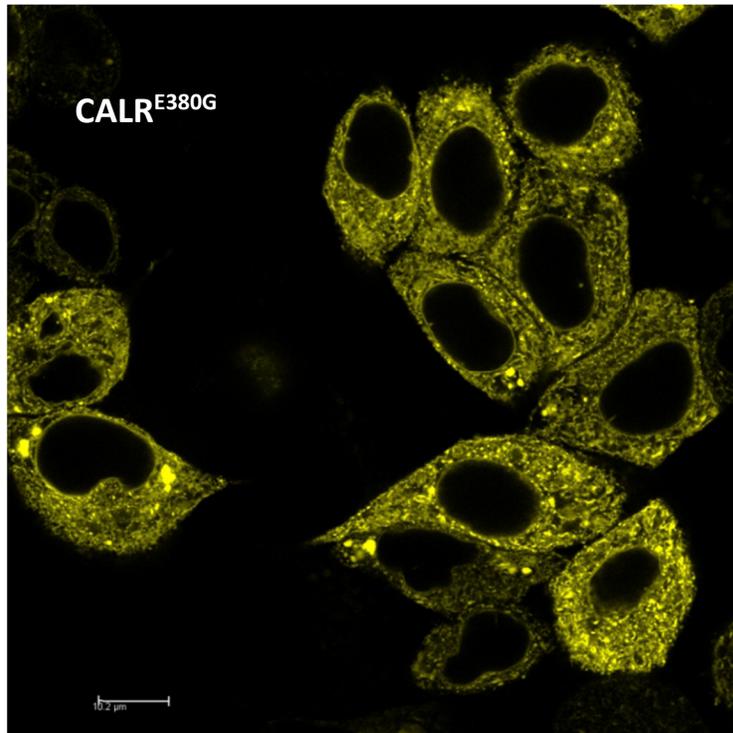
**Figure 4 c. Cellular localization of CALR<sup>D302N</sup>**



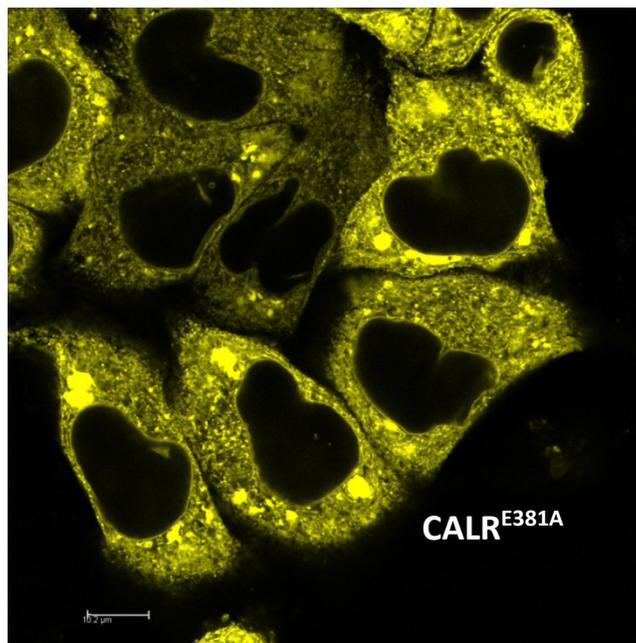
**Figure 4 d. Cellular localization of CALR<sup>S323C</sup>**



**Figure 4 e. Cellular localization of CALR<sup>K355 del</sup>**

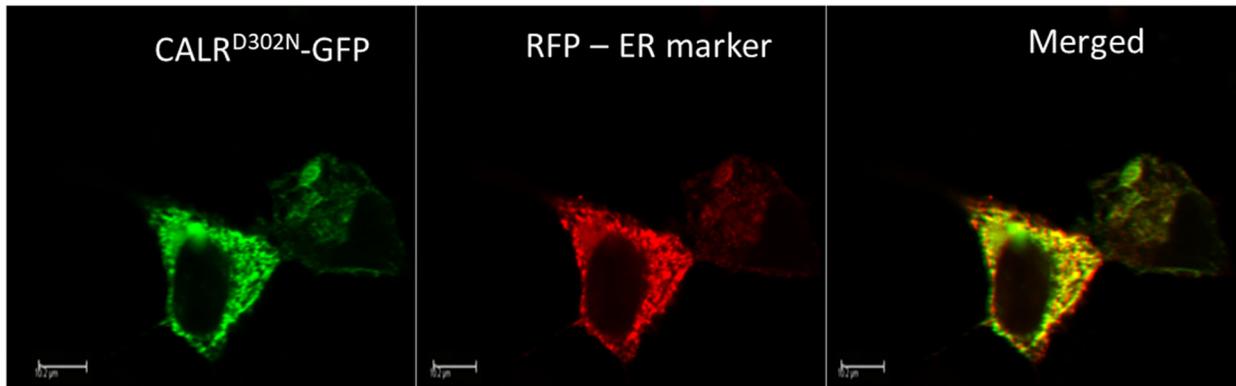


**Figure 4 f. Cellular localization of CALR<sup>E380G</sup>**

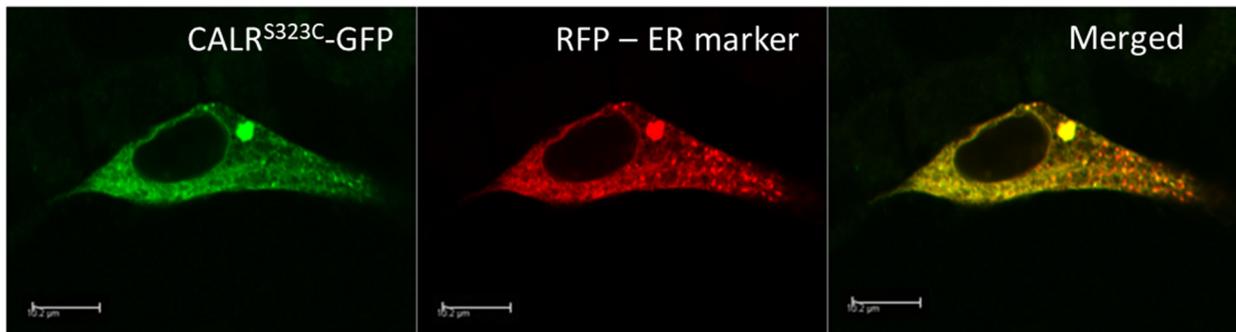


**Figure 4 g. Cellular localization of CALR<sup>E381A</sup>**

*A.*



*B.*



**Figure 5. Cellular localization of CALR<sup>D302N</sup> and CALR<sup>S323C</sup> and overlap with ER marker**

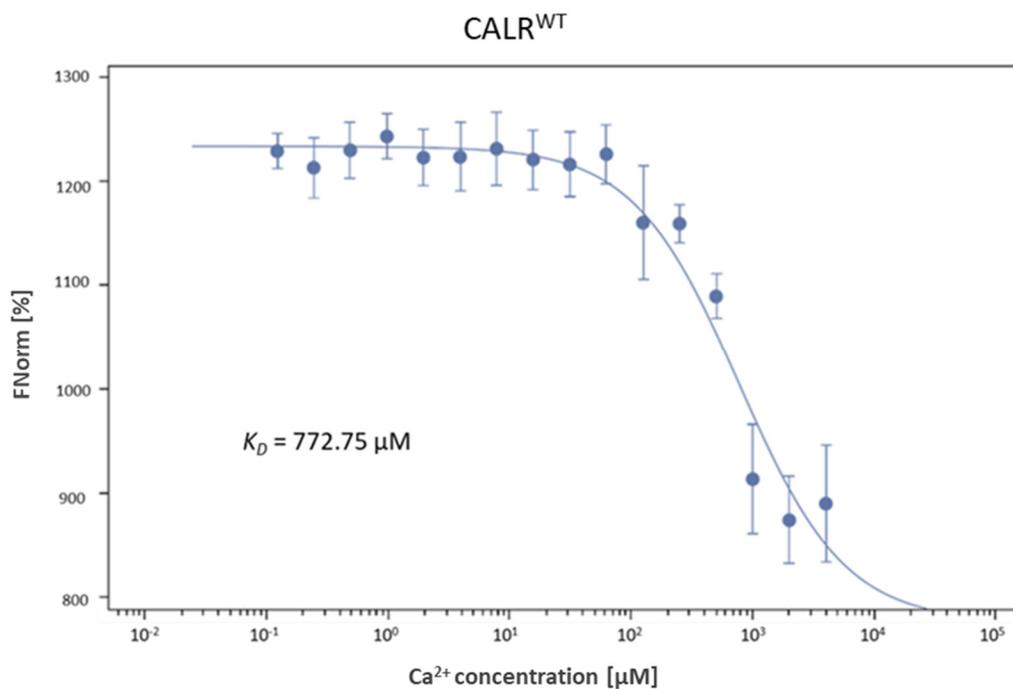
*A.* Overlap analysis of CALR<sup>D302N</sup>-GFP, or CALR<sup>S323C</sup>-GFP with RFP-ER marker in HEK293 cells. Pearson's coefficient = 0.88. *B.* Overlap analysis of CALR<sup>S323C</sup>-GFP with RFP-ER marker in HEK293 cells. Pearson's coefficient = 0.95. Scale bars: 10.2 μm. (Performed by Dr. Elzbieta Dudek).

## Ca<sup>2+</sup> binding

The ER is the main site of intracellular Ca<sup>2+</sup> storage [59], and CALR is the main ER Ca<sup>2+</sup>-binding protein [43, 45], therefore necessary for proper cellular homeostasis of this fundamental signaling molecule [84]. Considering that Ca<sup>2+</sup>, along with K<sup>+</sup> and Na<sup>+</sup>, are the main ions with imbalances involved in arrhythmia and SUD phenotypes [2, 4, 27], it was necessary to examine whether CALR mutants have disrupted ability to bind Ca<sup>2+</sup>.

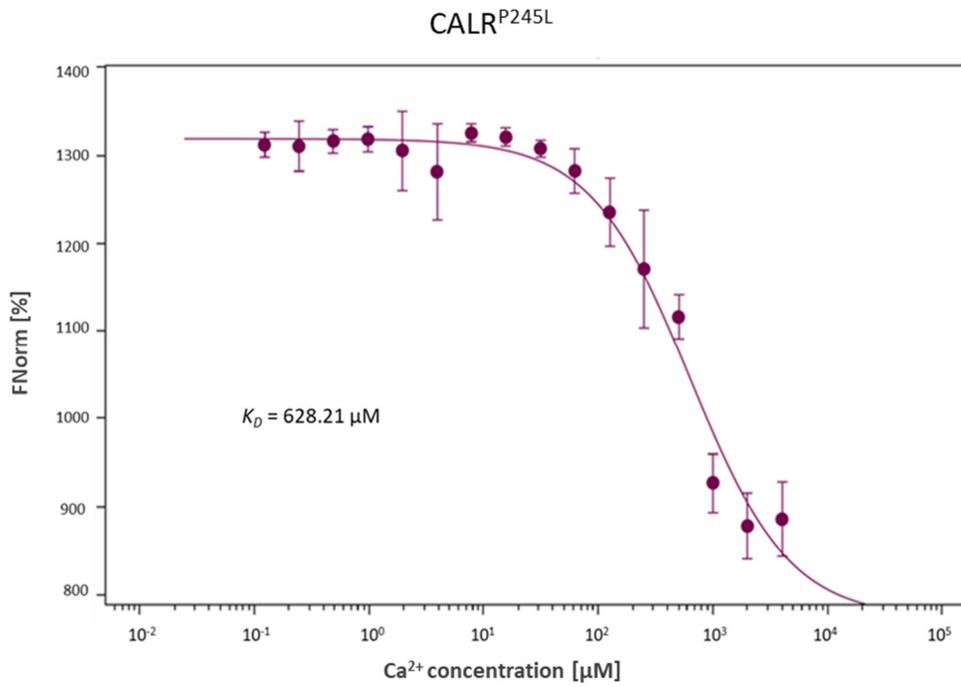
For this purpose, microscale thermophoresis (MST) was used for the *in vitro* analysis of CALR interaction with Ca<sup>2+</sup>. Briefly, purified CALR<sup>WT</sup> or CALR<sup>376fs</sup> were mixed with increasing concentrations of Ca<sup>2+</sup> in small capillaries. A temperature gradient was induced by an infrared laser, and changes in thermophoresis due to protein binding to ligand were detected by the Nanotemper Monolith NT.015T, and binding constants were determined using Nanotemper analysis software.

Binding curves are presented in Figure 6 (a-g), and binding constants are summarized in Table 3. All mutants were able to bind Ca<sup>2+</sup> similar to CALR<sup>WT</sup> ( $K_D = 772.75 \mu\text{M}$ ). However, MST analysis software calculated slight differences in mutants binding constants compared to wild-type protein, with the most marked differences in a relatively higher binding affinity by CALR<sup>D302N</sup> ( $K_D = 414.16 \mu\text{M}$ ) and relatively lower binding affinity by CALR<sup>355K del</sup> ( $K_D = 1576.2 \mu\text{M}$ ) and CALR<sup>S323C</sup> ( $K_D = 2130.56 \mu\text{M}$ ).

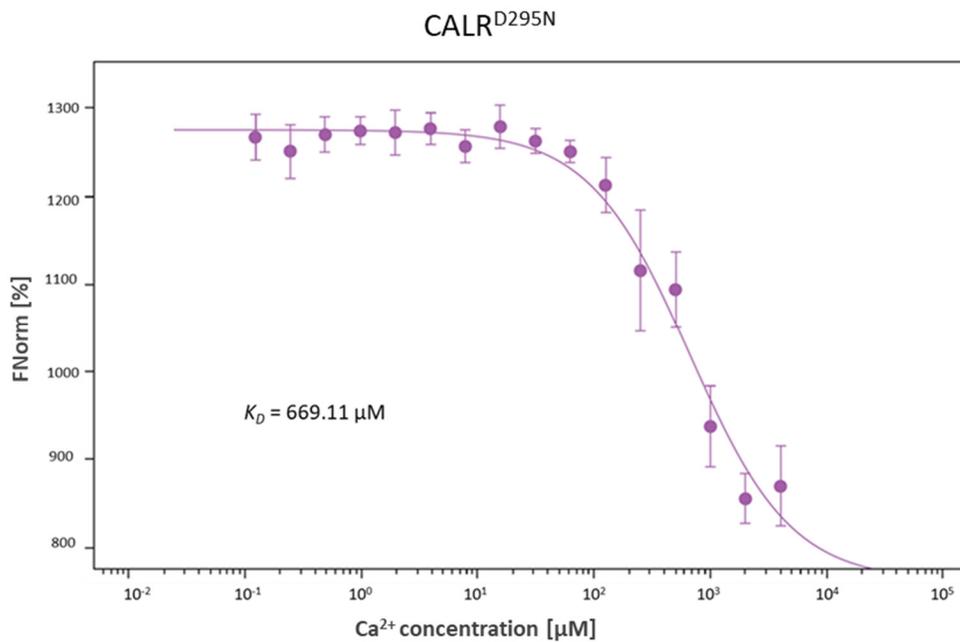


**Figure 6. Ca<sup>2+</sup> binding affinity of CALR<sup>WT</sup> and CALR mutants**

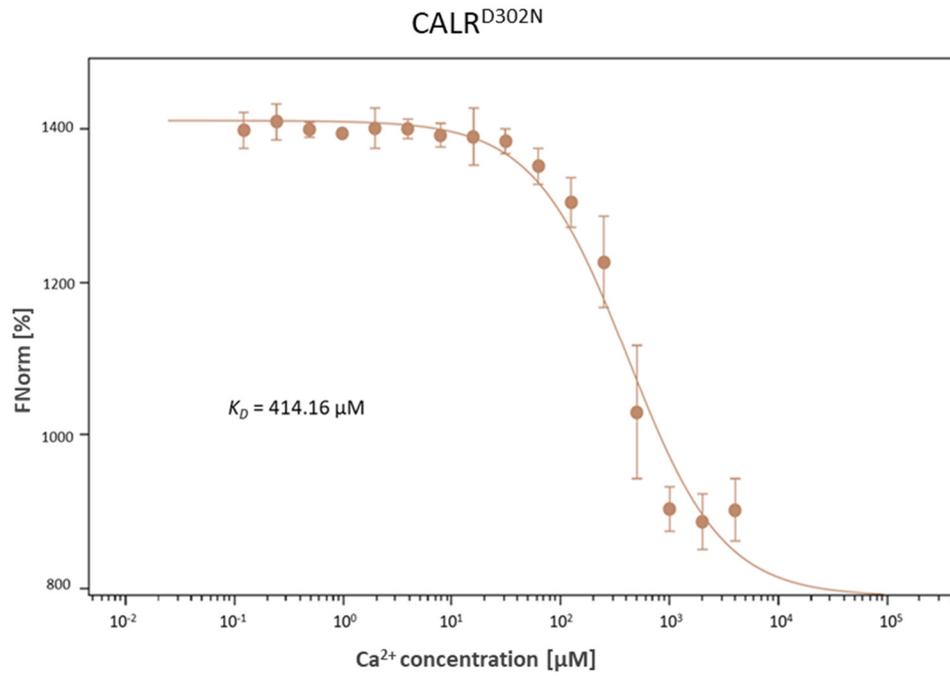
Microscale thermophoresis (MST) of the interaction of Ca<sup>2+</sup> with CALR<sup>WT</sup> and CALR mutants. Curves are derived from the specific change in thermophoretic mobility upon titration of ligand (Ca<sup>2+</sup>) to a constant CALR concentration.  $K_D$  values were calculated using NanoTemper analysis software. Error bars = s.d.; n = 5. (Performed by Dr. Elzbieta Dudek).



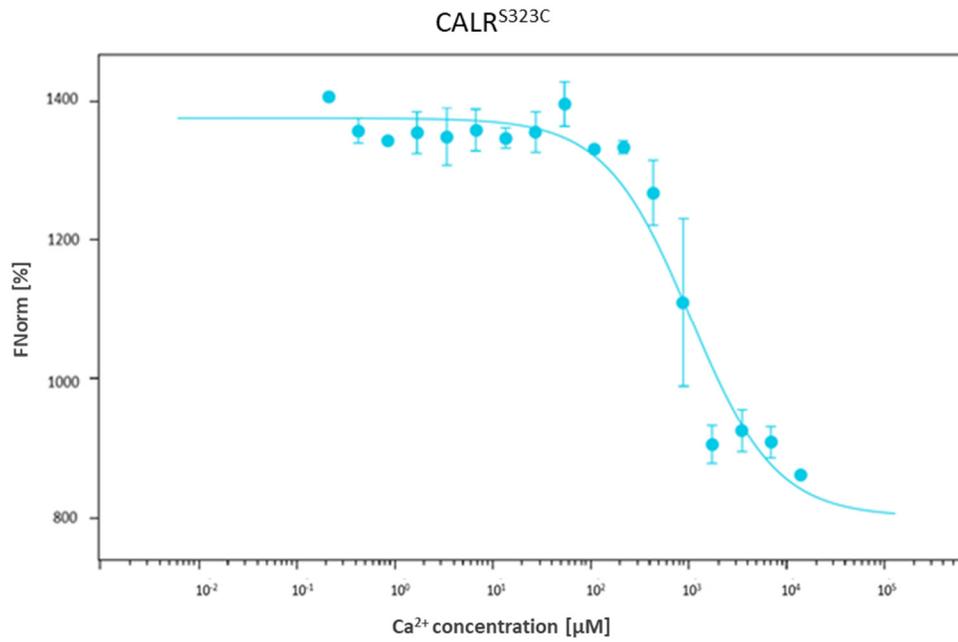
**Figure 6 a. Ca<sup>2+</sup> binding affinity of CALR<sup>P245L</sup>**



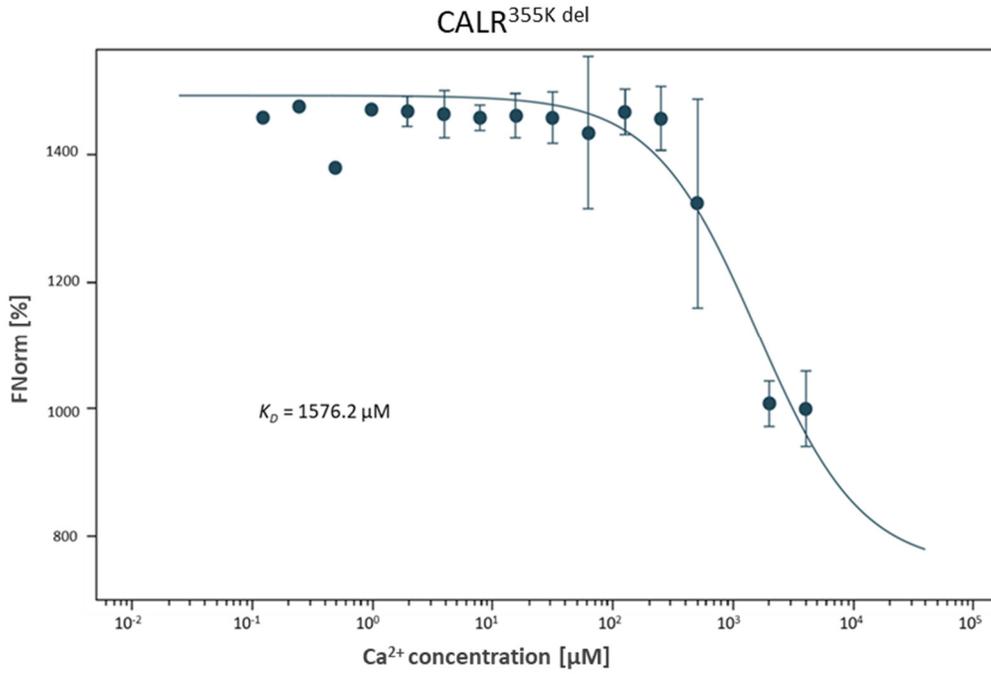
**Figure 6 b. Ca<sup>2+</sup> binding affinity of CALR<sup>D295N</sup>**



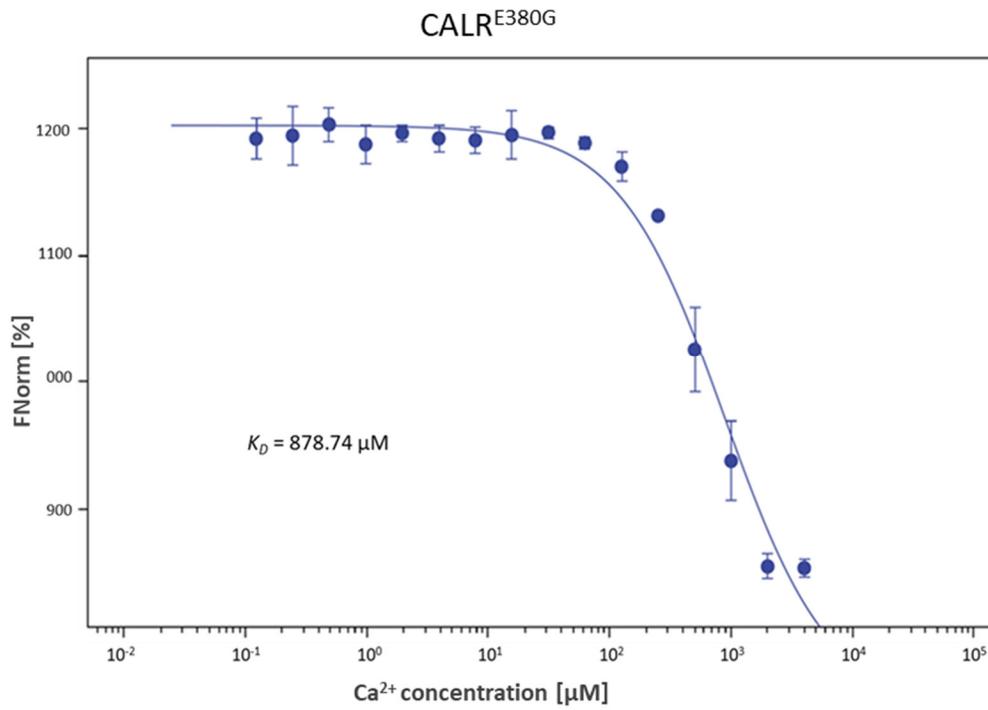
**Figure 6 c. Ca<sup>2+</sup> binding affinity of CALR<sup>D302N</sup>**



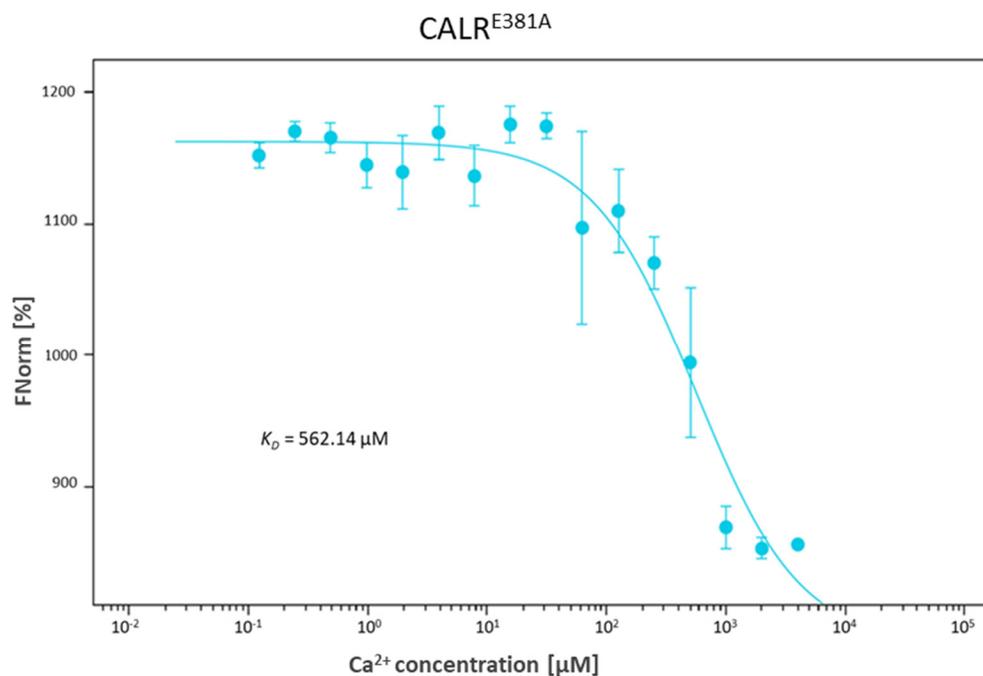
**Figure 6 d. Ca<sup>2+</sup> binding affinity of CALR<sup>S323C</sup>**



**Figure 6 e. Ca<sup>2+</sup> binding affinity of CALR<sup>355K del</sup>**



**Figure 6 f. Ca<sup>2+</sup> binding affinity of CALR<sup>E380G</sup>**



**Figure 6 g. Ca<sup>2+</sup> binding affinity of CALR<sup>E381A</sup>**

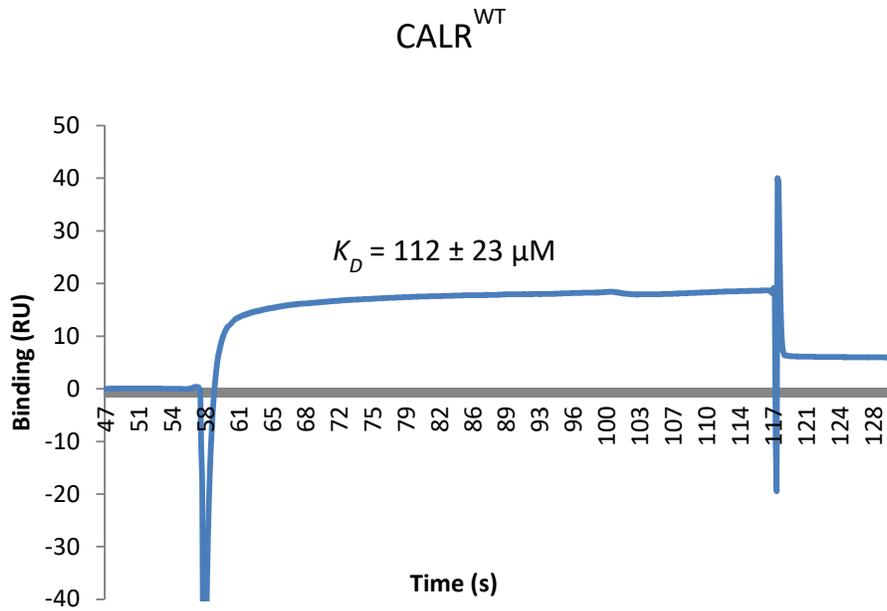
**Table 3. Summary of Ca<sup>2+</sup> binding affinities of CALR mutants**

CALR	Binding affinity to Ca <sup>2+</sup> (µM)
Wild-type	772.75 ± 28.50
P245L	628.21 ± 26.13
D295N	669.11 ± 23.24
D302N	414.16 ± 33.46
S323C	2130.56 ± 25.52
K355 del	1576.20 ± 24.21
E380G	878.74 ± 18.74
E381A	562.14 ± 23.93

### **Binding to oxidoreductase PDIA3**

The thiol oxidoreductase PDIA3 is an ER molecular chaperone that interacts with CALR for synergistic chaperone activity of glycosylated substrates, where PDIA3 aids in formation of disulfide bonds for faster folding rate. Among other functions, PDIA3 also binds to the ER  $\text{Ca}^{2+}$  sensor responsible for SOCE: STIM1, and binds to ER  $\text{Ca}^{2+}$  uptake protein SERCA, having therefore a possible role in modulation of  $\text{Ca}^{2+}$  homeostasis.

The *in vitro* interaction of PDIA3 with  $\text{CALR}^{\text{WT}}$  or mutants was examined by surface plasmon resonance (BIAcore). Purified CALR was immobilized to BIAcore sensor chips to test binding of purified PDIA3 at different concentrations. Representative binding curves sensorgrams from BIAcore software are presented in Figure 7 (a-g), and binding affinities are shown in Table 4. Except for  $\text{CALR}^{\text{E381A}}$ , all mutants bound PDIA3 with higher affinity than  $\text{CALR}^{\text{WT}}$ .



**Figure 7. Surface plasmon resonance (BIAcore) analysis of the interaction of PDIA3 with CALR<sup>WT</sup> and CALR mutants**

Purified PDIA3 was immobilized to a BIAcore CM5 chip and aliquots up to 2  $\mu\text{M}$  of CALR<sup>WT</sup> or CALR mutants were added as ligand. Representative sensorgrams of 3 independent analyses are shown.  $K_D$  values were calculated using BIAcore T200 software analysis. (Performed by Dr. Elzbieta Dudek).

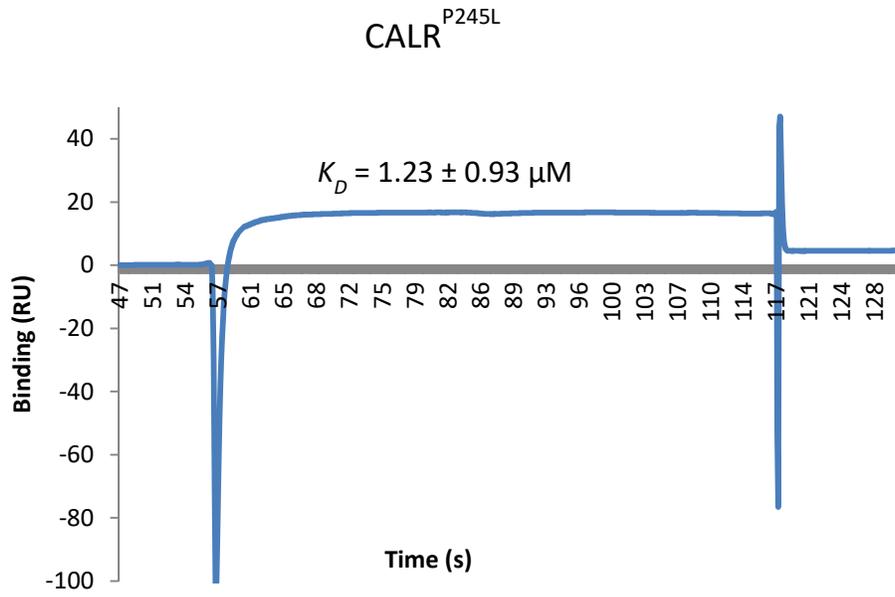


Figure 7 a. BiACore analysis of the interaction of PDIA3 with CALR<sup>P245L</sup>

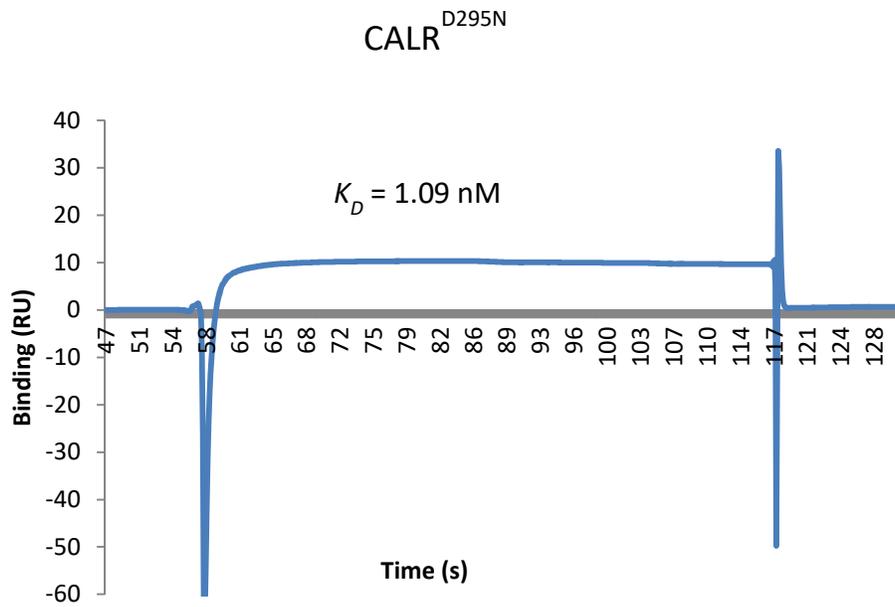


Figure 7 b. BiACore analysis of the interaction of PDIA3 with CALR<sup>D295N</sup>

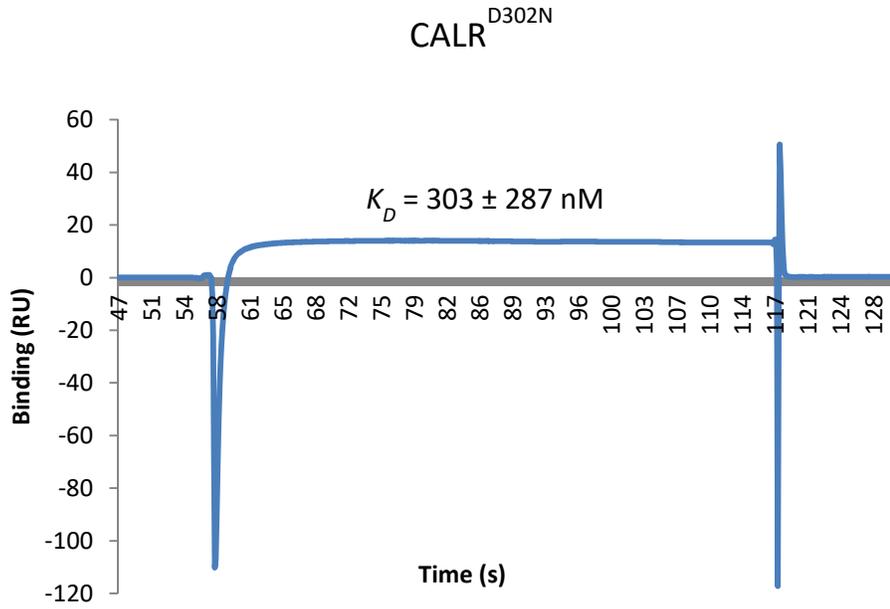


Figure 7 c. BiACore analysis of the interaction of PDIA3 with CALR<sup>D302N</sup>

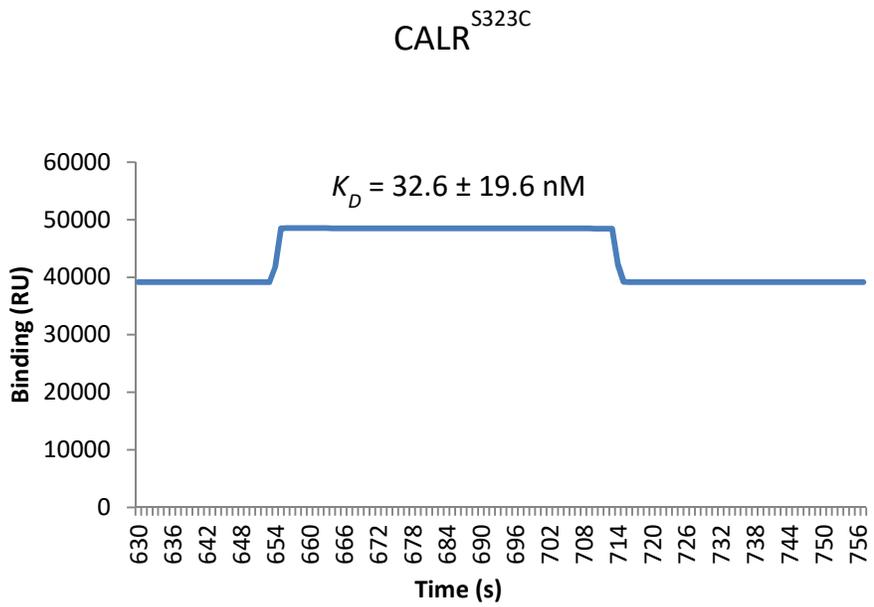


Figure 7 d. BiACore analysis of the interaction of PDIA3 with CALR<sup>S323C</sup>

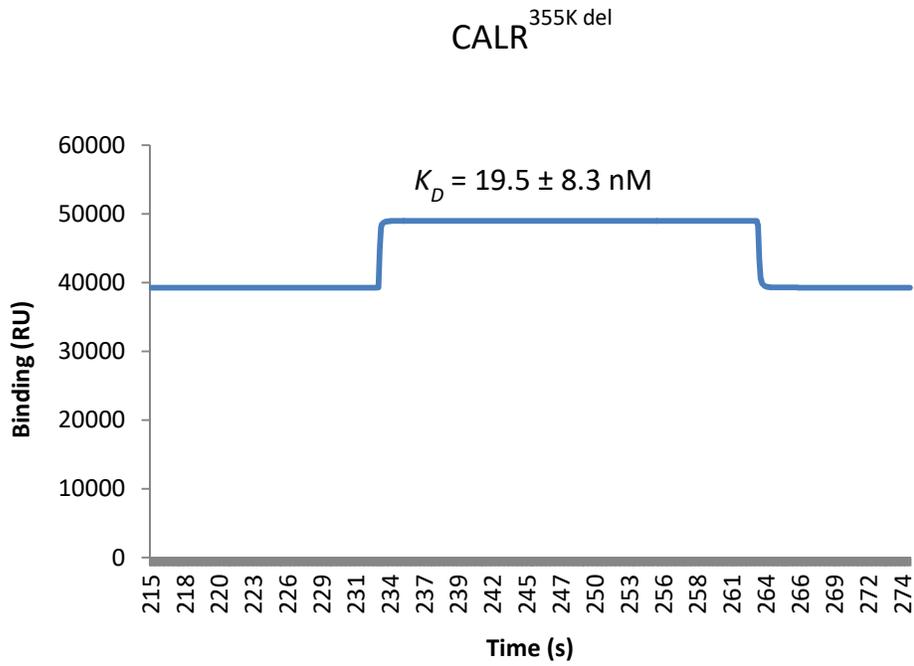


Figure 7 e. BiACore analysis of the interaction of PDIA3 with CALR<sup>K355 del</sup>

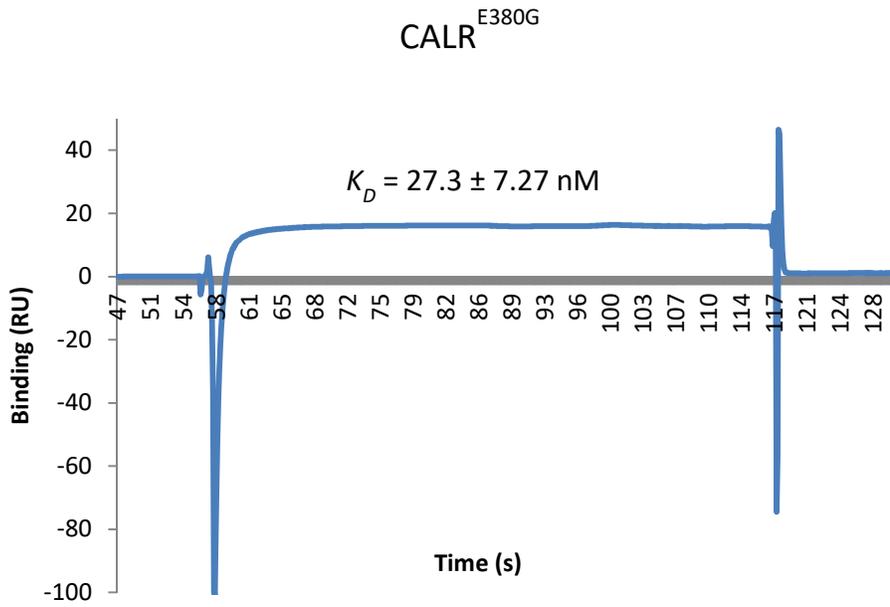
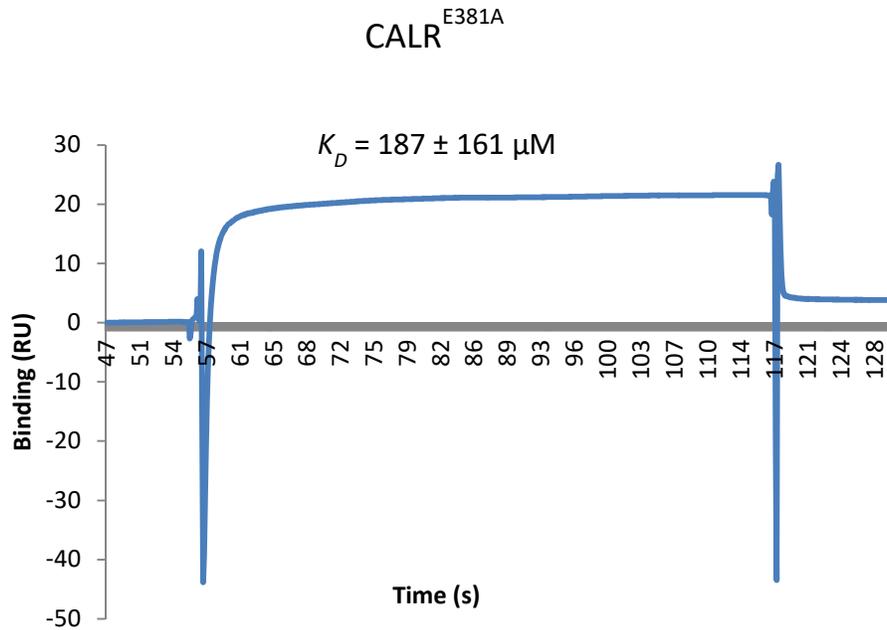


Figure 7 f. BiACore analysis of the interaction of PDIA3 with CALR<sup>E380G</sup>



**Figure 7 g. BIAcore analysis of the interaction of PDIA3 with CALR<sup>E381A</sup>**

**Table 4. Summary of PDIA3 binding affinities of CALR mutants**

CALR	Binding affinity to PDIA3
Wild-type	$112 \pm 23 \mu\text{M}$
P245L	$1.23 \pm 0.93 \mu\text{M}$
D295N	$1.09 \text{ nM}^*$
D302N	$303 \pm 287 \text{ nM}$
S323C	$32.6 \pm 19.6 \text{ nM}$
K355 del	$19.5 \pm 8.3 \text{ nM}$
E380G	$27.3 \pm 7.27 \text{ nM}$
E381A	$187 \pm 161 \mu\text{M}$

n = 3

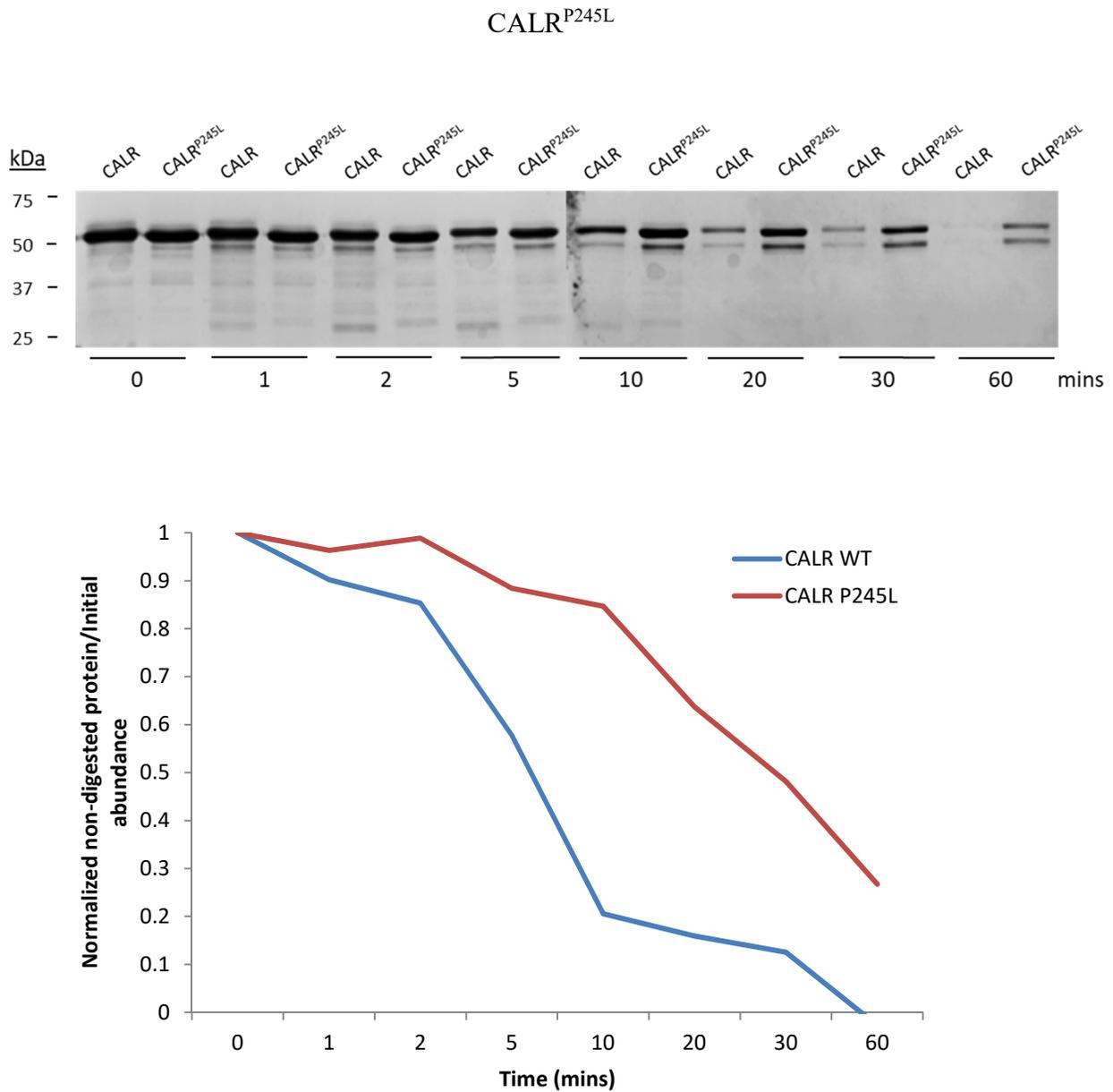
\*n = 1

## Folding of CALR

In order to perform their biological function, proteins must adopt their proper three-dimensional conformation. The information for this process is encoded in their amino acid sequence [107] and involves molecular chaperones and folding enzymes [108-110]. The mutated amino acid residues of CALR in SUD could alter the proper folding of the protein, and as molecular chaperone, this could impact the proper folding of substrate proteins and macromolecules.

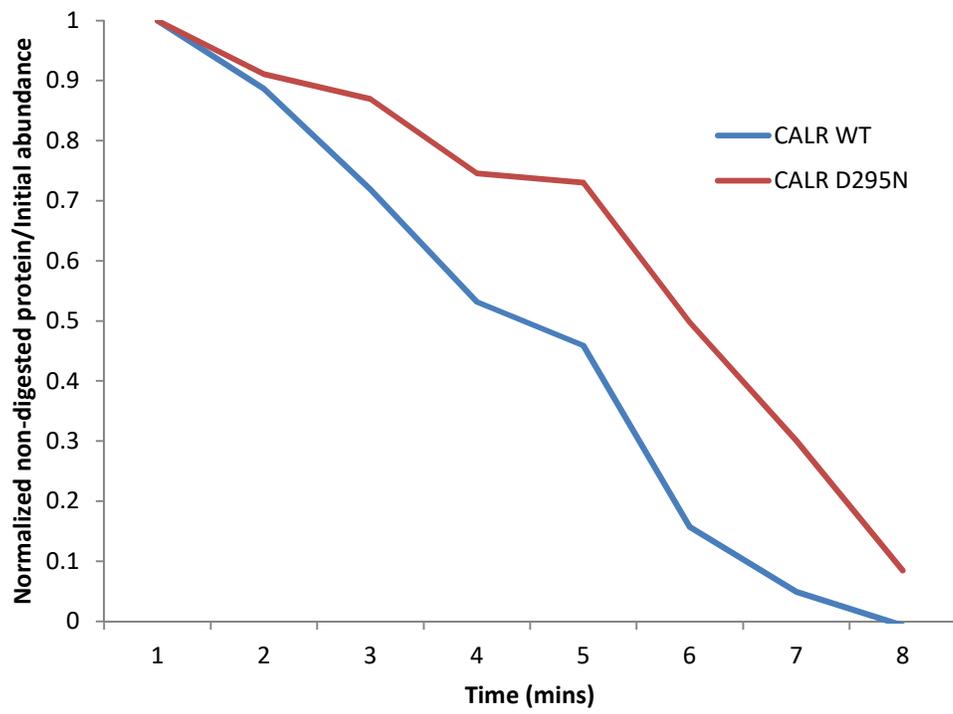
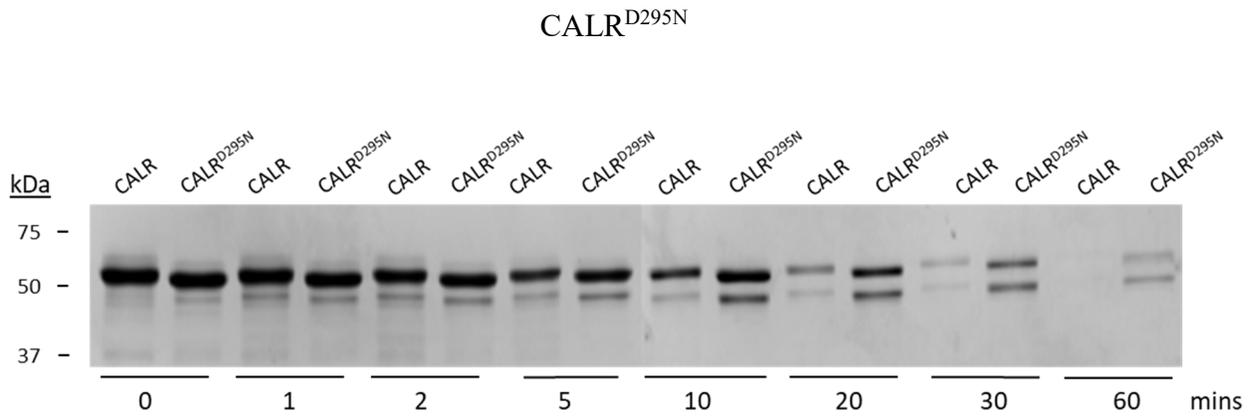
To investigate the effect of the mutations on CALR structure, the folding of CALR mutants was examined by comparing its resistance to proteolytic digestion by trypsin to that of CALR<sup>WT</sup>. CALR<sup>WT</sup> or mutants were incubated with trypsin for 1 hr, and aliquots were taken at different time-points and analyzed by SDS-PAGE, followed by Coomassie blue staining.

Remarkably, all CALR mutants, except for CALR<sup>K355 del</sup>, exhibited higher resistance to digestion by trypsin over time compared to CALR<sup>WT</sup> [Figure 8 (a-f), top bands at 55-kDa for full protein] indicating a different accessibility of the enzyme and thus a different folding arrangement of the mutant proteins. Overall, CALR<sup>K355 del</sup> seemed to be digested faster than CALR<sup>WT</sup>, also indicating a different folding arrangement, but there was a persistent non-cleaved peptide at 50-kDa after 1 hr, when all CALR<sup>WT</sup> had been cleaved. Interestingly, a similar unique resistant peptide at around 50-kDa was observed for CALR<sup>S323C</sup> and CALR<sup>E381A</sup>, but not the other mutants. CALR<sup>P245L</sup> displayed different mobility of a peptide at 28-kDa compared to CALR<sup>WT</sup>. The results indicate that CALR mutants are folded differently to wild-type protein, and there are some unique effects of the amino acid changes at specific locations of the protein on the accessibility of trypsin, and even on the folding pattern of the cleaved peptides.

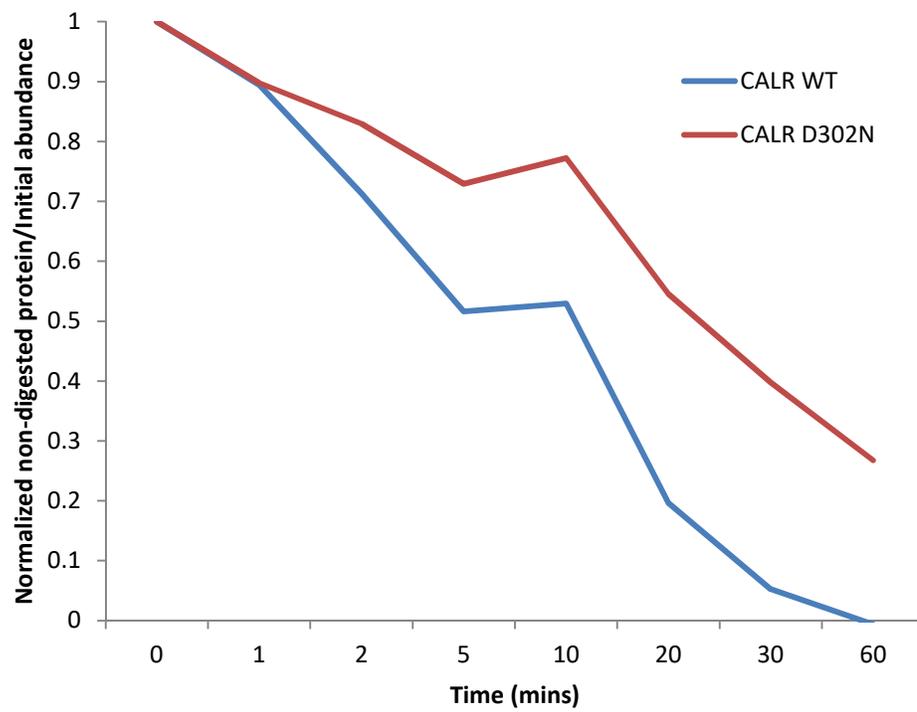
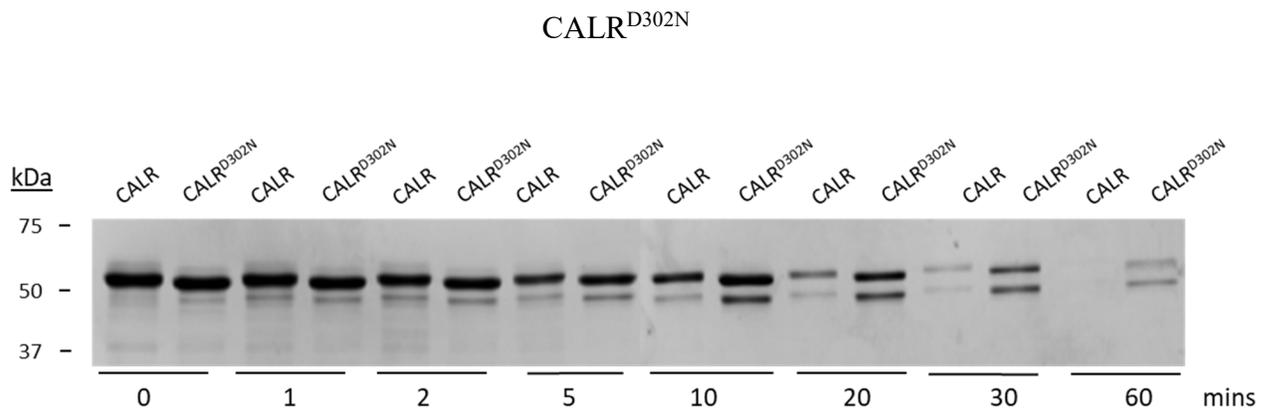


**Figure 8. Proteolytic digestion of CALR<sup>WT</sup> and CALR<sup>P245L</sup>**

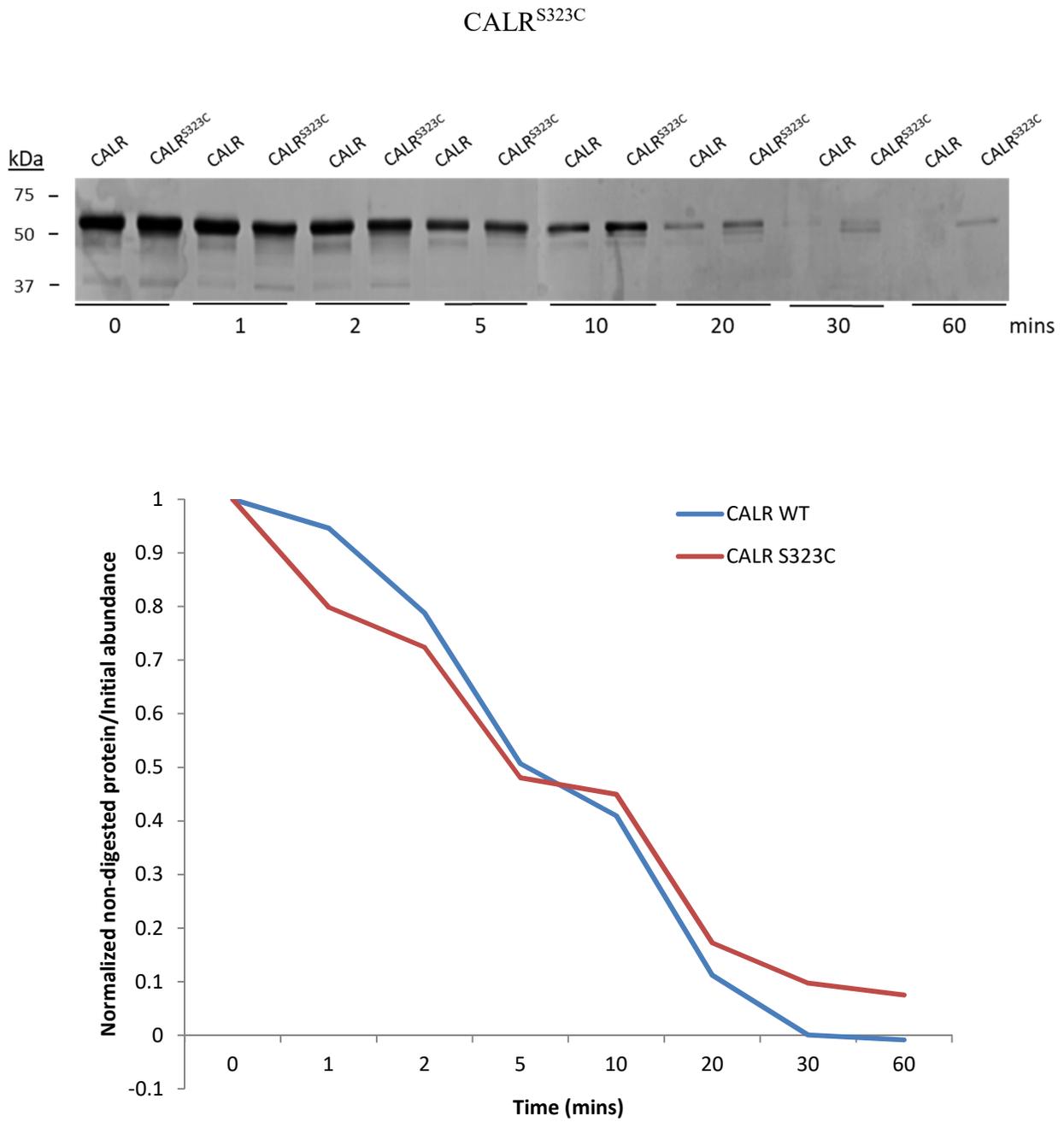
Trypsin digestion analysis of CALR<sup>WT</sup> and mutants (shown is CALR<sup>P245L</sup>). Purified proteins were incubated with trypsin at 1:100 (trypsin/protein; w/w) at 37° C. Aliquots were taken at the indicated time-points. The proteins were separated by SDS-PAGE and stained with Coomassie blue (top). Quantitative analysis of Coomassie blue staining showing the ratio of non-digested protein to initial abundance over time (bottom).



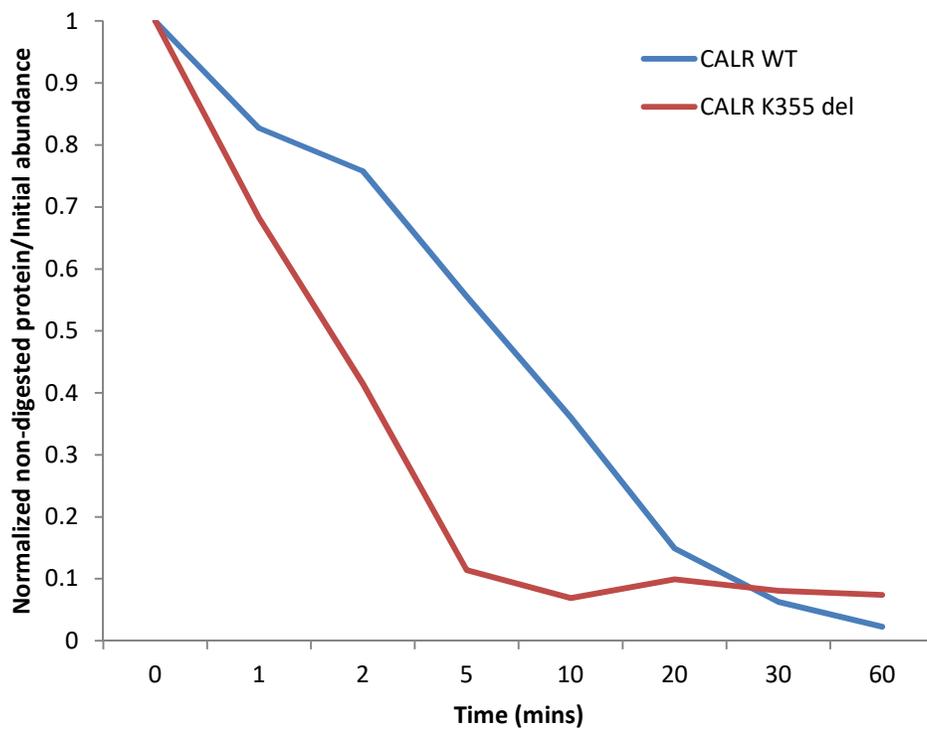
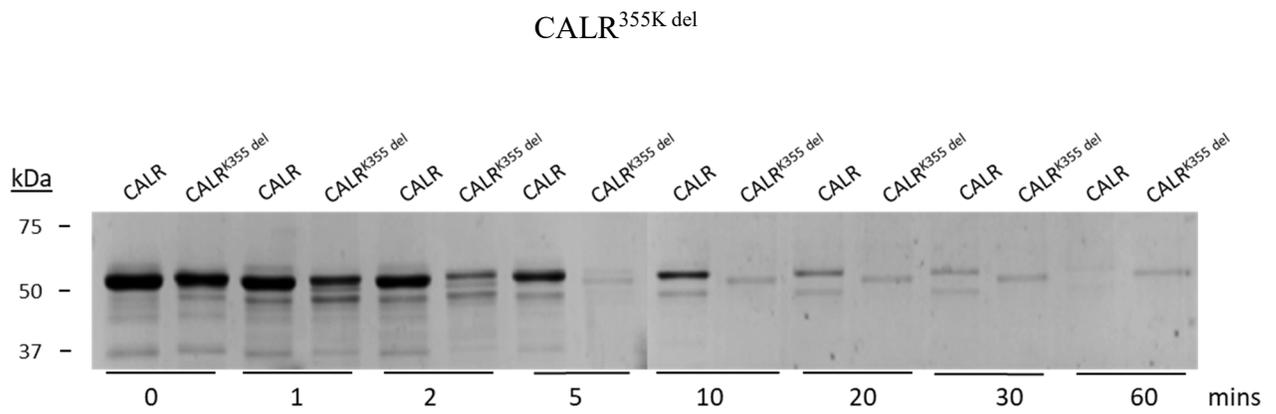
**Figure 8 a. Proteolytic digestion of CALR<sup>D295N</sup>**



**Figure 8 b. Proteolytic digestion of CALR<sup>D302N</sup>**



**Figure 8 c. Proteolytic digestion of CALR<sup>S323C</sup>**



**Figure 8 d. Proteolytic digestion of CALR<sup>K355del</sup>**

CALR<sup>E380G</sup>

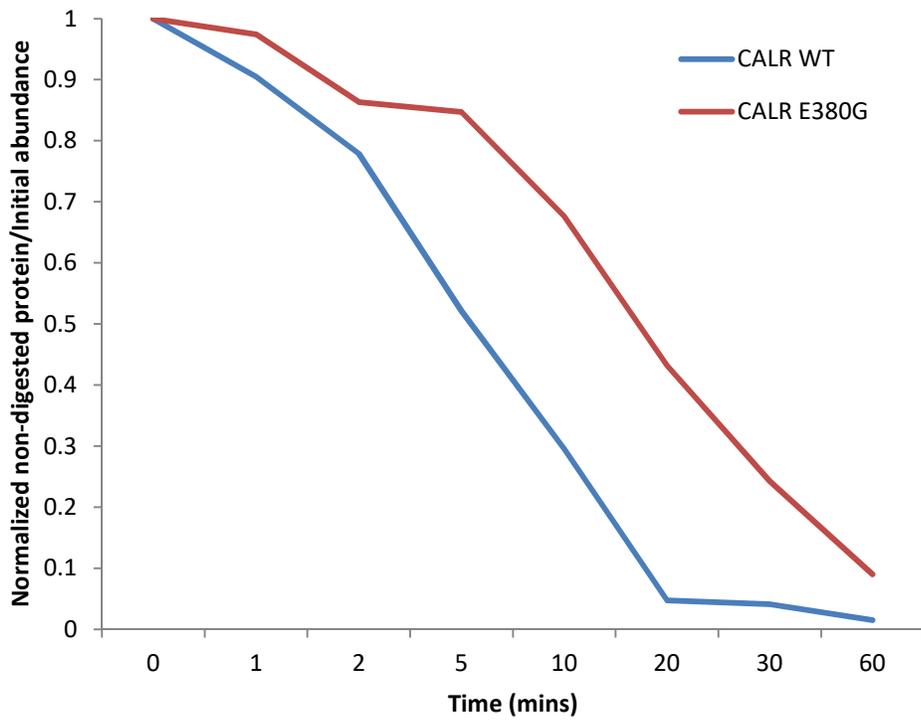
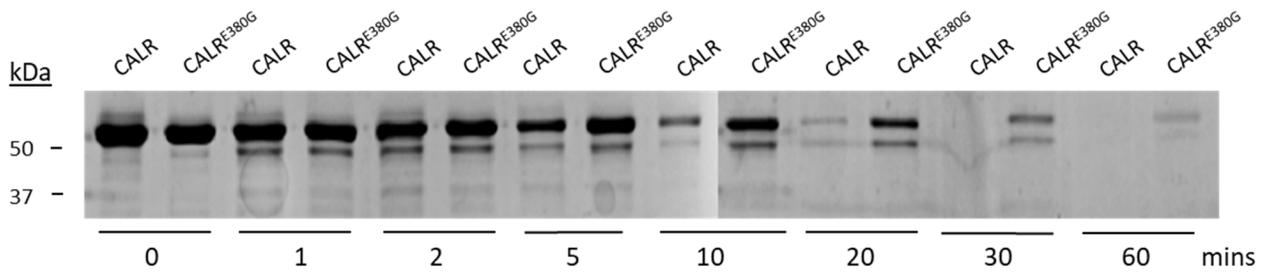


Figure 8 e. Proteolytic digestion of CALR<sup>E380G</sup>

CALR<sup>E381A</sup>

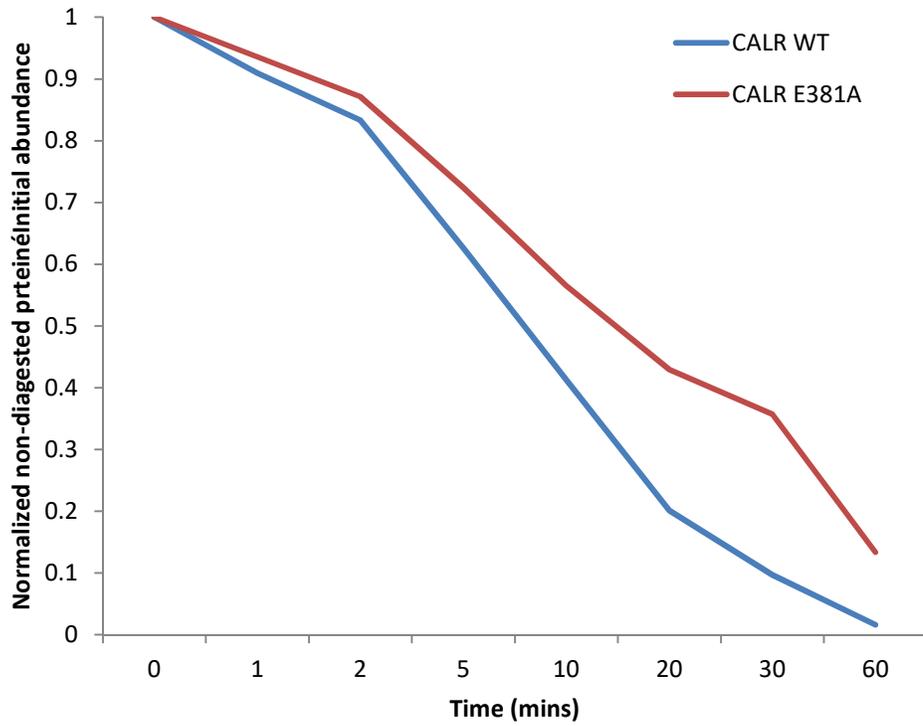
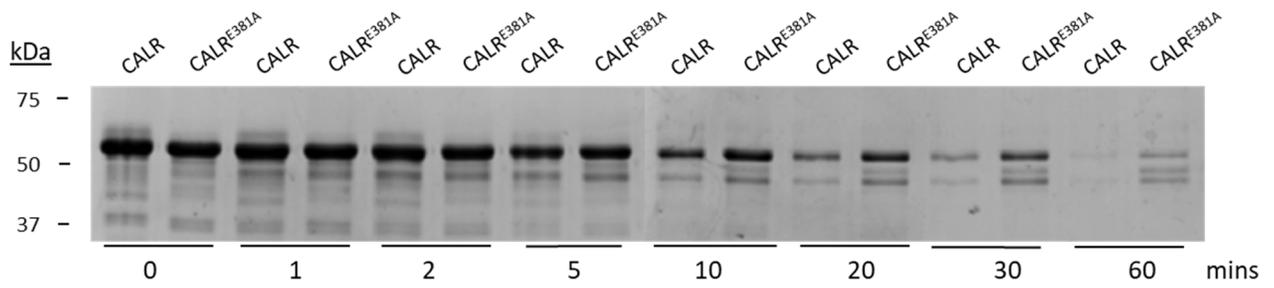


Figure 8 f. Proteolytic digestion of CALR<sup>E381A</sup>

## CHAPTER FOUR:

# Characterization of CALR<sup>376fs</sup> Mutant

---

## CALR<sup>WT</sup> vs CALR<sup>376fs</sup>

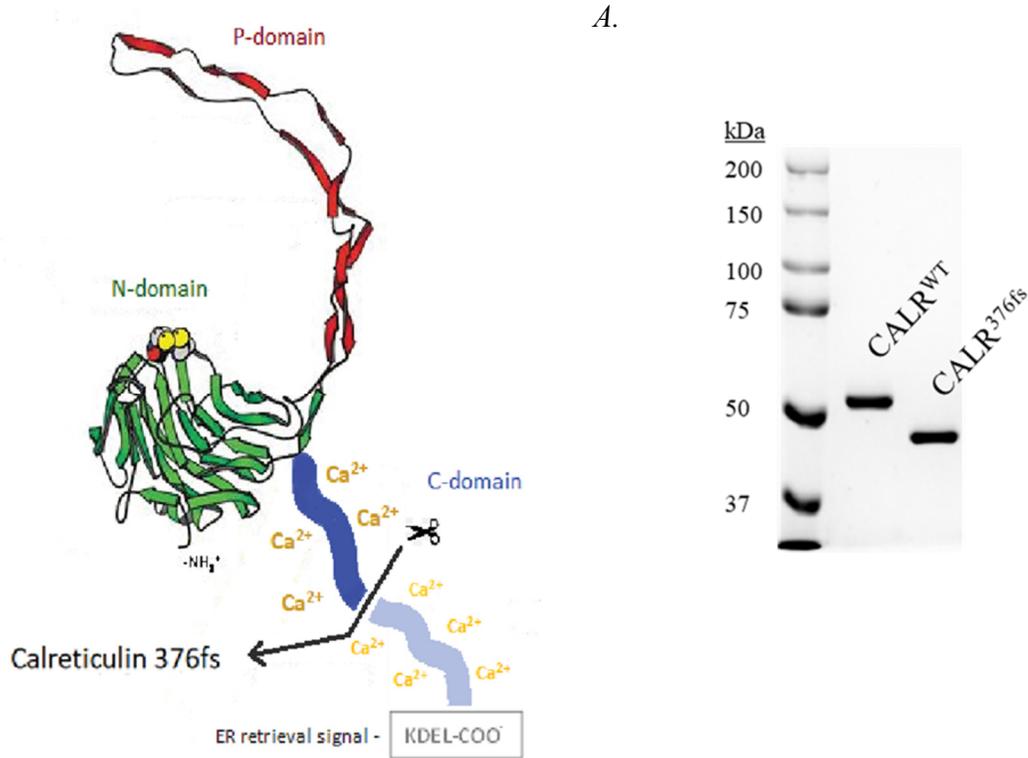
The focus of this project was an ultra-rare non-synonymous frame-shift mutation in the *CALR* gene (c.del1130\_1131AA) encoding the truncated mutant protein CALR<sup>376fs</sup> (Figure 9A).

CALR<sup>WT</sup> and CALR<sup>376fs</sup> proteins were purified at the Institute of Biotechnology (Vilnius, Lithuania) as outlined in methods and as reported previously [100]. The proteins were run in SDS-PAGE and stained with Coomassie blue, showing a single band with mobility at molecular weights of 55-kDa and 45-kDa for CALR<sup>WT</sup> and CALR<sup>376fs</sup>, respectively (Figure 9A).

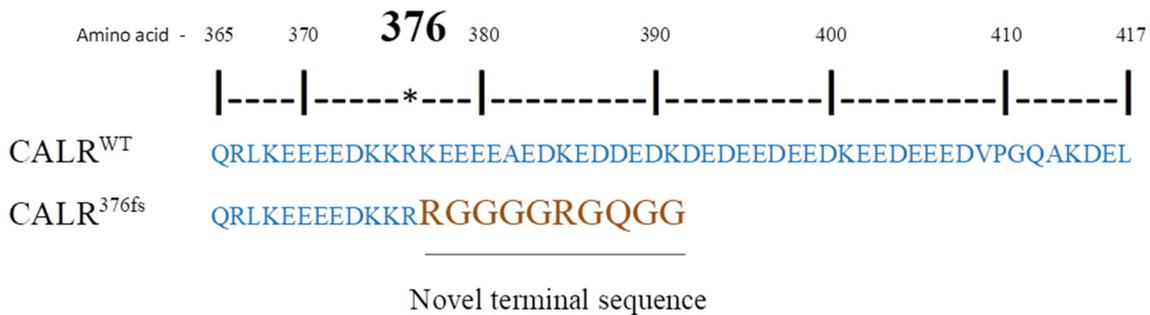
The CALR<sup>376fs</sup> mutant results in a shorter protein than CALR<sup>WT</sup> because of a deletion of 41 amino acids at the Ca<sup>2+</sup>-buffering C-terminus, ending with a novel basic amino acid motif (RGGGGRGQGG) added after arginine (R) 376, instead of a highly basic C-tail, with also loss of a KDEL ER retrieval signal. Therefore, CALR<sup>376fs</sup> is 31 amino acids shorter than CALR<sup>WT</sup> (Figure 9B).

Matrix assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) analysis was used to validate the identity of purified CALR<sup>WT</sup> and CALR<sup>376fs</sup> [111, 112], which confirmed human CALR as sole source of protein (Figure 9C, top). Overlap of CALR<sup>376fs</sup> tryptic peptide fingerprint obtained from MALDI-TOF MS against CALR sequence is shown (Figure 9C, bottom).

A comparison of basic properties of CALR<sup>WT</sup> and CALR<sup>376fs</sup> is shown in Table 5 .



**B.**



**Figure 9. CALR<sup>WT</sup> and CALR<sup>376fs</sup> structure and sequence comparison**

**A.** Three-dimensional model of CALR showing a cut at the C-domain for mutant CALR<sup>376fs</sup> (left), and mobility in SDS-PAGE of the purified proteins evaluated by Coomassie blue staining (right). **B.** CALR<sup>WT</sup> and CALR<sup>376fs</sup> sequences show C-terminus amino acids 365 to 417, indicating the novel terminal sequence CALR<sup>376fs</sup> after amino acid 376. **C.** Screenshots showing the result MALDI-TOF MS analysis (top) and tryptic peptides fingerprint identified in CALR<sup>376fs</sup> overlapped to the sequence of human CALR (bottom) (next page).

C.

1	Accession	Description	Score	Coverage	# Proteins	# Unique Peptides	# Peptides	# PSMs	# AAs	
2	P27797	Calreticulin OS=Homo sapiens	169.05	42.93	1	11	11	51	417	
3		A3	Sequence		# PSMs	# Proteins	# Protein Groups	Protein Group Accessions	Modifications	ΔCn
4		High	LFPNSLDQTDmHGDSEYNIImFGP		3	1	1	P27797	M11(Oxidation); M20(Oxidation)	0.0000
5		High	IDNSQVESGSLEDDWDFLPPK		3	1	1	P27797		0.0000
6		High	KPEDWDEEMDGEWEPPVIQNPE		2	1	1	P27797		0.0000
7		High	CKDDEFTHLYTLIVRPDNTYEVK		4	1	1	P27797	C1(Carbamidomethyl)	0.0000
8		High	AKIDDPDTSKPEDWDKPEHIPDPI		2	1	1	P27797		0.0000
9		High	KPEDWDEEMDGEWEPPVIQNPE		5	1	1	P27797	M9(Oxidation)	0.0000
10		High	LFPNSLDQTDmHGDSEYNIImFGP		6	1	1	P27797	M11(Oxidation); C26(Carbamidomethyl)	0.0000
11		High	FYALSASFEPFSNK		8	1	1	P27797		0.0000
12		High	IDNSQVESGSLEDDWDFLPPKK		5	1	1	P27797		0.0000
13		High	DDEFTHLYTLIVRPDNTYEVK		2	1	1	P27797		0.0000
14		High	HEQNIDcGGGYK		5	1	1	P27797	C7(Carbamidomethyl)	0.0000
15		High	IKDPDASKPEDWDER		4	1	1	P27797		0.0000
16		High	LFPNSLDQTDmHGDSEYNIImFGP		1	1	1	P27797	C26(Carbamidomethyl)	0.0000
17		High	GQTLVQFTVK		1	1	1	P27797		0.0000
18	P04264	Keratin, type II cytoskeletal 1	11.70	7.14	1	4	4	4	644	
27										
28										

```

10          20          30          40          50
MLLSVPLLLG LLGLAVAEPV VYFKEQFLDG DGWTSRWIES KHKSDFGKFFV
60          70          80          90          100
LSSGKFFYGDE EKDKGLQTSQ DARFYALSAS FEPFSNKGOT LVVQFTVKHE
110         120         130         140         150
QNIDCGGGYV KLFNSLDQT DMHGDSEYNI MFGPDICGPG TKRVHVFNFY
160         170         180         190         200
KGNVNLINKD IRCKDDEFTH LYTLIVRPDN TYEVKIDNSQ VESGSLEDDW
210         220         230         240         250
DFLPPKKIKD PDASKPEDWD ERAKIDDPDTSKPEDWDKPE HIPDPDAKPE
260         270         280         290         300
EDWDEEMDGE WEPPVIQNPE YKGEWKPRQI DNPDYKGTWI HPEIDNPEYS
310         320         330         340         350
PDPSIYAYDN FGVLGDLWQ VKSGTIFDNF LIINDEAYAE EFGNETWGVN
360         370         380         390         400
KAAEKQMKDK QDEEQRLKEE EEDKKRKEEE EAEDKEDDED KDEDEEDED
410
KEEDEEEDVP GQAKDEL

```

Figure 9 (continued from previous page)

**Table 5. Comparison of basic properties of CALR<sup>WT</sup> and CALR<sup>376fs</sup>**

<b>CALR</b>	<b>WT</b>	<b>376fs</b>
Length	417 aa	386 aa
Predicted mass	47-kDa	43-kDa
Mobility in SDS-PAGE	55-kDa	45-kDa
pI	4.44	4.71
Charge at pH 7.0	-57.3	-31.4

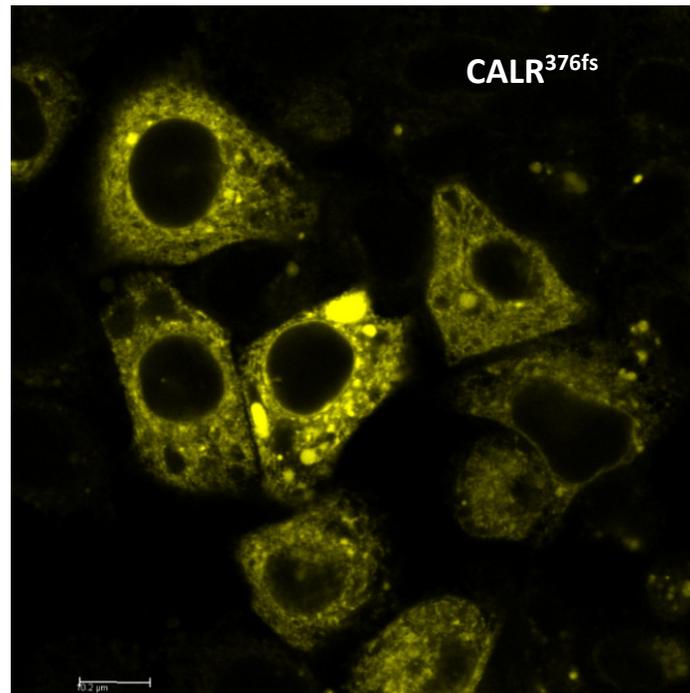
Note: Values obtained by inputting CALR<sup>WT</sup> and CALR<sup>376fs</sup> amino acid sequences in Protein Calculator v3.4 (<http://protecalc.sourceforge.net/>)

### **Cellular localization of the CALR<sup>376fs</sup> mutant**

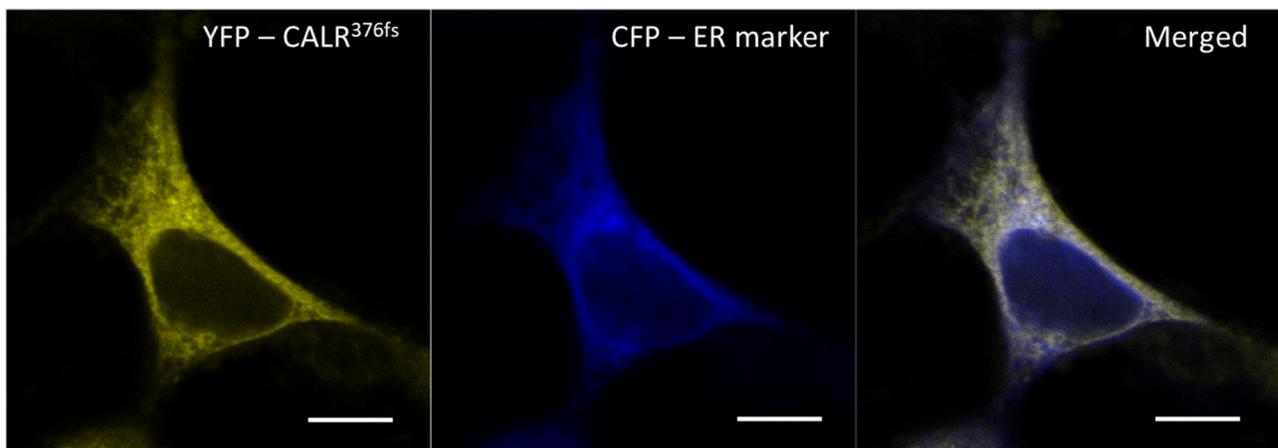
Considering that the CALR<sup>376fs</sup> mutant has lost part of its C-terminus, including the ER retrieval KDEL signal, it was necessary to check the cellular localization of the mutant. Other C-terminus deletion CALR mutants implicated in myeloproliferative neoplasms (MPNs) have been found outside of the ER [95, 96, 113].

A YFP-tagged CALR<sup>376fs</sup> mutant was expressed in HeLa cells (Figure 10A), or in HEK293 cells along with a CFP-ER marker (Figure 10B), and fluorescence analysis was performed by confocal microscopy. Overlap comparison of the fluorescence of the mutant to a CFP-ER marker indicated that CALR<sup>376fs</sup> localized to the ER, similar to the other mutants evaluated.

A.



B.



**Figure 10. Cellular localization of CALR<sup>376fs</sup>**

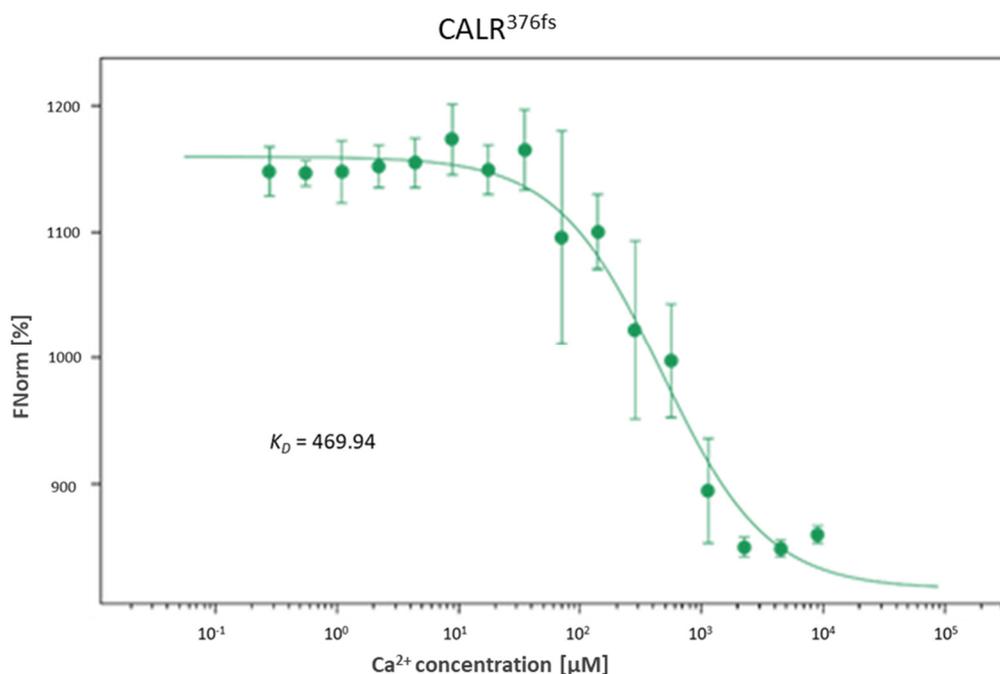
*A.* Intracellular distribution of YFP-CALR and mutants in HeLa cells visualized by fluorescent microscopy.

*B.* Intracellular distribution of YFP-CALR<sup>376fs</sup> mutant in HEK293 cells, with cyan fluorescent protein (CFP) targeted to the ER used as control. The overlap of YFP-CALR<sup>376fs</sup> and CFP-ER is shown. Pearson's coefficient = 0.84. Scale bars: 10 μM. (Performed by Dr. Elzbieta Dudek).

## Ca<sup>2+</sup> binding affinity to the CALR<sup>376fs</sup> mutant

Considering that the C-domain of CALR accounts for the high buffering activity of the protein, and since the CALR<sup>376fs</sup> mutant resulted in a shorter protein than CALR<sup>WT</sup> because of a deletion of 41 amino acids at the C-terminus, with a 10 non-acidic and slightly basic amino acid motif added, it was important to investigate whether the mutant had a decrease in Ca<sup>2+</sup>-binding ability compared to wild-type.

Surprisingly, like the other mutants CALR<sup>376fs</sup> retained the ability to bind Ca<sup>2+</sup>, and instead showed a higher binding affinity ( $K_D = 469.94 \pm 19.78 \mu\text{M}$ ) compared to CALR<sup>WT</sup> ( $K_D = 772.75 \pm 28.50 \mu\text{M}$ ) (Figure 11).

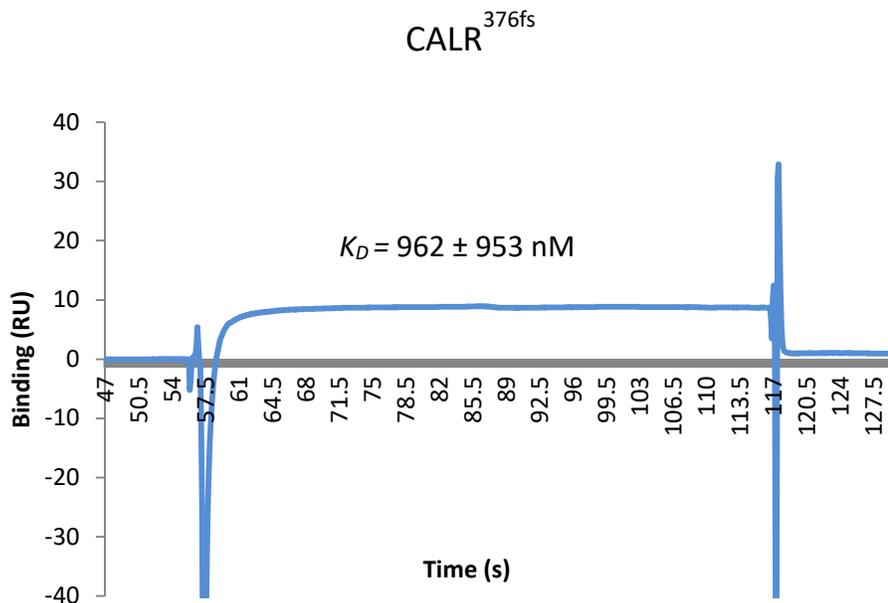


**Figure 11. Calcium binding affinity of CALR<sup>376fs</sup>**

Microscale thermophoresis (MST) of the interaction of Ca<sup>2+</sup> with CALR<sup>376fs</sup>. Curves are derived from the specific change in thermophoretic mobility upon titration of ligand (Ca<sup>2+</sup>) to a constant CALR<sup>376fs</sup> concentration.  $K_D$  value was calculated using NanoTemper analysis software. Error bars = s.d.;  $n = 7$ . (Performed by Dr. Elzbieta Dudek).

## Binding of oxidoreductase PDIA3 to the CALR<sup>376fs</sup> mutant

As done before for the other CALR mutants, in order to analyze the effect of a C-terminus truncation on the ability of CALR to bind a synergistic folding protein, the *in vitro* interaction of PDIA3 with CALR<sup>WT</sup> or CALR<sup>376fs</sup> was examined by BIAcore. Similar to most other mutants, it was found that CALR<sup>376fs</sup> had a stronger binding affinity to PDIA3 ( $K_D = 962 \pm 953$  nM) compared to CALR<sup>WT</sup> ( $K_D = 112 \pm 23$   $\mu$ M) (Figure 12).



**Figure 12. Surface plasmon resonance (BIAcore) analysis of the interaction between PDIA3 and CALR<sup>376fs</sup>**

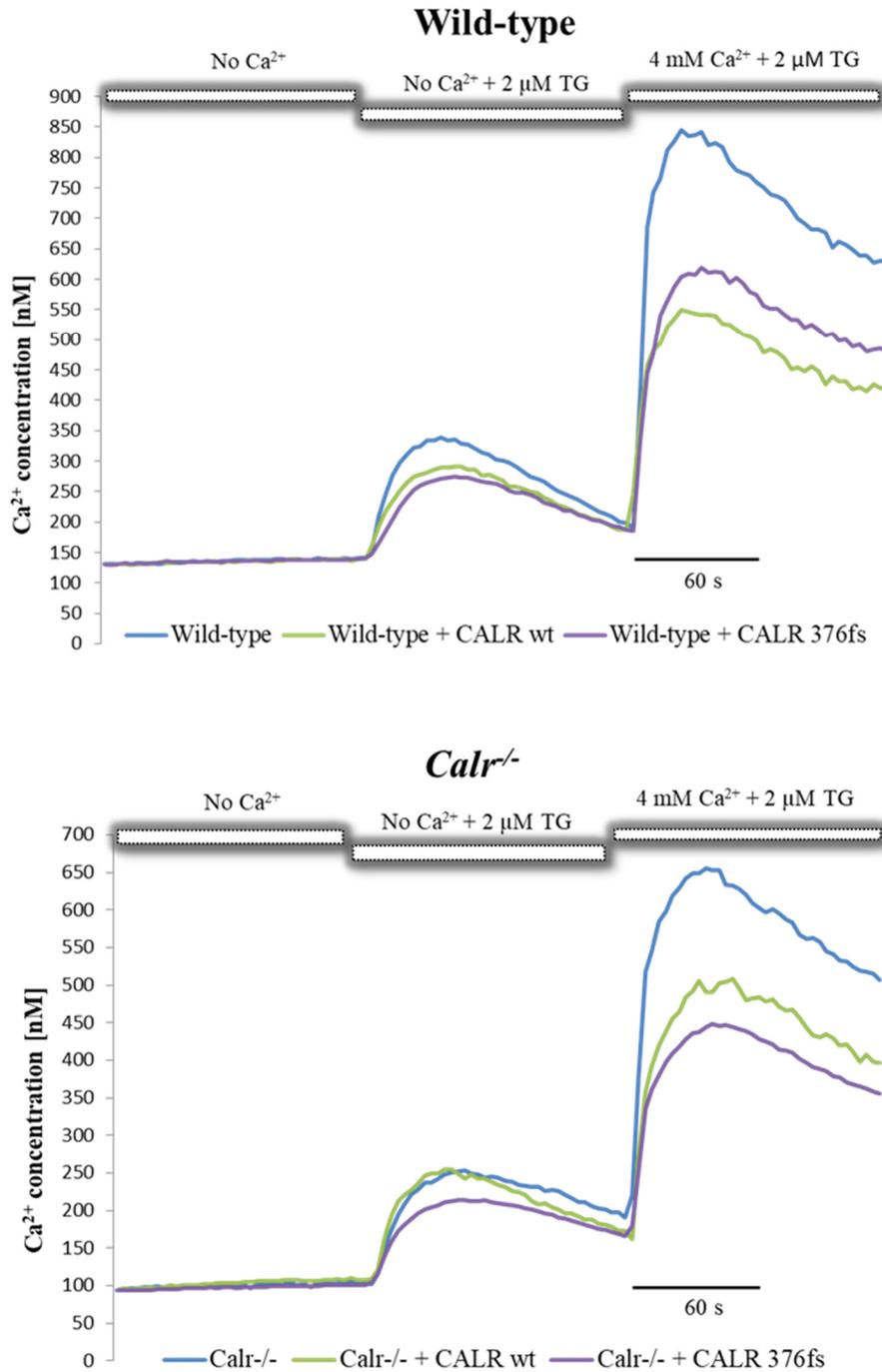
Purified CALR<sup>376fs</sup> was immobilized to a BIAcore CM5 chip and 2  $\mu$ M PDIA3 was added as ligand. Representative sensorgram of 3 independent analyses is shown.  $K_D$  value was calculated using BIAcore T200 software analysis. (Performed by Dr. Elzbieta Dudek).

## Ca<sup>2+</sup> dynamics of the CALR<sup>376fs</sup> mutant

Since the CALR<sup>376fs</sup> mutant had part of its Ca<sup>2+</sup>-buffering C-tail deleted, but it showed stronger binding affinity *in vitro* to Ca<sup>2+</sup> and PDIA3, the effect of CALR<sup>376fs</sup> on intracellular Ca<sup>2+</sup> dynamics was examined by transfecting the mutant into wild-type and *Calr* knockout (*Calr*<sup>-/-</sup>) mouse embryonic fibroblasts (MEFs). The cells were incubated in the fluorescent Ca<sup>2+</sup> indicator fura-2 AM and fluorescence was measured in a Ca<sup>2+</sup>-free buffer, inducing ER Ca<sup>2+</sup> release by adding thapsigargin, followed by SOCE by adding Ca<sup>2+</sup>.

Wild-type cells did not show any difference in free ER Ca<sup>2+</sup> release with CALR<sup>WT</sup> or CALR<sup>376fs</sup> mutant, and *Calr*<sup>-/-</sup> cells showed only a slight release decrease with the mutant (Figure 13A). Overall, there was no significant difference in Ca<sup>2+</sup> dynamics in cells expressing CALR<sup>WT</sup> or CALR<sup>376fs</sup>. However, both CALR<sup>WT</sup> and CALR<sup>376fs</sup> induced a decrease in SOCE in wild-type and *Calr*<sup>-/-</sup> cells. The differences in SOCE with CALR<sup>WT</sup> or CALR<sup>376fs</sup> were small and might be affected by expression of the proteins. Indeed, immunoblotting with anti-CALR antibodies indicated that the CALR<sup>376fs</sup> mutant was less abundant than CALR<sup>WT</sup> in both cell types (Figure 13B).

A.



**Figure 13. Ca<sup>2+</sup> dynamics of cells expressing the CALR<sup>376fs</sup> mutant**

A. ER luminal Ca<sup>2+</sup> and SOCE of wild-type (top) and *Calr*<sup>-/-</sup> (bottom) cells expressing CALR<sup>WT</sup> or CALR<sup>376fs</sup> were analyzed by incubation with the fluorescent cytoplasmic Ca<sup>2+</sup> indicator fura-2 AM and by measuring fluorescence at 340 nm in a Ca<sup>2+</sup>-free buffer. Thapsigargin (TG) and Ca<sup>2+</sup> were added at the indicated time-points to induce ER Ca<sup>2+</sup> release and SOCE, respectively. B. Immunoblot analysis of CALR<sup>WT</sup> and CALR<sup>376fs</sup> mutant (next page).

B.

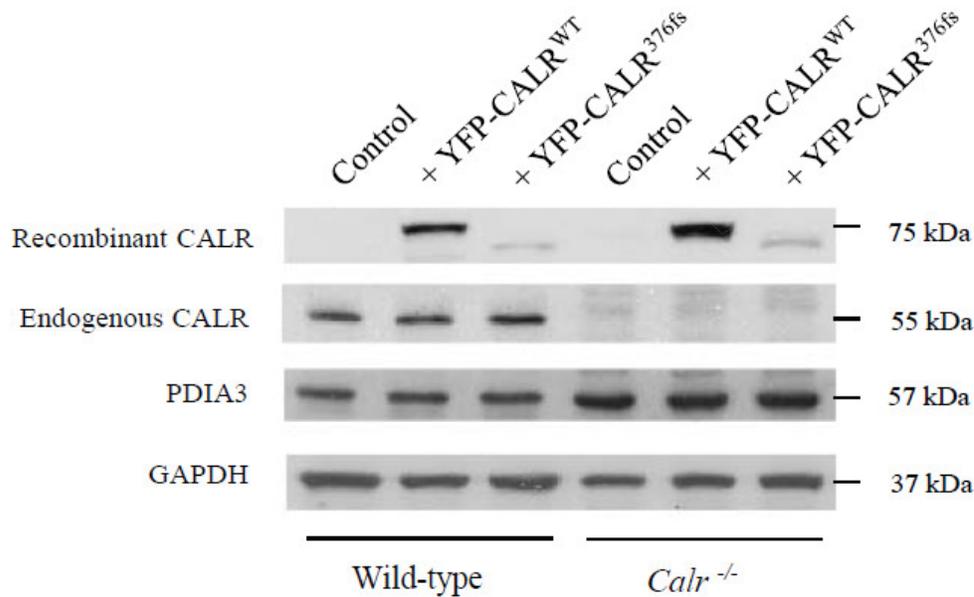


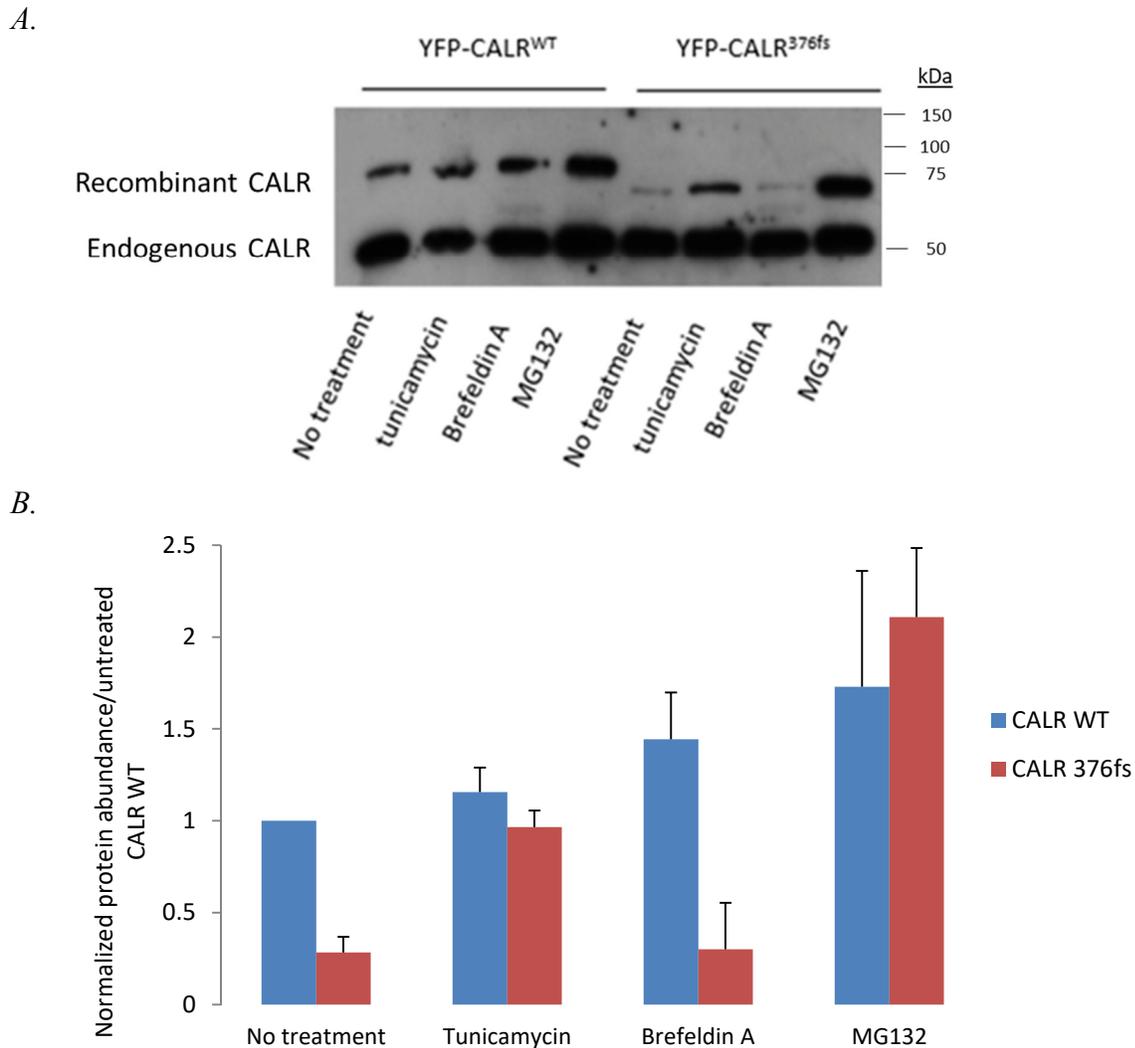
Figure 13 (continued from previous page)

### Processing of the CALR<sup>376fs</sup> mutant

The observed low abundance of CALR<sup>376fs</sup> in MEFs could be due to different processing of the mutant compared to wild-type protein. Furthermore, the tighter binding of CALR<sup>376fs</sup> to Ca<sup>2+</sup> and PDIA3 *in vitro*, but not much difference in Ca<sup>2+</sup> dynamics or in PDIA3 levels with expression of CALR<sup>376fs</sup> in cells, could similarly be due to alternative processing of CALR<sup>376fs</sup>.

In order to examine factors affecting processing and abundance, CALR<sup>376fs</sup> or CALR<sup>WT</sup> were expressed in HEK293 cells, followed by treatment with the N-glycosylation blocker tunicamycin, ER to Golgi transport blocker Brefeldin A, or proteasome inhibitor MG132. This was followed by immunoblotting with anti-CALR antibodies of obtained lysates.

CALR<sup>376fs</sup> seemed to be less abundant than CALR<sup>WT</sup> in HEK293 cells (Figure 14). There was some increase in mutant abundance with tunicamycin treatment, but not with BFA. However, treatment with MG132 led to a very high increase in abundance of the mutant. This indicates that the CALR<sup>376fs</sup> mutant was mostly being degraded.

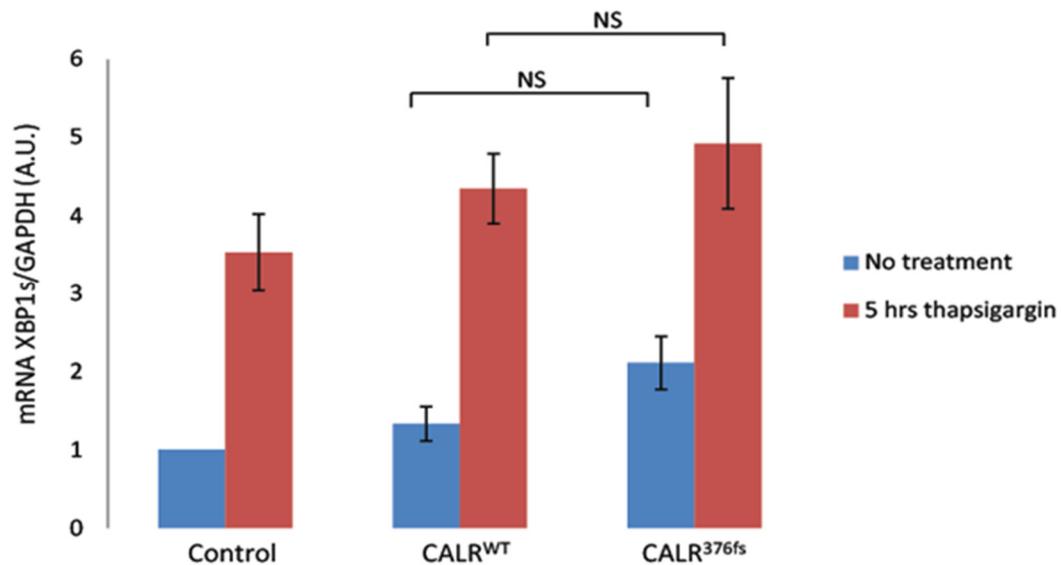


**Figure 14. Processing of CALR<sup>376fs</sup> in cells**

**A.** Representative immunoblot analysis of HEK293 cells expressing CALR<sup>WT</sup> or CALR<sup>376fs</sup> mutant and treated with N-glycosylation blocker tunicamycin (10  $\mu$ g/ml), ER to Golgi transport blocker Brefeldin A (5  $\mu$ g/ml), and proteasome inhibitor MG132 (1  $\mu$ M) for 16 hrs. **B.** Quantitative analysis of immunoblots showing the ratio of recombinant CALR to CALR<sup>WT</sup> with no treatment (n=2).

## Induction of UPR in cells expressing the CALR<sup>376fs</sup> mutant

Analysis of ER stress/UPR by spliced XBP1 (XBP1s) of HEK293 cells transfected with CALR<sup>WT</sup> or CALR<sup>376fs</sup> showed that the mutant caused a higher UPR than wild-type protein, however this was not statistically significant (Figure 15). Induction of UPR was confirmed by thapsigargin treatment. Therefore, the CALR<sup>376fs</sup> mutant is highly degraded, but does not greatly induce UPR.



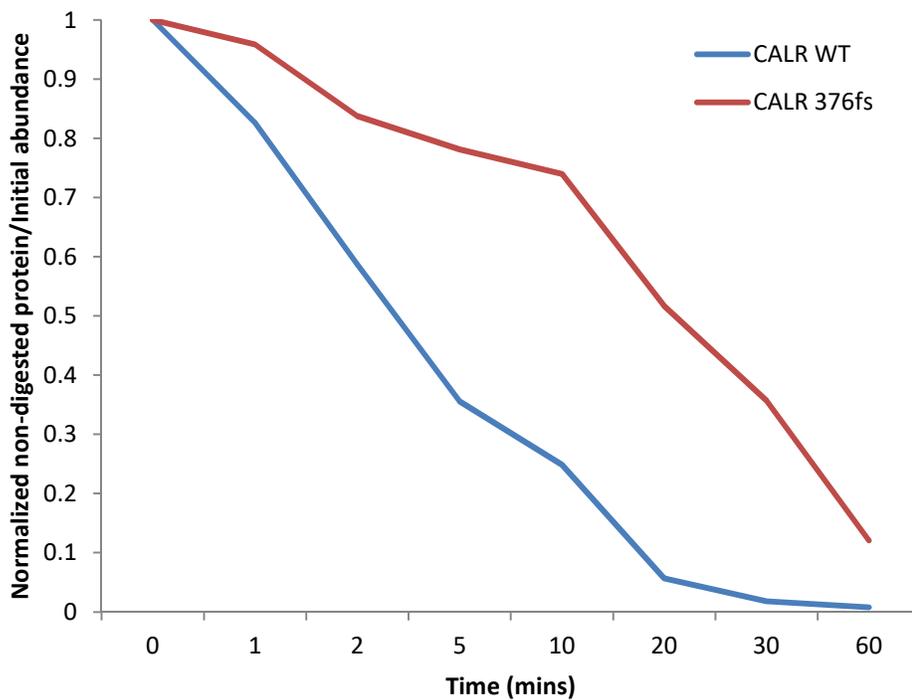
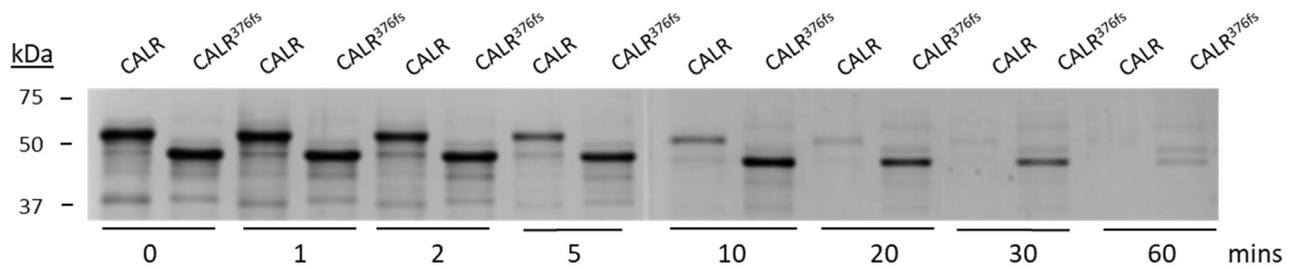
**Figure 15. Induction of UPR by CALR<sup>376fs</sup>**

Real-time Q-PCR analysis of spliced XBP1 (XBP1s) in HEK293 cells expressing CALR<sup>WT</sup> or CALR<sup>376fs</sup>, and treated with 0.5  $\mu$ M thapsigargin for 5 hrs. Values were normalized to glyceraldehyde 3-phosphate dehydrogenase (GAPDH). Error bars = s.d.; NS, not significant (ANOVA). Representative of 6 biological samples.

### **Protein folding of the CALR<sup>376fs</sup> mutant**

The CALR<sup>376fs</sup> mutant lacks part of the C-tail, bound Ca<sup>2+</sup> and PDIA3 with higher affinity than CALR<sup>WT</sup> *in vitro*, but did not affect Ca<sup>2+</sup> dynamics in cells, and showed low expression in cells due to degradation. The disrupted function of the mutant and apparent degradation could be due to disrupted structure. Therefore, as done before for the other CALR mutants, the folding of CALR<sup>376fs</sup> was examined by comparing its resistance to proteolytic digestion by trypsin to that of CALR<sup>WT</sup>.

Indeed, the CALR<sup>376fs</sup> mutant exhibited higher resistance to digestion by trypsin over time compared to CALR<sup>WT</sup> (Figure 16, top bands at around 45- and 55-kDa, respectively) indicating a different accessibility of the enzyme and different folding arrangement of the mutant protein.

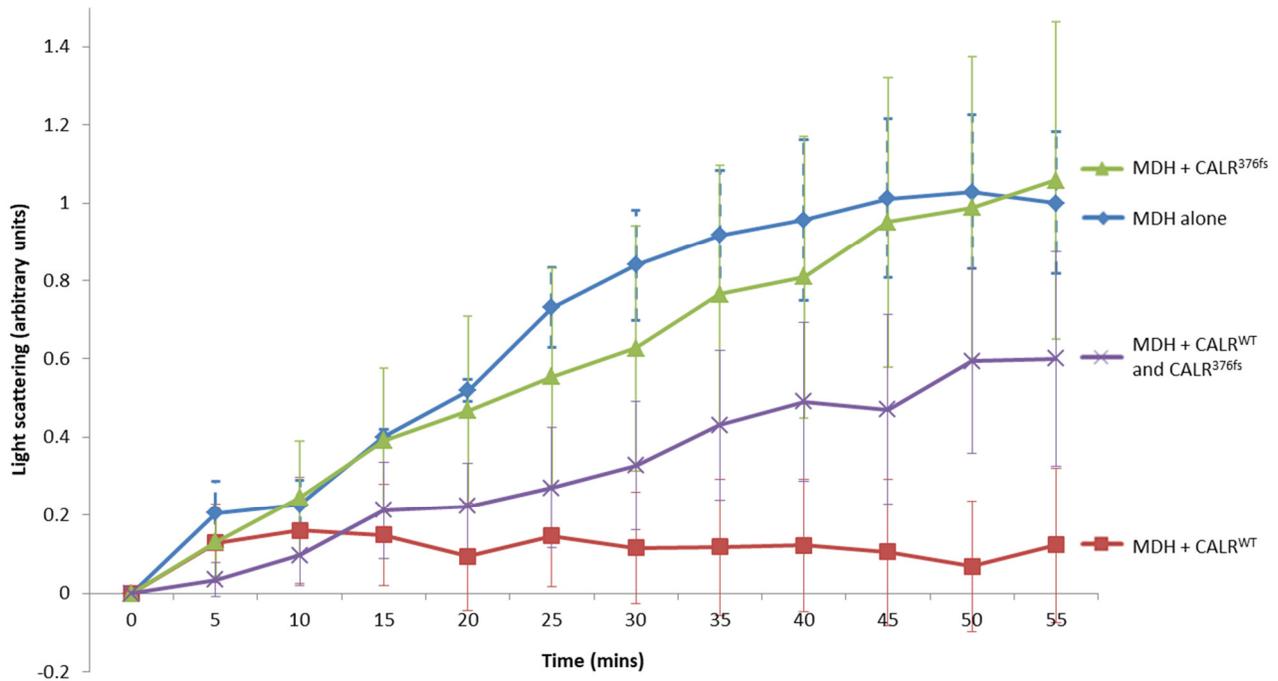


**Figure 16. Proteolytic digestion of the CALR<sup>376fs</sup> mutant**

Trypsin digestion analysis of CALR<sup>WT</sup> and CALR<sup>376fs</sup>. Purified proteins were incubated with trypsin at 1:100 (trypsin/protein; w/w) at 37° C. Aliquots were taken at the indicated time-points. The proteins were separated by SDS-PAGE and stained with Coomassie blue (top). Quantitative analysis of Coomassie blue staining showing the ratio of non-digested protein to initial abundance over time (bottom).

## Chaperone function of the CALR<sup>376fs</sup> mutant

CALR is a molecular chaperone and important component of the quality control system in the ER for folding, retention, and assembly of secretory and membrane proteins. However, CALR<sup>376fs</sup> showed to be folded differently compared to CALR<sup>WT</sup>. Therefore, it was investigated whether the disrupted structure of the CALR<sup>376fs</sup> mutant could be affecting its function as a molecular chaperone *in vitro*. To do so, the thermal aggregation of misfolded protein malate dehydrogenase (MDH) was analyzed by measuring increases in light scattering at 45°C in incubation with CALR<sup>WT</sup> or CALR<sup>376fs</sup> mutant for 1 hr. As shown by a decrease in light scattering of MDH, CALR<sup>WT</sup> inhibited thermal aggregation of MDH, while CALR<sup>376fs</sup> mutant did not (Figure 17). Moreover, incubation of MDH with CALR<sup>WT</sup> in combination with CALR<sup>376fs</sup> resulted in partial aggregation of MDH, showing a half decrease in inhibition by CALR<sup>WT</sup>.



**Figure 17. Aggregation suppression activity of CALR<sup>376fs</sup> on malate dehydrogenase (MDH)**

One  $\mu\text{M}$  MDH was incubated in presence or absence of  $0.2 \mu\text{M}$  CALR<sup>WT</sup>,  $0.2 \mu\text{M}$  CALR<sup>376fs</sup>, or  $0.1 \mu\text{M}$  CALR<sup>WT</sup> plus  $0.1 \mu\text{M}$  CALR<sup>376fs</sup>. Aggregation was recorded at  $45^\circ\text{C}$  by monitoring light scattering at  $360 \text{ nm}$ , and is shown as fraction of MDH alone at end point. Data points are shown every 5 mins. Error bars = s.d.  $n = 4$ .

# CHAPTER FIVE:

## General Discussion

---

## **CALR mutants localize to the ER**

Confocal fluorescence analysis of HeLa cells indicated that neither CALR<sup>WT</sup> nor mutants localized to the cell surface in intact cells, and instead displayed a typical ER pattern distribution with multiple contact sites surrounding the nucleus. Nevertheless, when the mutant CALR<sup>D302N</sup> was evaluated against an RFP-ER marker in HEK293 cells, it showed a small number of clusters that did not overlap with the marker when merging fluorescent images. It is a possibility that the amino acid change from aspartate (D) to asparagine (N) could insert a glycosylation site in the protein that led to higher degradation, and the reason it is seen as clusters. Alternatively, the heterogeneous nature of a vast organelle such as the ER is ideal for spatial and temporal specialized activities of proteins, therefore it is also possible that CALR could form microdomains that do not fully overlap with an ER marker. Several clusters are also evident in most mutants.

Similarly, HEK293 cells transfected with CALR<sup>376fs</sup> and CFP-ER marker indicated that the truncated mutant localized to the ER, despite loss of Golgi-to-ER retrieval KDEL signal. This suggests that CALR<sup>376fs</sup>, as well as wild-type protein, might possess an ER retention signal in its amino acid sequence, or be kept in the ER due to interactions with other proteins such as PDIA3, or with Ca<sup>2+</sup> [114].

## **CALR mutants bind Ca<sup>2+</sup> and PDIA3**

The CALR mutants showed similar ability to bind Ca<sup>2+</sup> compared to CALR<sup>WT</sup>, with small differences in binding affinities: some bound Ca<sup>2+</sup> tighter, some had looser binding. Surprisingly, CALR<sup>376fs</sup> exhibited slightly higher binding affinity to Ca<sup>2+</sup> ( $K_D = 469.94 \mu\text{M}$ ) than CALR<sup>WT</sup> ( $K_D = 772.75 \mu\text{M}$ ) despite lacking part of the high buffering C-tail, and having a predicted overall less negative charge. This is unlikely to be due to high affinity binding sites in the globular domain,

which have a  $\text{Ca}^{2+}$  affinity of  $<50 \mu\text{M}$  [48, 115], indicating that the mutant should have  $\text{Ca}^{2+}$  buffering ability in the ER.

It would be interesting to determine the total buffering capacity of  $\text{CALR}^{376\text{fs}}$  compared to that of  $\text{CALR}^{\text{WT}}$  (25 molecules of  $\text{Ca}^{2+}$ /mol of protein) [43]. Nevertheless, studies on a mouse *Calr* construct lacking part of the C-tail (1-362) have also shown higher binding affinity to  $\text{Ca}^{2+}$  compared to full-length protein, with  $K_D$  values of  $507 \mu\text{M}$  vs  $590 \mu\text{M}$ , respectively [48]. Surprisingly, the construct also had higher binding capacity than wild-type protein (5.9 moles of  $\text{Ca}^{2+}$ /mol of protein vs 3.8 moles of  $\text{Ca}^{2+}$ /mol of protein).

It was suggested that the N-terminal region of the acidic domain was sufficient to account for  $\text{Ca}^{2+}$  binding to low affinity sites, with 4–6 enthalpically driven low affinity sites measured within the C-domain of *Calr* (WT) as well as *Calr* (1–362). However, the differences observed in binding affinity and capacity to those previously reported [41, 115] might be because of different effects on folding of the proteins, due to different conditions in the techniques used (isothermal titration calorimetry vs equilibrium dialysis assay, enthalpy-driven vs entropy-driven binding). Similarly, the technique used in this study, microscale thermophoresis, might influence the folding and binding of mutant compared to wild-type protein. Regardless, this supports the importance of folding of  $\text{CALR}$  on  $\text{Ca}^{2+}$ -binding ability.

Most mutants also bound oxidoreductase PDIA3 with higher affinity than  $\text{CALR}^{\text{WT}}$ , with the possibility of this having effects on chaperone function and  $\text{Ca}^{2+}$  fluxes. It is also possible that  $\text{CALR}$  mutants could partially sequester  $\text{Ca}^{2+}$  and PDIA3 availability from  $\text{CALR}^{\text{WT}}$  activity.

**Cells expressing  $\text{CALR}^{376\text{fs}}$  have normal  $\text{Ca}^{2+}$  dynamics**

The CALR<sup>376fs</sup> mutant had part of its Ca<sup>2+</sup>-buffering C-tail deleted, but it showed stronger binding affinity *in vitro* to Ca<sup>2+</sup> and PDIA3, which could possibly affect Ca<sup>2+</sup> dynamics in cells. However, transient transfection of wild-type and *Calr*<sup>-/-</sup> cells with CALR<sup>WT</sup> or CALR<sup>376fs</sup> expression vectors showed very small and not significant differences in ER Ca<sup>2+</sup> release or SOCE when expressing CALR<sup>376fs</sup> compared to CALR<sup>WT</sup>.

The observed results might be affected by abundance of CALR<sup>376fs</sup>, or partial compensatory activity by endogenous CALR in wild-type cells or of other Ca<sup>2+</sup>-buffering proteins in *Calr*<sup>-/-</sup> cells, which could help regulate ER Ca<sup>2+</sup> content. For example, the molecular chaperone and Ca<sup>2+</sup>-binding protein PDIA3 displayed higher expression in *Calr*<sup>-/-</sup> cells. In any event, CALR<sup>376fs</sup> expression displayed normal Ca<sup>2+</sup> dynamics in wild-type and *Calr*<sup>-/-</sup> cells compared to CALR<sup>WT</sup>.

### **The CALR<sup>376fs</sup> mutant protein is highly degraded**

Consistent with the immunoblot of transfected wild-type and *Calr*<sup>-/-</sup> cells, CALR<sup>376fs</sup> seemed to be less expressed than CALR<sup>WT</sup> in HEK293 cells. Furthermore, consistent with fluorescence data, there was no increase of CALR<sup>376fs</sup> mutant expression with BFA treatment, indicating that the mutant was not being transported outside of the ER along the secretory pathway.

However, treatment with tunicamycin led to some increase in CALR<sup>376fs</sup> mutant abundance, while treatment with MG132 led to a very high increase in abundance of the mutant. MG132 treatment suggests that the CALR<sup>376fs</sup> mutant was mostly being degraded by the proteasome, possibly through ERAD mechanisms.

Intriguingly, treatment with tunicamycin, an N-glycosylation inhibitor and ER stress inducer, led to some increase in CALR<sup>376fs</sup> abundance, but this did not happen with BFA, an ER-

to-Golgi transport inhibitor that is also an ER stress inducer. The difference is surprising considering that endogenous *CALR* and other molecular chaperones are ER stress response genes, which could help delay degradation by retention of the exogenous *CALR*<sup>376fs</sup> mutant. This difference could be due to the type of activity of each drug and time requirement in leading to ER stress and UPR activation, with consequent *CALR* and chaperones gene response. The N-glycosylation blocking activity of tunicamycin could have a faster and greater immediate requirement for chaperone activity, whereas ER-to-Golgi transport blocking activity of BFA could be slower or have other preferential UPR requirements, such as ERAD.

Alternatively, *CALR*<sup>376fs</sup> might interact with the accumulated non-glycosylated misfolded proteins leading to increased abundance, and further contribute to ER stress by defective chaperone activity. Lastly, it is possible that tunicamycin could directly bind and stabilize the mutant *CALR*<sup>376fs</sup>, decreasing degradation.

### **The *CALR*<sup>376fs</sup> mutant does not induce higher UPR in cells**

As a Ca<sup>2+</sup>-buffer and molecular chaperone with a folding defect that is highly degraded, possibly affecting the quality control of the ER, it was hypothesized that *CALR*<sup>376fs</sup> expression in cells would lead to high UPR. Despite a partial increase in UPR compared to *CALR*<sup>WT</sup> expression, this was not statistically significant. The high degradation of a misfolded *CALR*<sup>376fs</sup>, and probably compensation of other ER chaperones, might help mitigate ER stress and UPR activation.

### ***CALR* mutants have a folding defect**

The higher resistance to trypsin digestion of *CALR* mutants compared to *CALR*<sup>WT</sup> suggest a different folding arrangement of the mutants, due to different accessibility of trypsin to cleaving sites on the protein. This is not surprising for *CALR*<sup>376fs</sup> considering that it is a shorter protein and

shows different mobility in SDS-PAGE. Moreover, the folding characteristics and stability of CALR have been previously reported to be affected by the C-domain and  $\text{Ca}^{2+}$  levels [46-48, 116].

It has been proposed that under physiological conditions the primary sequence of many proteins is not sufficient to determine a unique three-dimensional structure [117, 118]. Such proteins contain intrinsically disordered regions that provide a flexible character for a structure that can exist in dynamic equilibrium among multiple conformational states in response to environmental cues [119]. This appears to be a feature of several molecular chaperones [120-122].

Not surprising for a protein that has not been possible to be crystallized, structural and conformational computational analyses of CALR predict it to have several intrinsic disordered regions, indicating it is a highly disordered protein [119]. Some of these regions are able to undergo disorder-to-order transitions as result of binding, suggested to be crucial for recognition, regulation, and signaling function. Such regions of proteins are known as molecular recognition features (MoRFs).

Using an established specialized algorithm called ANCHOR, five MoRFs were identified in CALR: residues 170–176, 198–204, 259–269, 304–352, 409–417, located at N, P, and C-domains, and suggested as important for protein-protein interactions [119]. An alternative computational tool called MoRFPred supported this analysis of disorder-to-order regions by identifying several MoRFs in CALR: residues 1–10, 19–25, 50–55, 76–80, 199–206, 221–225, 236–242, 247–252, 312–314, 332–334, 378–389, and 409–417. While ANCHOR-based predicted residues 170-176 and 259-269 were missing from MoRFs, they showed significantly higher average scores than adjacent regions, suggesting increased binding potential.

These analyses contribute to the view of CALR as a dynamic, multifunctional, and highly promiscuous protein with structural flexibility that can interact with a wide range of binding partners by means of intrinsically disordered regions, identified as MoRFs.

Intriguingly, none of the whole set of mutants identified in SUD localize to this broad range of MoRFs. Only CALR<sup>E380G</sup> and CALR<sup>E381A</sup>, which were used for characterization studies, fall within the predicted residues by the alternative tool MoRFPred. This suggests that the mutants might not necessarily be involved in direct interaction with misfolded proteins or binding partners. However, it indicates that the mutated residues are located by structurally ordered regions of CALR. Disruptions of such ordered regions by amino acid changes might have striking consequences in the structure of the protein.

This would lead to folding defects as observed in the trypsin digestion analyses of all mutants and is likely to affect CALR function in part by altering the accessibility of disordered regions, as well as ability to bind interacting partners and become stabilized.

### **The CALR<sup>376fs</sup> mutant does not efficiently suppress aggregation of MDH**

The aggregation assay demonstrated that the CALR<sup>376fs</sup> mutant protein did not properly prevent aggregation of misfolded protein MDH or failed to be a substrate for proper folding or assembly of subunits for oligomer formation, suggesting an impaired chaperone function.

Moreover, incubation of MDH with CALR<sup>WT</sup> in combination with CALR<sup>376fs</sup> mutant seemed to affect the proper inhibition of aggregation of CALR<sup>WT</sup>. This could indicate that the mutant is indeed binding to the misfolded protein but does not aid in proper folding, while preventing access to CALR<sup>WT</sup>, or that the mutant is directly interfering with CALR<sup>WT</sup> and affecting its chaperone activity. Alternatively, the amount of functional chaperone, CALR<sup>WT</sup>, may just not

be enough to prevent aggregation when complemented by a defective chaperone, CALR<sup>376fs</sup>, representing heterozygous conditions.

### **Increased activity of L-type Ca<sup>2+</sup> channel in presence of the CALR<sup>376fs</sup> mutant**

The CALR<sup>376fs</sup> mutant was discovered in a victim of sudden unexplained death in the young (SUDY), which was due to a cardiac arrhythmia induced by long QT syndrome (LQTS). The LQTS has been proposed by Ackerman *et al.* to be due to an increased activity of L-type Ca<sup>2+</sup> ion channel (LTCC). Analysis of cardiac LTCC subtype Cav1.2 co-expressed with CALR<sup>WT</sup> in cells and in zebrafish has showed lower plasma membrane localization by immunoblot analysis and lower current activity by electrophysiological studies, compared to Cav1.2 higher plasma membrane localization and higher current activity when co-expressed in combination with CALR<sup>376fs</sup>, similar to Cav1.2 expressed alone (Ackerman *et al.*, personal communication). This suggests that CALR<sup>WT</sup> expression leads to a decrease in LTCC activity, while CALR<sup>376fs</sup> does not.

The impaired chaperone function of CALR<sup>376fs</sup> suggested by inability to prevent aggregation of MDH could explain the higher plasma membrane localization and activity of Cav1.2. While CALR<sup>WT</sup> had higher ER retention and quality control of Cav1.2, the defective mutant CALR<sup>376fs</sup> had lower ER retention and quality control in regulating Cav1.2 trafficking, leading to higher current.

### **Future studies**

The experiments conducted in CALR<sup>376fs</sup> should be similarly performed in the other CALR mutants (Table 6). Importantly, the aggregation assay should be performed with the mutants to evaluate a common defect in chaperone function. More work of chaperone function in cells, with also folding and trafficking of Cav1.2 and other plasma membrane proteins, including ion

channels, and gap junction proteins connexin 40 and connexin 43, are required to confirm the defect of CALR mutants in contributing to SUD. Furthermore, a mouse model of CALR<sup>376fs</sup> and CALR mutants, and cardiomyocytes analysis could shed more light on the LQTS-induced SUD phenotype and the contributing molecular mechanisms.

**Table 6. Available data of the CALR mutants**

Mutant	Cellular localization	Ca <sup>2+</sup> binding affinity	PDIA3 binding affinity	Trypsin digestion	Ca <sup>2+</sup> dynamics in cells	Processing in cells	UPR activation in cells	Aggregation suppression	Effect on LTCC activity
L11F									
Y57C									
P114R									
S189T									
P245L	x	x	x	x					
D295N	x	x	x	x					
D302N	x	x	x	x					
S323C	x	x	x	x					
K355 del	x	x	x	x					
376 fs	x	x	x	x	x	x	x	x	x
E380G	x	x	x	x					
E381A	x	x	x	x					
398-400 del									

## Conclusions

The ER contains a broad set of molecular chaperones and folding enzymes for quality control of its extensive protein folding needs, but the organization, coordination, distribution, and

specificity of such specialized proteins remains poorly characterized [123-125]. For instance, one specialized family of ER quality control, the CALR/CANX cycle, composed of highly homologous proteins CALR and CANX, can interact with a significantly different spectrum of proteins, or interact with specific subunits at different stages of folding process, or for different periods of time [72, 126-128]. These differences are suggested to be in part due to N-domain characteristics unique to CALR that confer specialized folding of specific substrates, and to differential accessibility of substrate components from luminal versus membrane-proximal sides.

The finding of CALR mutants in contributing to arrhythmia and SUD phenotype by increased activity of an LTCC might elucidate a specialized function of CALR in the heart. CALR has been previously implicated in impaired expression and trafficking of plasma membrane proteins [73, 74, 129-132], including that of neuroendocrine LTCC Cav1.3 [133]. Therefore, the folding and arrangement of cardiac LTCC Cav1.2 subunits might have lower ER quality control by a possibly important specific chaperone being CALR, with Cav1.2 having higher trafficking to the plasma membrane and increased current activity in presence of the mutant CALR<sup>376fs</sup>.

It is possible that endogenous Calr expression is downregulated in late embryonic stages, and protein abundance is low in adult mice hearts, in part to allow higher activity of LTCC in activation of sinus and AV nodes. Calr overexpression in adult mice hearts, and defective mutant CALR<sup>376fs</sup> in the present study have a similar SUD phenotype, indicating that tight control of CALR levels and activity in the heart are critical. A similar regulation of Calr levels and activity has been proposed in vomeronasal sensory neurons (VSNs) in mice. VSNs showed low abundance of Calr and instead high abundance of a VSN-specific Calr homolog, Calr4, which allows for traffic of vomeronasal type 2 pheromone receptor, since Calr was shown to affect its export [73, 129].

The involvement of the ER in the cardiac channelopathy LQTS is not well studied, however it has been demonstrated that ~90% of missense mutations of one of the main genes responsible for LQTS, *KCNH2* (causing LQT2) decrease the folding efficiency of K<sup>+</sup> channel Kv11.1 and increase its retention in the ER by quality control mechanisms [134]. Defects in ion channel trafficking, including ER retention, are considered one of the less understood but important contributing mechanisms leading to arrhythmias [32, 135].

Therefore, quality control in the ER is a fundamental cellular process that can lead to disruption of proteostasis and cause disease under unfavorable conditions, for example when either chaperone or substrate are mutated, as also suggested in the current project.

It is intriguing that Ca<sup>2+</sup> dynamics were not affected by the truncated CALR<sup>376fs</sup> mutant in possibly contributing to arrhythmia and SUD. However, this makes sense taking into consideration other CALR mutants identified in cases of SUD, which might not necessarily have a decrease in Ca<sup>2+</sup>-buffering activity, as was estimated for a truncated mutant that lacks part of the buffering C-tail. Moreover, the observed phenotype seems to be restricted to cardiac electric function in absence of structural abnormalities, while Ca<sup>2+</sup> homeostasis defects from a ubiquitous Ca<sup>2+</sup>-handling protein would probably lead to a wider range of phenotypes. A similar conclusion can be made from the observation of a not significant UPR activation. However, the possibility of a small continuous and persistent Ca<sup>2+</sup> imbalance and UPR activation inflicted by CALR<sup>376fs</sup> in a busy and sensitive cardiomyocyte environment, as also contributing to the SUD phenotype should not be eliminated.

In heterozygous conditions *in vivo*, it is conceivable that the disrupted CALR<sup>376fs</sup> mutant function effect be minimal because of degradation due to misfolding, and compensated for by

CALR<sup>WT</sup>. However, under high-demand for protein folding needs, such as in stressful conditions in the heart, CALR<sup>WT</sup> might not be enough to fully compensate the defect in chaperone function of CALR<sup>376fs</sup>, as observed in the *in vitro* assay, and lead to pathological high activity of LTCC due to impaired retention.

CALR was here presented as a novel protein implicated in SUD. Characterization of mutants identified in victims of SUD, with focus on analysis of truncated mutant CALR<sup>376fs</sup>, elucidated defects in folding, decreased abundance, and impaired chaperone activity as main factors that may contribute to the SUD phenotype. CALR<sup>376fs</sup> is specifically suggested to lead to a decrease in ER quality control, and to higher current activity of LTCC as mechanisms contributing to SUD. This study also implicates the ER and molecular chaperones as cellular components whose investigation could help decrease uncertainty around causes of channelopathies such as LQTS, and of mechanisms responsible for SUD.

## References

1. Pachon, M. and J. Almendral, *Sudden death: managing the patient who survives*. Heart, 2011. **97**(19): p. 1619-25.
2. Fernandez-Falgueras, A., et al., *Cardiac Channelopathies and Sudden Death: Recent Clinical and Genetic Advances*. Biology (Basel), 2017. **6**(1).
3. Black, M. and D.I. Graham, *Sudden unexplained death in adults caused by intracranial pathology*. J Clin Pathol, 2002. **55**(1): p. 44-50.
4. Magi, S., et al., *Sudden cardiac death: focus on the genetics of channelopathies and cardiomyopathies*. J Biomed Sci, 2017. **24**(1): p. 56.
5. Semsarian, C. and R.M. Hamilton, *Key role of the molecular autopsy in sudden unexpected death*. Heart Rhythm, 2012. **9**(1): p. 145-50.
6. Tester, D.J. and M.J. Ackerman, *Postmortem long QT syndrome genetic testing for sudden unexplained death in the young*. J Am Coll Cardiol, 2007. **49**(2): p. 240-6.
7. Semsarian, C., J. Ingles, and A.A. Wilde, *Sudden cardiac death in the young: the molecular autopsy and a practical approach to surviving relatives*. Eur Heart J, 2015. **36**(21): p. 1290-6.
8. Liberthson, R.R., *Sudden death from cardiac causes in children and young adults*. N Engl J Med, 1996. **334**(16): p. 1039-44.
9. Doolan, A., N. Langlois, and C. Semsarian, *Causes of sudden cardiac death in young Australians*. Med J Aust, 2004. **180**(3): p. 110-2.
10. Basso, C., et al., *Sudden cardiac death with normal heart: molecular autopsy*. Cardiovasc Pathol, 2010. **19**(6): p. 321-5.
11. Tester, D.J. and M.J. Ackerman, *The molecular autopsy: should the evaluation continue after the funeral?* Pediatr Cardiol, 2012. **33**(3): p. 461-70.

12. Cote, A., P. Russo, and J. Michaud, *Sudden unexpected deaths in infancy: what are the causes?* J Pediatr, 1999. **135**(4): p. 437-43.
13. Arnestad, M., A. Vege, and T.O. Rognum, *Evaluation of diagnostic tools applied in the examination of sudden unexpected deaths in infancy and early childhood.* Forensic Sci Int, 2002. **125**(2-3): p. 262-8.
14. Goldstein, R.D., et al., *Overall Postneonatal Mortality and Rates of SIDS.* Pediatrics, 2016. **137**(1).
15. Matthews, T.J., M.F. MacDorman, and M.E. Thoma, *Infant Mortality Statistics From the 2013 Period Linked Birth/Infant Death Data Set.* Natl Vital Stat Rep, 2015. **64**(9): p. 1-30.
16. Goldstein, R.D., H.C. Kinney, and M. Willinger, *Sudden Unexpected Death in Fetal Life Through Early Childhood.* Pediatrics, 2016. **137**(6).
17. Baruteau, A.E., et al., *Sudden infant death syndrome and inherited cardiac conditions.* Nat Rev Cardiol, 2017. **14**(12): p. 715-726.
18. Guntheroth, W.G. and P.S. Spiers, *The triple risk hypotheses in sudden infant death syndrome.* Pediatrics, 2002. **110**(5): p. e64.
19. Filiano, J.J. and H.C. Kinney, *A perspective on neuropathologic findings in victims of the sudden infant death syndrome: the triple-risk model.* Biol Neonate, 1994. **65**(3-4): p. 194-7.
20. Ackerman, M.J., et al., *Molecular diagnosis of the inherited long-QT syndrome in a woman who died after near-drowning.* N Engl J Med, 1999. **341**(15): p. 1121-5.
21. Priori, S.G., et al., *Brugada syndrome and sudden cardiac death in children.* Lancet, 2000. **355**(9206): p. 808-9.
22. Ackerman, M.J., D.J. Tester, and D.J. Driscoll, *Molecular autopsy of sudden unexplained death in the young.* Am J Forensic Med Pathol, 2001. **22**(2): p. 105-11.

23. Tester, D.J., et al., *Targeted mutational analysis of the RyR2-encoded cardiac ryanodine receptor in sudden unexplained death: a molecular autopsy of 49 medical examiner/coroner's cases*. Mayo Clin Proc, 2004. **79**(11): p. 1380-4.
24. Ackerman, M.J., et al., *Postmortem molecular analysis of SCN5A defects in sudden infant death syndrome*. JAMA, 2001. **286**(18): p. 2264-9.
25. Chugh, S.S., et al., *Postmortem molecular screening in unexplained sudden death*. J Am Coll Cardiol, 2004. **43**(9): p. 1625-9.
26. Ackerman, M.J., *Cardiac channelopathies: it's in the genes*. Nat Med, 2004. **10**(5): p. 463-4.
27. Lieve, K.V. and A.A. Wilde, *Inherited ion channel diseases: a brief review*. Europace, 2015. **17 Suppl 2**: p. ii1-6.
28. Schwartz, P.J., *The congenital long QT syndromes from genotype to phenotype: clinical implications*. J Intern Med, 2006. **259**(1): p. 39-47.
29. Schwartz, P.J., et al., *Prevalence of the congenital long-QT syndrome*. Circulation, 2009. **120**(18): p. 1761-7.
30. Kobza, R., et al., *Prevalence of long and short QT in a young population of 41,767 predominantly male Swiss conscripts*. Heart Rhythm, 2009. **6**(5): p. 652-7.
31. Funada, A., et al., *Assessment of QT intervals and prevalence of short QT syndrome in Japan*. Clin Cardiol, 2008. **31**(6): p. 270-4.
32. Delisle, B.P., et al., *Biology of cardiac arrhythmias: ion channel protein trafficking*. Circ Res, 2004. **94**(11): p. 1418-28.
33. Sanchez, O., et al., *Natural and Undetermined Sudden Death: Value of Post-Mortem Genetic Investigation*. PLoS One, 2016. **11**(12): p. e0167358.
34. Anderson, J.H., et al., *Whole-Exome Molecular Autopsy After Exertion-Related Sudden Unexplained Death in the Young*. Circ Cardiovasc Genet, 2016. **9**(3): p. 259-65.

35. Narula, N., et al., *Post-mortem Whole exome sequencing with gene-specific analysis for autopsy-negative sudden unexplained death in the young: a case series*. *Pediatr Cardiol*, 2015. **36**(4): p. 768-78.
36. Ackerman, M.J., *Genetic purgatory and the cardiac channelopathies: Exposing the variants of uncertain/unknown significance issue*. *Heart Rhythm*, 2015. **12**(11): p. 2325-31.
37. Wendelsdorf, K. and S. Shah, *Empowered genome community: leveraging a bioinformatics platform as a citizen-scientist collaboration tool*. *Appl Transl Genom*, 2015. **6**: p. 7-10.
38. Lek, M., et al., *Analysis of protein-coding genetic variation in 60,706 humans*. *Nature*, 2016. **536**(7616): p. 285-91.
39. Antzelevitch, C., et al., *Brugada syndrome: report of the second consensus conference*. *Heart Rhythm*, 2005. **2**(4): p. 429-40.
40. Petrovski, S., et al., *Genic intolerance to functional variation and the interpretation of personal genomes*. *PLoS Genet*, 2013. **9**(8): p. e1003709.
41. Michalak, M., et al., *Calreticulin: one protein, one gene, many functions*. *Biochem J*, 1999. **344 Pt 2**: p. 281-92.
42. McCauliffe, D.P., et al., *The 5'-flanking region of the human calreticulin gene shares homology with the human GRP78, GRP94, and protein disulfide isomerase promoters*. *J Biol Chem*, 1992. **267**(4): p. 2557-62.
43. Michalak, M., et al., *Calreticulin, a multi-process calcium-buffering chaperone of the endoplasmic reticulum*. *Biochem J*, 2009. **417**(3): p. 651-66.
44. Chouquet, A., et al., *X-ray structure of the human calreticulin globular domain reveals a peptide-binding area and suggests a multi-molecular mechanism*. *PLoS One*, 2011. **6**(3): p. e17886.
45. Nakamura, K., et al., *Functional specialization of calreticulin domains*. *J Cell Biol*, 2001. **154**(5): p. 961-72.

46. Corbett, E.F., et al., *The conformation of calreticulin is influenced by the endoplasmic reticulum luminal environment*. J Biol Chem, 2000. **275**(35): p. 27177-85.
47. Conte, I.L., et al., *The interplay between calcium and the in vitro lectin and chaperone activities of calreticulin*. Biochemistry, 2007. **46**(15): p. 4671-80.
48. Wijeyesakere, S.J., A.A. Gafni, and M. Raghavan, *Calreticulin is a thermostable protein with distinct structural responses to different divalent cation environments*. J Biol Chem, 2011. **286**(11): p. 8771-85.
49. Villamil Giraldo, A.M., et al., *The structure of calreticulin C-terminal domain is modulated by physiological variations of calcium concentration*. J Biol Chem, 2010. **285**(7): p. 4544-53.
50. Boelt, S.G., et al., *Mapping the Ca(2+) induced structural change in calreticulin*. J Proteomics, 2016. **142**: p. 138-48.
51. Gold, L.I., et al., *Calreticulin: non-endoplasmic reticulum functions in physiology and disease*. FASEB J, 2010. **24**(3): p. 665-83.
52. Baumann, O. and B. Walz, *Endoplasmic reticulum of animal cells and its organization into structural and functional domains*. Int Rev Cytol, 2001. **205**: p. 149-214.
53. English, A.R. and G.K. Voeltz, *Endoplasmic reticulum structure and interconnections with other organelles*. Cold Spring Harb Perspect Biol, 2013. **5**(4): p. a013227.
54. Rizzuto, R., M.R. Duchen, and T. Pozzan, *Flirting in little space: the ER/mitochondria Ca<sup>2+</sup> liaison*. Sci STKE, 2004. **2004**(215): p. re1.
55. Frischauf, I., et al., *The STIM/Orai coupling machinery*. Channels (Austin), 2008. **2**(4): p. 261-8.
56. Braakman, I. and N.J. Balleid, *Protein folding and modification in the mammalian endoplasmic reticulum*. Annu Rev Biochem, 2011. **80**: p. 71-99.
57. Papp, S., et al., *Is all of the endoplasmic reticulum created equal? The effects of the heterogeneous distribution of endoplasmic reticulum Ca<sup>2+</sup>-handling proteins*. J Cell Biol, 2003. **160**(4): p. 475-9.

58. Berridge, M.J., P. Lipp, and M.D. Bootman, *The versatility and universality of calcium signalling*. Nat Rev Mol Cell Biol, 2000. **1**(1): p. 11-21.
59. Meldolesi, J. and T. Pozzan, *The endoplasmic reticulum Ca<sup>2+</sup> store: a view from the lumen*. Trends Biochem Sci, 1998. **23**(1): p. 10-4.
60. Miyawaki, A., et al., *Fluorescent indicators for Ca<sup>2+</sup> based on green fluorescent proteins and calmodulin*. Nature, 1997. **388**(6645): p. 882-7.
61. Hogan, P.G., R.S. Lewis, and A. Rao, *Molecular basis of calcium signaling in lymphocytes: STIM and ORAI*. Annu Rev Immunol, 2010. **28**: p. 491-533.
62. Putney, J.W., Jr. and R.R. McKay, *Capacitative calcium entry channels*. Bioessays, 1999. **21**(1): p. 38-46.
63. John, L.M., J.D. Lechleiter, and P. Camacho, *Differential modulation of SERCA2 isoforms by calreticulin*. J Cell Biol, 1998. **142**(4): p. 963-73.
64. Kaufman, R.J., *Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls*. Genes Dev, 1999. **13**(10): p. 1211-33.
65. Hebert, D.N. and M. Molinari, *In and out of the ER: protein folding, quality control, degradation, and related human diseases*. Physiol Rev, 2007. **87**(4): p. 1377-408.
66. Groenendyk, J., et al., *Biology of endoplasmic reticulum stress in the heart*. Circ Res, 2010. **107**(10): p. 1185-97.
67. Mesaeli, N., et al., *Calreticulin is essential for cardiac development*. J Cell Biol, 1999. **144**(5): p. 857-68.
68. Luo, S., et al., *GRP78/BiP is required for cell proliferation and protecting the inner cell mass from apoptosis during early mouse embryonic development*. Mol Cell Biol, 2006. **26**(15): p. 5688-97.
69. Wanderling, S., et al., *GRP94 is essential for mesoderm induction and muscle development because it regulates insulin-like growth factor secretion*. Mol Biol Cell, 2007. **18**(10): p. 3764-75.

70. Coe, H., et al., *ERp57 modulates STAT3 signaling from the lumen of the endoplasmic reticulum*. J Biol Chem, 2010. **285**(9): p. 6725-38.
71. Williams, D.B., *Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum*. J Cell Sci, 2006. **119**(Pt 4): p. 615-23.
72. Molinari, M., et al., *Contrasting functions of calreticulin and calnexin in glycoprotein folding and ER quality control*. Mol Cell, 2004. **13**(1): p. 125-35.
73. Dey, S. and H. Matsunami, *Calreticulin chaperones regulate functional expression of vomeronasal type 2 pheromone receptors*. Proc Natl Acad Sci U S A, 2011. **108**(40): p. 16651-6.
74. Popescu, C.I., et al., *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, 2005. **280**(14): p. 13833-40.
75. Knee, R., et al., *Compromised calnexin function in calreticulin-deficient cells*. Biochem Biophys Res Commun, 2003. **304**(4): p. 661-6.
76. Rizvi, S.M., et al., *A polypeptide binding conformation of calreticulin is induced by heat shock, calcium depletion, or by deletion of the C-terminal acidic region*. Mol Cell, 2004. **15**(6): p. 913-23.
77. Coe, H. and M. Michalak, *ERp57, a multifunctional endoplasmic reticulum resident oxidoreductase*. Int J Biochem Cell Biol, 2010. **42**(6): p. 796-9.
78. Li, Y. and P. Camacho, *Ca<sup>2+</sup>-dependent redox modulation of SERCA 2b by ERp57*. J Cell Biol, 2004. **164**(1): p. 35-46.
79. Prins, D., et al., *Modulation of STIM1 and capacitative Ca<sup>2+</sup> entry by the endoplasmic reticulum luminal oxidoreductase ERp57*. EMBO Rep, 2011. **12**(11): p. 1182-8.
80. Groenendyk, J., L.B. Agellon, and M. Michalak, *Coping with endoplasmic reticulum stress in the cardiovascular system*. Annu Rev Physiol, 2013. **75**: p. 49-67.

81. Hetz, C., E. Chevet, and S.A. Oakes, *Proteostasis control by the unfolded protein response*. Nat Cell Biol, 2015. **17**(7): p. 829-38.
82. Eggleton, P., et al., *Calreticulin, a therapeutic target?* Expert Opin Ther Targets, 2016. **20**(9): p. 1137-47.
83. Wang, W.A., J. Groenendyk, and M. Michalak, *Calreticulin signaling in health and disease*. Int J Biochem Cell Biol, 2012. **44**(6): p. 842-6.
84. Corbett, E.F. and M. Michalak, *Calcium, a signaling molecule in the endoplasmic reticulum?* Trends Biochem Sci, 2000. **25**(7): p. 307-11.
85. Wiersma, V.R., et al., *Mechanisms of Translocation of ER Chaperones to the Cell Surface and Immunomodulatory Roles in Cancer and Autoimmunity*. Front Oncol, 2015. **5**: p. 7.
86. Guo, L., et al., *Cardiac-specific expression of calcineurin reverses embryonic lethality in calreticulin-deficient mouse*. J Biol Chem, 2002. **277**(52): p. 50776-9.
87. Nakamura, K., et al., *Complete heart block and sudden death in mice overexpressing calreticulin*. J Clin Invest, 2001. **107**(10): p. 1245-53.
88. Michalak, M., et al., *Calreticulin in the heart*. Mol Cell Biochem, 2004. **263**(1): p. 137-42.
89. Guo, L., et al., *COUP-TF1 antagonizes Nkx2.5-mediated activation of the calreticulin gene during cardiac development*. J Biol Chem, 2001. **276**(4): p. 2797-801.
90. Beard, N.A., D.R. Laver, and A.F. Dulhunty, *Calsequestrin and the calcium release channel of skeletal and cardiac muscle*. Prog Biophys Mol Biol, 2004. **85**(1): p. 33-69.
91. Michalak, M. and M. Opas, *Endoplasmic and sarcoplasmic reticulum in the heart*. Trends Cell Biol, 2009. **19**(6): p. 253-9.
92. Klampfl, T., et al., *Somatic mutations of calreticulin in myeloproliferative neoplasms*. N Engl J Med, 2013. **369**(25): p. 2379-90.

93. Nangalia, J., et al., *Somatic CALR mutations in myeloproliferative neoplasms with nonmutated JAK2*. *N Engl J Med*, 2013. **369**(25): p. 2391-2405.
94. Marty, C., et al., *Calreticulin mutants in mice induce an MPL-dependent thrombocytosis with frequent progression to myelofibrosis*. *Blood*, 2016. **127**(10): p. 1317-24.
95. Chachoua, I., et al., *Thrombopoietin receptor activation by myeloproliferative neoplasm associated calreticulin mutants*. *Blood*, 2016. **127**(10): p. 1325-35.
96. Araki, M., et al., *Activation of the thrombopoietin receptor by mutant calreticulin in CALR-mutant myeloproliferative neoplasms*. *Blood*, 2016. **127**(10): p. 1307-16.
97. Martin, V., et al., *Identification by mutational analysis of amino acid residues essential in the chaperone function of calreticulin*. *J Biol Chem*, 2006. **281**(4): p. 2338-46.
98. Guo, L., et al., *Identification of an N-domain histidine essential for chaperone function in calreticulin*. *J Biol Chem*, 2003. **278**(50): p. 50645-53.
99. Saito, Y., et al., *Calreticulin functions in vitro as a molecular chaperone for both glycosylated and non-glycosylated proteins*. *EMBO J*, 1999. **18**(23): p. 6718-29.
100. Ciplys, E., et al., *High-level secretion of native recombinant human calreticulin in yeast*. *Microb Cell Fact*, 2015. **14**: p. 165.
101. Ciplys, E., E. Zitkus, and R. Slibinskas, *Native signal peptide of human ERp57 disulfide isomerase mediates secretion of active native recombinant ERp57 protein in yeast *Saccharomyces cerevisiae**. *Protein Expr Purif*, 2013. **89**(2): p. 131-5.
102. Jerabek-Willemsen, M., et al., *Molecular interaction studies using microscale thermophoresis*. *Assay Drug Dev Technol*, 2011. **9**(4): p. 342-53.
103. Pocanschi, C.L., et al., *Structural and functional relationships between the lectin and arm domains of calreticulin*. *J Biol Chem*, 2011. **286**(31): p. 27266-77.

104. Ellgaard, L., et al., *Three-dimensional structure topology of the calreticulin P-domain based on NMR assignment*. FEBS Lett, 2001. **488**(1-2): p. 69-73.
105. Norgaard Toft, K., et al., *Small angle X-ray scattering study of calreticulin reveals conformational plasticity*. Biochim Biophys Acta, 2008. **1784**(9): p. 1265-70.
106. Qiu, Y. and M. Michalak, *Transcriptional control of the calreticulin gene in health and disease*. Int J Biochem Cell Biol, 2009. **41**(3): p. 531-8.
107. Anfinsen, C.B., et al., *The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain*. Proc Natl Acad Sci U S A, 1961. **47**: p. 1309-14.
108. Hartl, F.U., A. Bracher, and M. Hayer-Hartl, *Molecular chaperones in protein folding and proteostasis*. Nature, 2011. **475**(7356): p. 324-32.
109. Saibil, H., *Chaperone machines for protein folding, unfolding and disaggregation*. Nat Rev Mol Cell Biol, 2013. **14**(10): p. 630-42.
110. Kim, Y.E., et al., *Molecular chaperone functions in protein folding and proteostasis*. Annu Rev Biochem, 2013. **82**: p. 323-55.
111. Webster, J. and D. Oxley, *Protein identification by MALDI-TOF mass spectrometry*. Methods Mol Biol, 2012. **800**: p. 227-40.
112. Falkner, J.A., et al., *Validated MALDI-TOF/TOF mass spectra for protein standards*. J Am Soc Mass Spectrom, 2007. **18**(5): p. 850-5.
113. Han, L., et al., *Calreticulin-mutant proteins induce megakaryocytic signaling to transform hematopoietic cells and undergo accelerated degradation and Golgi-mediated secretion*. J Hematol Oncol, 2016. **9**(1): p. 45.
114. Sonnichsen, B., et al., *Retention and retrieval: both mechanisms cooperate to maintain calreticulin in the endoplasmic reticulum*. J Cell Sci, 1994. **107 ( Pt 10)**: p. 2705-17.

115. Baksh, S. and M. Michalak, *Expression of calreticulin in Escherichia coli and identification of its Ca<sup>2+</sup> binding domains*. J Biol Chem, 1991. **266**(32): p. 21458-65.
116. Li, Z., W.F. Stafford, and M. Bouvier, *The metal ion binding properties of calreticulin modulate its conformational flexibility and thermal stability*. Biochemistry, 2001. **40**(37): p. 11193-201.
117. Dunker, A.K., et al., *Intrinsically disordered protein*. J Mol Graph Model, 2001. **19**(1): p. 26-59.
118. Uversky, V.N. and A.K. Dunker, *Multiparametric analysis of intrinsically disordered proteins: looking at intrinsic disorder through compound eyes*. Anal Chem, 2012. **84**(5): p. 2096-104.
119. Varricchio, L., et al., *Calreticulin: Challenges Posed by the Intrinsically Disordered Nature of Calreticulin to the Study of Its Function*. Front Cell Dev Biol, 2017. **5**: p. 96.
120. Tompa, P. and P. Csermely, *The role of structural disorder in the function of RNA and protein chaperones*. FASEB J, 2004. **18**(11): p. 1169-75.
121. Tompa, P. and D. Kovacs, *Intrinsically disordered chaperones in plants and animals*. Biochem Cell Biol, 2010. **88**(2): p. 167-74.
122. Kovacs, D. and P. Tompa, *Diverse functional manifestations of intrinsic structural disorder in molecular chaperones*. Biochem Soc Trans, 2012. **40**(5): p. 963-8.
123. Gidalevitz, T., F. Stevens, and Y. Argon, *Orchestration of secretory protein folding by ER chaperones*. Biochim Biophys Acta, 2013. **1833**(11): p. 2410-24.
124. Koldewey, P., S. Horowitz, and J.C.A. Bardwell, *Chaperone-client interactions: Non-specificity engenders multifunctionality*. J Biol Chem, 2017. **292**(29): p. 12010-12017.
125. Bose, D. and A. Chakrabarti, *Substrate specificity in the context of molecular chaperones*. IUBMB Life, 2017. **69**(9): p. 647-659.
126. Danilczyk, U.G. and D.B. Williams, *The lectin chaperone calnexin utilizes polypeptide-based interactions to associate with many of its substrates in vivo*. J Biol Chem, 2001. **276**(27): p. 25532-40.

127. *Calnexin, calreticulin and the folding of glycoproteins*. Trends Cell Biol, 1997. **7**(5): p. 193-200.
128. Nauseef, W.M., S.J. McCormick, and M. Goedken, *Coordinated participation of calreticulin and calnexin in the biosynthesis of myeloperoxidase*. J Biol Chem, 1998. **273**(12): p. 7107-11.
129. Jiang, Y., S. Dey, and H. Matsunami, *Calreticulin: roles in cell-surface protein expression*. Membranes (Basel), 2014. **4**(3): p. 630-41.
130. Zhang, L., et al., *Calreticulin promotes folding/dimerization of human lipoprotein lipase expressed in insect cells (sf21)*. J Biol Chem, 2003. **278**(31): p. 29344-51.
131. Theocharides, A.P., et al., *Homozygous calreticulin mutations in patients with myelofibrosis lead to acquired myeloperoxidase deficiency*. Blood, 2016. **127**(25): p. 3253-9.
132. Cazzola, M., *Mutant calreticulin: when a chaperone becomes intrusive*. Blood, 2016. **127**(10): p. 1219-21.
133. Karnabi, E., et al., *Calreticulin negatively regulates the surface expression of Cav1.3 L-type calcium channel*. Biochem Biophys Res Commun, 2013. **437**(4): p. 497-501.
134. Smith, J.L., et al., *Molecular pathogenesis of long QT syndrome type 2*. J Arrhythm, 2016. **32**(5): p. 373-380.
135. Harkcom, W.T. and G.W. Abbott, *Emerging concepts in the pharmacogenomics of arrhythmias: ion channel trafficking*. Expert Rev Cardiovasc Ther, 2010. **8**(8): p. 1161-73.