

**IN SILICO ANALYSIS OF CODING SINGLE-NUCLEOTIDE  
POLYMORPHISM (SNP)'S IN GLUCOCEREBROSIDASE (GBA)  
GENE AND THEIR IMPACT ON THE BIOLOGICAL FUNCTIONS  
OF THE CELL**

by  
VEYSEL OĞULCAN KAYA

Submitted to the Graduate School of Engineering and Natural Sciences  
in partial fulfilment of  
the requirements for the degree of Master of Science

Sabancı University  
August 2020

**IN SILICO ANALYSIS OF CODING SINGLE-NUCLEOTIDE POLYMORPHISM (SNP)'S IN  
GLUCOCEREBROSIDASE (GBA) GENE AND THEIR IMPACT ON THE BIOLOGICAL  
FUNCTIONS OF THE CELL**

APPROVED BY:

Assoc. Prof. Özlem Kutlu  
(Thesis Supervisor)

A blue ink signature of Assoc. Prof. Özlem Kutlu, written over a horizontal dotted line.

Asst. Prof. Emre Taşkın

A blue ink signature of Asst. Prof. Emre Taşkın, written over a horizontal dotted line.

Asst. Prof. Öznur Bayraktar

A blue ink signature of Asst. Prof. Öznur Bayraktar, written over a horizontal dotted line.

DATE OF APPROVAL: 14.08.2020



THESIS AUTHOR 2020 ©

All Rights Reserved

## ABSTRACT

### IN SILICO ANALYSIS OF CODING SINGLE-NUCLEOTIDE POLYMORPHISM (SNP)'S IN GLUCOCEREBROSIDASE (GBA) GENE AND THEIR IMPACT ON THE BIOLOGICAL FUNCTIONS OF THE CELL

VEYSEL OĞULCAN KAYA

Molecular Biology, Genetics and Bioengineering M.S. THESIS, AUGUST 2020

Thesis Supervisor: Assoc. Prof. Özlem Kutlu

Keywords: Gaucher Disease, GBA, SNP, glucocerebrosidase, Lysosomal  
Re-acidification, V-ATPases

Gaucher disease (GD, ORPHA355) is a rare, autosomal recessive genetic disorder. It is caused by an insufficiency of the lysosomal enzyme, glucocerebrosidase (GCase), due to the severity of GBA1 gene's mutations. Such conditions lead to GD through massive accumulation of GCase substrate, glucosylceramide, in the lysosomes. Accordingly, this research analyzed the most frequent Gaucher Disease-linked single nucleotide polymorphisms (SNPs) on the GBA1 gene by applying various bioinformatics algorithms. We have classified and characterized the L296V mutation in the GD-linked deleterious SNPs spectrum. We also showed an increase in in-vitro enzymatic activity after lysosomal reacidification and evaluated the mutant GCases' reactions to constant pH experiments. Our results showed that L296V, N370S, L444P, and D409H variants are harmful mutations with different levels of hydrolysis disturbance, tertiary structure unbalance, activation interruption, and transportation of the GCase protein as well as enzymatic efficiency reduction. Yet, we have seen lysosomes' reacidification by over-expressing hydrogen pumps (V-ATPases), which reverted the decreased stability and increased enzymatic activity. These findings may suggest a research background for the therapeutic applications of Gaucher Disease through lysosomal re-acidification, and can be extended for future studies.

## ÖZET

### GLUKOSEREBROSİDAZ (GBA) GENİNDEKİ KODLANAN TEK NUKLEOTİD POLİMORFİZMLERİNİN (SNP) IN SİLİKO YONTEMLER İLE ANALİZİ VE HUCRENİN BİYOLOJİK FONKSİYONLARI ÜZERİNDEKİ ETKİSİ

VEYSEL OĞULCAN KAYA

Moleküler Biyoloji, Genetik ve Biyomühendislik YÜKSEK LİSANS TEZİ,  
AGUSTOS 2020

Tez Danışmanı: Doç. Dr. Özlem Kutlu

Anahtar Kelimeler: Gaucher Hastalığı, GBA, Tek Nükleotid Polimorfizmleri,  
V-ATPase, Glukoserebrosidaz

Gaucher hastalığı (GH, ORPHA355) nadir görülen, otozomal resesif geçişli bir genetik hastalıktır, ve GBA1 gen mutasyonlarının etkisi sonucu lizozomal enzim olan glukoserebrosidazın (GCase) yetersizliğinden kaynaklanmaktadır. Bu tür koşullar, lizozomlarda, GCase substratı olan glukosilseramidin büyük miktarda birikimi yoluyla GH'ye yol açmaktadır. Bu araştırmada, çeşitli biyoinformatik algoritmalarını uygulayarak GBA1 genindeki, Gaucher Hastalığına bağlı en sık görülen tek nükleotid polimorfizmleri (SNP'ler) analiz edilmiştir. Ayrıca, GH bağlantılı zararlı SNP'ler spektrumu içerisinde L296V mutasyonunu sınıflandırılmıştır ve karakterize edilmiştir. Bununla beraber, lizozomların yeniden asitleştirilmesinden sonra in vitro enzimatik aktivitede bir artış gösterilmiştir ve mutant GCase'lerin sabit pH deneylerine tepkileri değerlendirilmiştir. Araştırmalarımız sonucu, L296V, N370S, L444P ve D409H varyantlarının, farklı seviyelerde hidroliz bozukluğu, üçüncül yapı dengesizliği, aktivasyon yetersizliği ve GCase proteininin taşınmasının yanı sıra enzimatik verimlilik azalması gibi farklı niteliklerdeki etkilere yol açtığı görülmüştür. Fakat, ekspresyon seviyesi arttırılmış hidrojen pompaları (V-ATPases) yoluyla lizozomların yeniden asitleştirilmesi sonucu azalmış stabilitenin ve enzimatik aktivitenin geri kazanıldığı görülmüştür. Bu bulgular, Gaucher Hastalığının lizozomal yeniden asitleştirme yoluyla terapötik uygulamaları için bir araştırma arka planı önermektedir ve gelecekteki çalışmalar için genişletilebilecek potansiyel taşımaktadır.

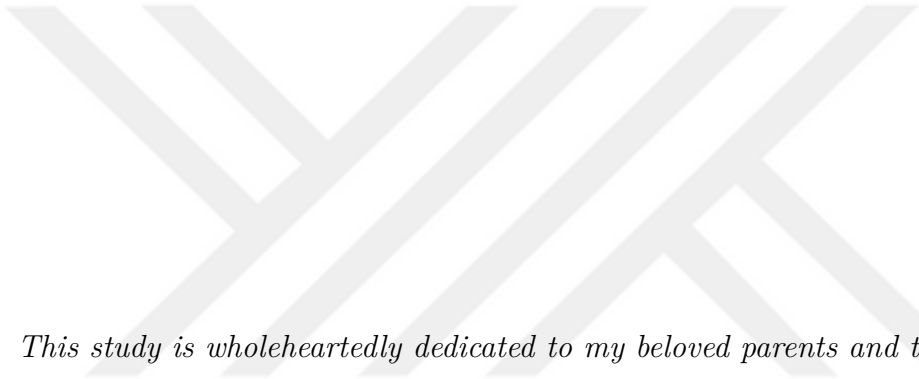
## ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisor, Özlem Kutlu, for the continuous support of my MSc study and related research, for her patience, motivation, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis.

I want to express my most profound thankfulness to my parents and my girlfriend. Without their love, support, and understanding over the years, none of this would have been possible. They have always been there for me, and I am thankful for everything they have helped me achieve.

My gratitude also extends to my laboratory colleagues. I want to thank them for their great understanding, patience, support, and joyful moments that we shared in our most intense times.

This project is supported by TUBITAK under the Project No: 112T130. I thank Biorender.com for assisting me in drawing molecular pathways' figures with their user-friendly editor.



*This study is wholeheartedly dedicated to my beloved parents and to my girlfriend,  
for all her love and support*

## TABLE OF CONTENTS

|                                                                        |           |
|------------------------------------------------------------------------|-----------|
| <b>LIST OF TABLES .....</b>                                            | <b>x</b>  |
| <b>LIST OF FIGURES .....</b>                                           | <b>xi</b> |
| <b>1. INTRODUCTION.....</b>                                            | <b>1</b>  |
| 1.1. Lysosomal Storage Diseases.....                                   | 2         |
| 1.1.1. Sphingolipidoses.....                                           | 3         |
| 1.2. Gaucher Disease .....                                             | 4         |
| 1.2.1. Pathophysiology of Gaucher Disease .....                        | 5         |
| 1.2.2. Therapeutic Approaches .....                                    | 10        |
| 1.2.2.1. Enzyme Replacement Therapy .....                              | 10        |
| 1.2.2.2. Substrate Reduction Therapy .....                             | 10        |
| 1.2.2.3. Gene Therapy .....                                            | 12        |
| 1.3. Lysosomal Biogenesis .....                                        | 12        |
| 1.3.1. Lysosomal pH Regulation .....                                   | 13        |
| 1.3.2. V-ATPase Regulation .....                                       | 14        |
| 1.4. Aim of the Thesis .....                                           | 15        |
| <b>2. MATERIALS and METHODS .....</b>                                  | <b>16</b> |
| 2.1. Wet-Lab Analysis .....                                            | 16        |
| 2.1.1. Primary Cell Culture.....                                       | 16        |
| 2.1.2. Gene Transfection .....                                         | 18        |
| 2.1.3. Lysosomal pH Measurement Assay.....                             | 18        |
| 2.1.4. B-glucocerebrosidase Enzyme Activity Assay .....                | 19        |
| 2.1.5. In-direct Immunofluorescence Assay and Confocal Microscopy      | 19        |
| 2.1.6. Site-directed Mutagenesis of L444P .....                        | 19        |
| 2.1.7. Statistical analyses .....                                      | 21        |
| 2.2. Dry-Lab Analysis .....                                            | 21        |
| 2.2.1. Sequence recovery .....                                         | 21        |
| 2.2.2. Analysis of SNPs in human GBA .....                             | 21        |
| 2.2.3. Identification of mutant SNPs position in different domains ... | 22        |

|           |                                                                                                  |           |
|-----------|--------------------------------------------------------------------------------------------------|-----------|
| 2.2.4.    | Evolutionary phylogenetic analysis of human GBA protein....                                      | 22        |
| 2.2.5.    | Extrapolation of amino acid changes and disease phenotypes..                                     | 22        |
| 2.2.6.    | Stability analysis of human GBA protein .....                                                    | 23        |
| 2.2.7.    | Prediction of post translational modification sites for GBA ...                                  | 23        |
| 2.2.8.    | Prediction of structural effect of point mutation on human<br>GBA protein .....                  | 23        |
| 2.2.9.    | Prediction of secondary structure .....                                                          | 23        |
| 2.2.10.   | Protein-protein interactions prediction .....                                                    | 24        |
| <b>3.</b> | <b>RESULTS .....</b>                                                                             | <b>25</b> |
| 3.1.      | SNP dataset consists of most frequent Gaucher Disease linked GBA<br>mutations .....              | 25        |
| 3.2.      | Analysis of SNPs in human GBA .....                                                              | 27        |
| 3.3.      | Identification of domains in GBA protein .....                                                   | 29        |
| 3.4.      | High-risk SNPs show variable conservation profile.....                                           | 31        |
| 3.5.      | Analysis of structural effects of high-risk SNPs in GBA .....                                    | 33        |
| 3.6.      | Identification of disease phenotype relation with our SNP dataset in<br>human GBA .....          | 35        |
| 3.7.      | Prediction of post translational modification sites for human GBA<br>protein .....               | 36        |
| 3.8.      | Effects of mutations on protein stability and protein-protein interac-<br>tion of human GBA..... | 36        |
| 3.9.      | Lysosomal re-acidification increases GBA enzymatic activity in-vitro .                           | 38        |
| 3.10.     | Lysosomal re-acidification alters GBA sub-localization .....                                     | 41        |
| <b>4.</b> | <b>DISCUSSION .....</b>                                                                          | <b>44</b> |

## LIST OF TABLES

|                                                                                                                                                                             |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Table 2.1. The relation between SNPs and the symptoms of GD-patients.                                                                                                       | 17 |
| Table 2.2. Site-directed mutagenesis reaction mixture.....                                                                                                                  | 20 |
| Table 2.3. Site-directed mutagenesis reaction thermal cycler protocol.....                                                                                                  | 20 |
| Table 3.1. Functional consequence types of the most common SNPs<br>present in the human GBA gene principal isoforms.....                                                    | 26 |
| Table 3.2. L444P, N370S, L296V, and D409H are the pathogenic amino<br>acid substitutions with several dangerous consequences on structure,<br>function, and stability. .... | 28 |
| Table 3.3. Results from the Analysis of human GBA protein by Project<br>HOPE reveals the structural effects of high-risk SNPs in GBA.....                                   | 34 |
| Table 3.4. Results from the Analysis of our SNP dataset by MutPred<br>enlights potential reasons for protein instability. ....                                              | 35 |
| Table 3.5. Results from the Analysis of our SNP dataset with I-Mutant<br>shows a general decrease with introduced variants.....                                             | 37 |
| Table 3.6. Normalization of the changes in the enzyme activity suits the<br>results for appropriate interpretation between SNP groups. ....                                 | 40 |

## LIST OF FIGURES

|                                                                                                                                                                  |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1.1. Alternative allele frequencies of the most frequent GBA variants. ....                                                                               | 4  |
| Figure 1.2. Schematic of the trafficking of glucocerebrosidase presence in the cell. ....                                                                        | 6  |
| Figure 1.3. Saposin C activates glucocerebrosidase (GCCase) for the hydrolysis of glucosylceramides (GlcCer). ....                                               | 8  |
| Figure 3.1. InterPro analysis revealed the GBA protein domains belonging to Glycoside hydrolase family 30. ....                                                  | 29 |
| Figure 3.2. Secondary structure of GCCase protein. ....                                                                                                          | 30 |
| Figure 3.3. Results from the Analysis of human GBA protein by ConSurf reports the divergence in the conservation spectrum of our SNP dataset. ....               | 32 |
| Figure 3.4. Protein–protein interaction network of GBA protein shown by STRING. ....                                                                             | 37 |
| Figure 3.5. Lysosomal re-acidification increases GCCase enzyme activity. ....                                                                                    | 39 |
| Figure 3.6. Confocal microscopy results of patient derived Gaucher fibroblasts reveals the change of GBA sub-localization after lysosomal re-acidification. .... | 42 |

## 1. INTRODUCTION

Gaucher Disease (GD) is the most prevalent lysosomal storage disease, mainly arise from the insufficiency of one lysosomal enzyme, glucocerebrosidase. Glucosylceramidase beta (GBA) gene encodes the lysosomal enzyme, producing a 497 amino acid long membrane glycoprotein, with a molecular weight of 65 kDa. The enzyme functions as a catalyst for glucosylceramide (GlcCer) hydrolysis into glucose and ceramide with the activation of saposin C. In The Human Gene Mutation Database [1], there are more than 400 mutations cataloged to the GBA gene which spans 1q21 on the human genome. GD linked variants, and their clinical types are constituted based on neurological involvement, namely: non-neuronopathic groups and neuronopathic groups. [2] GD linked GBA variants with genotype-phenotype correspondences have been established in the literature. [3] For instance, N370S is a common variant and associated with Type 1, or the L444P amino acid substitution links with neuronopathic forms of GD. [4] Furthermore, homozygous D409H variant is affiliated with a particular Type 3 exhibiting critical cardiac involvement and oculomotor apraxia. [5] Expression analyses of GBA mutants have been conducted in vitro in order to interpret the effect of the mutation on enzymatic activity and to investigate phenotype-genotype correlations. [6]

In this research, we have analyzed the most frequent Gaucher Disease-linked single nucleotide polymorphisms (SNPs) located in the GBA gene by applying various bioinformatics algorithms. We have classified and characterized the deleterious SNPs and evaluated the impact of the mutation with constant pH experiments. In addition, we showed an increase at in-vitro enzymatic activity after lysosomal reacidification.

Our results show that L296V [46], N370S, L444P, and D409H variants are harmful mutations with different levels of hydrolysis disturbance, tertiary structure unbalance, activation interruption, and transportation of the GBA protein as well as enzymatic efficiency reduction. Yet, elaborating on re-acidification of lysosomes with over-expressed hydrogen pumps reverts the decreased stability and increases enzymatic activity. These findings may suggest a therapeutic application for Gaucher

Disease and can be extended for future studies.

## 1.1 Lysosomal Storage Diseases

Lysosomal Storage Diseases (LSDs) result from heritable errors in metabolism, change the lysosome's function. Related to lysosomal catabolism, LSDs consist of 70 disorders, which are mostly caused by inborn errors linked to autosomal recessive traits. Mutations in the genes encoding the proteins incorporated in the lysosomal catabolism invoke lysosomal failure and the continuous increase of the substrates inside the lysosome. Such instances will eventually bring cell dysfunction and cell death.

Patients suffering from LSDs, present a diverse range of a spectrum, mostly represent a low representation with the driving mutation and dysfunctional proteins. Yet, the patients are generally categorized with the type of the disease and the age of the initial disease-related symptoms. Infantile patients, frequently, show the symptoms in the most severe conditions, compared to adult types. Diagnosis is mostly based on genetic screening through sequencing and enzymatic analysis. [7] With the considerably recent advancements in sequencing technology, whole-exome sequencing is a wide-spread application in the clinics. Yet, LSD patients with moderate symptoms are doubtful for specific diagnoses due to showing diverse signs of different diseases.

The sphingolipidoses is a genetically heterogeneous type of LSDs, and present function disorder affecting enzyme degradation capabilities of the metabolites, as critical components of cell membranes and regulators of numerous downstream pathways.

### 1.1.1 Sphingolipidoses

Sphingolipids are commonly located in the plasma membranes of eukaryotic cells.[8] A membrane sphingolipid has a common form of a two-tailed lipid consisted of a ceramide membrane anchor, and it is either linked to a phosphorylcholine complex to present sphingomyelin or attached to an oligosaccharide to present glycosphingolipids (GSLs). Sphingosine and sphingosine-1-phosphate are the one-tailed sphingolipids and consist of a lesser complex form while degrading or synthesizing sphingolipids. Besides sphingosine and sphingosine-1-phosphate, ceramide is also found in the lipid signaling mechanism with respect to intracellular regulation and physiological functions. The continuous lysosomal degradation of sphingolipids begins from the plasma membrane. In order to maintain cellular equilibrium, internalized membrane partitions are digested with various lipases and proteases in an acidic endo-lysosomal environment. [9]

Sphingolipidoses or lysosomal sphingolipid storage disorders result from a malfunction in sphingolipid degradation arising with genetic fault for involved enzymes or required cofactors. Massive accumulation of an intermediate belonging to sphingolipid degradation in the lysosomes will eventually lead to cellular decay. In the literature, the accumulation of each intermediate has been pointed to various sphingolipid storage diseases. Moreover, each variation shapes a different characteristic of the disease by altering the form of the intermediates and presenting a harsh form of sphingolipidose.

Unfortunately, the severe forms of sphingolipidoses lead to early death through neuronopathy. Patients with milder conditions are not characterized by neuronopathy and can expect longer lives. The diseases identified with milder forms of accumulated intermediates can evolve to more severe conditions, including neuronopathy. Sphingolipidoses are rare diseases, but most probably, the patients suffering sphingolipidoses are significantly higher in number, since the exact diagnosis should be accompanied by enzyme activity tests and genetic screening.

## 1.2 Gaucher Disease

One of the widespread sphingolipidosis remains Gaucher disease (GD). Philippe Gaucher recognized the first patient suffering from splenomegaly in 1882. GD is a genetic disorder with rare, recessive, and autosomal traits. The disease arises from the mutations located in the GBA gene, which spans on chromosome 1 region 1q21. Variants present in the GBA gene cause a remarkable decrease in glucocerebrosidase (GCase) function, which functions on hydrolyzing glucosylceramide into ceramide and glucose. It is also known that, on rare conditions, dysfunctional GCase activator, saposin C, can invoke Gaucher Disease. [10] Even though the phenotype is volatile, GD is classified into three clinical subgroups. The most common form of GD is type 1, without neurological impairment. However, types 2 and 3 are linked with neurological damage. In general, GD's presence in a population is around 1/50,000 births but can strike 1/10 births with Ashkenazi Jewish.

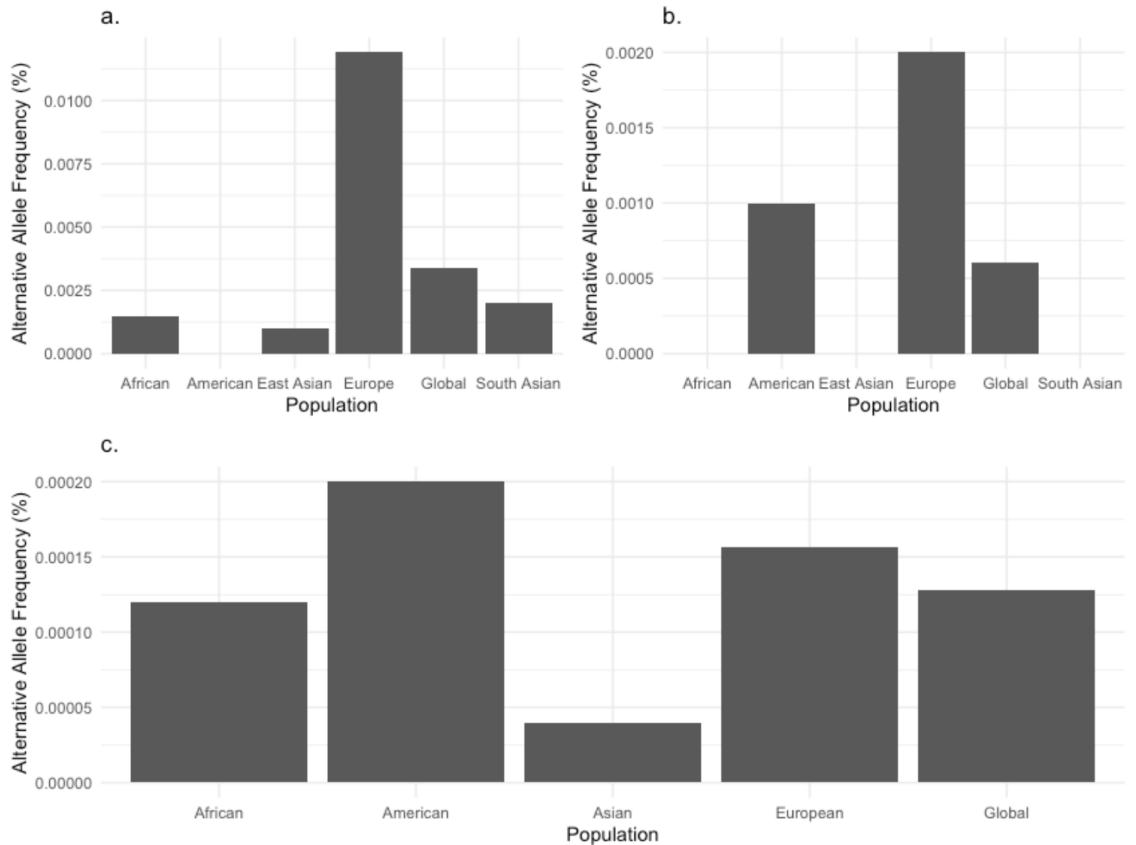


Figure 1.1 Alternative allele frequencies of the most frequent GBA variants.

The information plotted in a and b, has been retrieved from 1000 Genomes Project [11] and represents L444P and N370S respectively, whereas information plotted in c has been retrieved from gnomAd [12], and represents D409H.

### 1.2.1 Pathophysiology of Gaucher Disease

GCcase becomes dysfunctional with the severity of the mutations present on the GBA gene. Such variation will lead to the GCcase substrate's accumulation, glucosylceramide, and transform the cells into Gaucher cells, which exhibit heterogeneous patterns resulting from condensing cytoplasm and fibrillar rearrangement. The leading factor of the disorder's signs can be considered as the emergence of Gaucher cells in the bone marrow and internal organs such as the spleen and liver. Eventually, glucosylceramide increase in Gaucher cells provokes the necrotic difficulties through preoccupation of the bones and driving vascular complications. [13] Yet, neurological involvement of the Gaucher cells demands more explanation. For instance, neuronal GlcCer degradation is at a relatively lower rate, and its accumulation is notable with reduced GCcase activity due to GBA gene mutations. [14] Accordingly, *Drosophila* models of mutant GCcase bearing fly brains revealed autophagy impairment. [15] Importantly, GD may occur due to the variations located in the PSAP gene with the generated dysfunctional saposin C, which results in the loss of ability to activate GCcase. This finding proves that the only reason for GD is not only due to the changes in the GBA gene, but also its activator protein. Furthermore, patients carrying PSAP gene mutations display neurological symptoms comparable to type-3. Due to the notable interaction of GCcase and SapC, we have evaluated the mutant GCcase performance on the interaction with the opened SapC to form a functional complex through docking experiments.

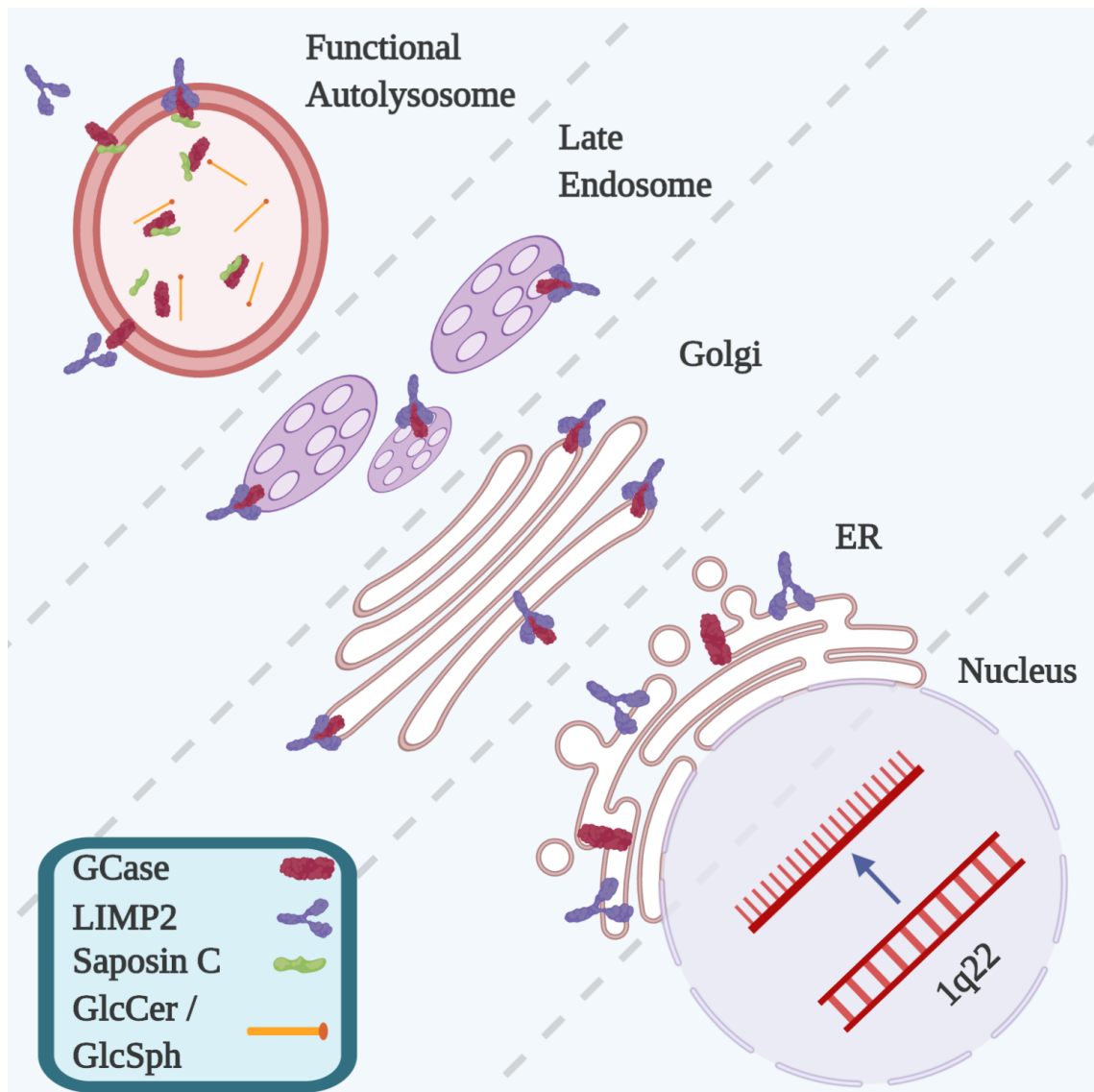


Figure 1.2 Schematic of the trafficking of glucocerebrosidase presence in the cell.

The mRNA of the GCase enzyme is transcribed from the GBA1 gene located at 1q22 and moved out of the nucleus to the endoplasmic reticulum. GCase enzyme is produced in the ER, and it binds with LIMP2 protein in the cytoplasm to be moved through the Golgi. Glucocerebrosidase enzyme is later shifted to a late endosome. During the late-stage endosome blends with a lysosome to build an autolysosome, LIMP2 protein releases GCase because of a pH drop. In the lysosomes, Saposin C activates the GCase enzyme to hydrolyze GlcCer and GlcSph substrates.

GlcCer substrate is also accompanied in another cellular pathway, which a ceramidase alters the substrate into glucosylsphingosine. After the modification, decreased hydrophobicity causes the adjusted form to disperse into fluids. Deficiency in Glucocerebrosidase enzyme as in GD, usually promotes the stated pathway where GBA2 gene is involved. This other GCase is functional under neutral pH conditions and

operates glucosylsphingosine where the product of the enzymatic reaction is sphingosine and finally sphingosine-1-phosphate. [16] Thus, the product sphingosine can inflict toxicity to the bone tissues, where GBA2 gene removal can substitute the GD signs involving bone deformities. Furthermore, the neuronal malfunction can be the result of the gathering of glucosylsphingosine, and shows primarily the causes of GD associate neurological signs. [17] For instance, the presence of glucosylsphingosine is usually low in the brain tissues, yet, the lesions of patients with GD show a more significant accumulation of glucosylsphingosine, which becomes a more precise biomarker when compared to chitotriosidase or CCL18 genes expression as to assess GD severity in clinical practice. [18]

Glucocerebrosidase deficiency is not only a single result of the loss of intrinsic enzymatic capabilities but also can be the outcome of the problems in transportation of the GCase towards the lysosome. For instance, the misfolding of GCase commences to its proteasomal degradation while delivering the endoplasmic reticulum. [19] LIMP-2 (Lysosomal Integral Membrane Protein 2), is highly essential for the transport of glucocerebrosidase to the lysosome and inactivates GCase for stability during transportation. The acidic environment of the lysosomes invokes the delinking process and further activation of the enzyme. [20] SCARB2 gene encoding LIMP-2 has various impacts on the GD phenotypes while modifying the symptoms due to the SCARB2 gene's mutations. These modifications are suggested to establish a base for turning a patient's GD phenotypes from Type-1 to Type-3. [21] Additionally, progranulin is a chaperone protein that accompanies GCase and LIMP-2 complex. Progranulin loss of GD patients demonstrated unusual ER trafficking, while the GCase and LIMP-2 complex accumulated in the cytoplasm, therefore resulted in degradation of the complex. [22] Consequently, patients with the same mutations are expected to expose different phenotypes which are highly dependent on molecular co-abnormalities.

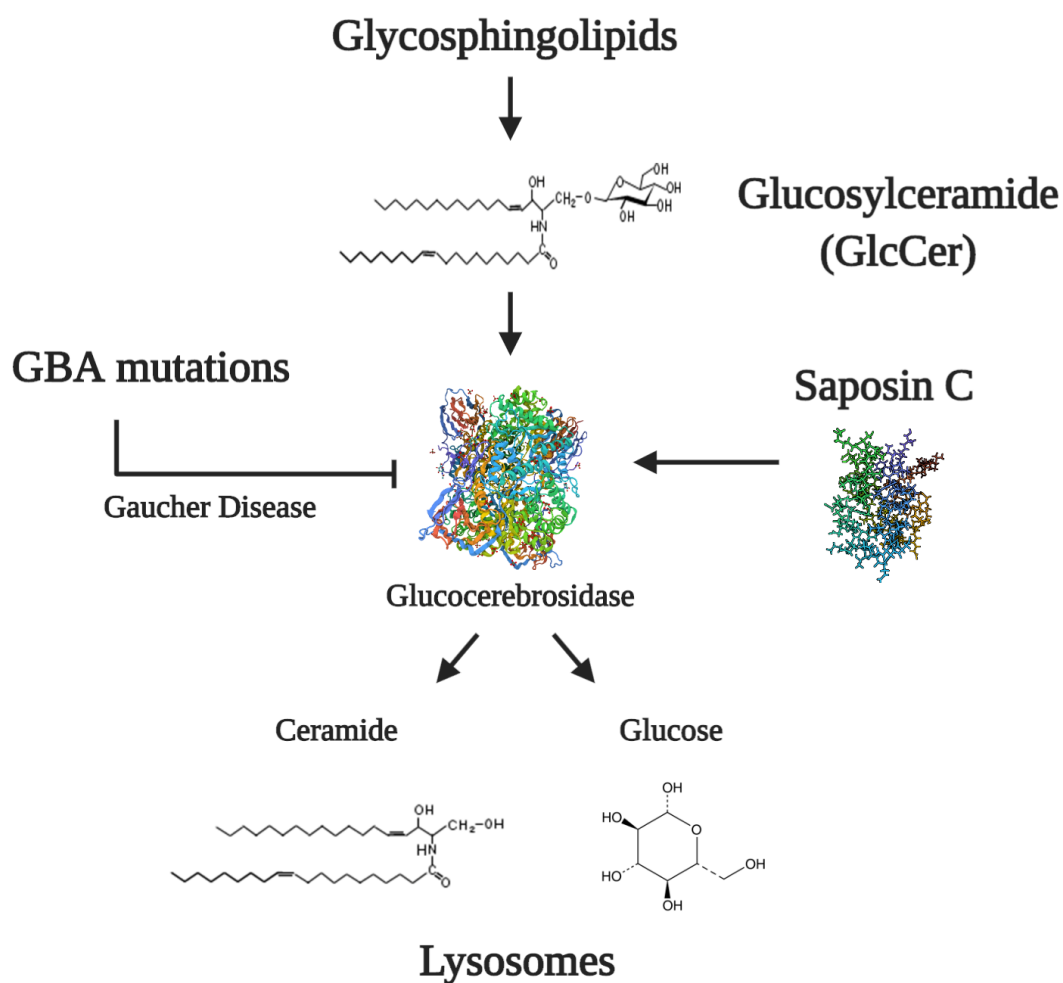


Figure 1.3 Saposin C activates glucocerebrosidase (GCase) for the hydrolysis of glucosylceramides (GlcCer).

The aggregation of the substrates in the lysosome is caused by an enzyme deficiency which eventually leads to lysosomal storage diseases. In Gaucher disease, the insufficiency of functioning glucocerebrosidase directs to the accumulation of glucosylceramide substrate, which accordingly presents arrangements of fibrillar aggregates in the cells.

Hepatosplenomegaly is a condition that causes enlargement and swelling of the liver and spleen, and the situation is linked with Gaucher disease. Yet, additional symptoms are accompanied with GD in the disease types. In some forms, severe neurological impairment is demonstrated with cytopenia and bone involvement. Thus, the variability of the symptoms are the result of a changing phenotype of the disease through the life of a patient. [23]

The most frequent type of GD is Type-1 GD, regularly recognized with the lack of neurological impairment. The clinical appearance has a diverse range of spectrum from asymptomatic to the forms diagnosed at early childhood. Primary signs differ considerably, and cases are diagnosed without any strict age, yet with an average age of 10 to 20 [24]. Patients suffering from Type-1 GD, usually accommodate fatigue, and children exhibit growth retardation and delayed puberty. [25] The most common symptom related to Type-1 GD is splenomegaly, and the sign accompanies more than %90 of the patients. Additionally, another common symptom of Type-1 GD becomes hepatomegaly covering more than %60 of the cases.

Type-2 GD is usually identified with severe neurological impairment commencing in infancy. Further symptoms such as hepatosplenomegaly follow the phenotype. Neck rigidity, bulbar palsy, and oculomotor deficiency are the most familiar symptoms. In children, psychomotor development is heavily neutered, while resulting in developing a lack of learning skills. If a patient is resistant to antiepileptic drugs, seizures usually become a part of the symptoms. Similar to Type-1, splenomegaly is present in most of the cases, and in children, growth retardation is a considerable sign. Type-2 GD usually involves the patient's death before the age of 3 and is considered the most severe form [26].

Type-3 GD usually presents the signs demonstrated in type-1 GD with addition of oculomotor neurological involvement. More than %5 of the GD patients are associated with type-3 GD, and the symptoms are generally seen before the age of 20. [27] As in type-1 GD, type-3 related symptoms are very varied and especially concerns neurological involvement, which most of the time comes into view as horizontal ophthalmoplegia. Additional symptoms originating from neurological impairment may be listed as myoclonus epilepsy, cerebellar ataxia or spasticity, and few cases with dementia. [28] Neurological symptoms can be seen after physical signs as described in type-1 GD, and such extensions cause misclassification of the patients.

### 1.2.2 Therapeutic Approaches

Depending on the varying phenotypes, particular medication can not be maintained in all cases. Once GD symptoms have begun, treatment is commonly offered throughout the patient's whole life. Today, two particular types of treatment are present and can be listed as; enzyme replacement therapy and substrate reduction therapy. The aim is to operate patients before the origin of harsh symptoms, and in order to prevent massive complications such as splenomegaly that can not be reversed by further treatments.

#### 1.2.2.1 Enzyme Replacement Therapy

The enzyme replacement therapy (ERT) system, is used to aid the GCase deficiency, by providing functional GCase to assist the Gaucher cells. Following employing an obtained enzyme from the placenta, early trials has started to produce a recombinant GCase. These enzymes, imiglucerase (Cerezyme®, Sanofi-Genzyme), are deglycosylated to demonstrate their mannose residues to recognize their capturing by macrophage receptors. Furthermore, the enzyme is brought to lysosomes. Additional recombinant enzymes are delivered to the market originating from different sources such as Chinese Hamster Ovary derived imiglucerase, and human fibroblasts derived velaglucerase (Vpriv®, Shire) and carrot cells derived taliglucerase (Elelyso®, Pfizer). Despite efforts to develop enzymes for replacement therapy, these enzymes demonstrated an insignificant influence on the accelerated progress of GD-linked cruel neurological symptoms with neither stabilization nor reversion. [29]

#### 1.2.2.2 Substrate Reduction Therapy

The goal of substrate reduction therapy (SRT), is to lessen massive amounts of GlcCer by reducing its production. An example stands as Miglustat (Zavesca®, Actelion), a synthase inhibitor for glucosylceramide (GlcCer) aiming to diminish the produced GlcCer in the cell and employed as a therapy for mild and moderate levels of Type-1 GD. Yet, this therapy is a choice under the circumstances where enzyme replacement therapy is not available. [30] The substrate inhibitor is benefi-

cial for the prevention of size increase in the liver resulting from hepatomegaly while reducing the chitotriosidase amounts in the cells. Additionally, a GlcCer synthase inhibitor, eliglustat (Cerdelga®), Sanofi-Genzyme), is a precise ceramide analogous and demonstrated efficacy on the long-term treatments of adult Type-1 GD patients [31].



### 1.2.2.3 Gene Therapy

A preceding gene transfer experiment was investigated on Type-3 GD patients [32]. Transfection of the GBA1 gene aimed to introduce wild-type genes with hematopoietic cells and insert the repaired cells in the patients showing the GD-linked phenotype. Yet, the demonstrated results presented insufficient treatment effects, where the glucocerebrosidase presence in the cells was too low, and weak clinical outcomes have been captured. Additionally, lentiviral vector gene transfer methods using gammaretroviral vectors have been performed on mice models of Type-1 GD, and showed correction of the enzyme deficiency [33].

## 1.3 Lysosomal Biogenesis

Lysosomes establish the prominent degradative organelles of the cell. They accept their substrates by endocytosis, phagocytosis, or autophagy. Lysosomes and lysosome-related organelles are required for numerous physiological pathways, such as plasma membrane repair, bone and tissue remodeling, pathogen defense, cell death, and cell signaling. These multiple roles present the lysosome as a central and active organelle in the faith of a cell.

Soluble lysosomal hydrolases and integral lysosomal membrane proteins are two classes of proteins crucial for lysosomal function. Lysosomal hydrolases target particular substrates for degradation, while their combined performance holds lysosomes' catabolic role. Besides their degradation roles, they are also included in starting apoptosis, extracellular matrix (ECM) degradation, and antigen processing [35]. Lysosomal membrane proteins exist essentially in the lysosomal limiting membrane while having different roles such as lysosomal lumen acidification, protein import from the cytosol, membrane fusion, and transportation of degradation products to the cytoplasm [36]. The most abundant lysosomal membrane proteins are lysosome-associated membrane protein 1 (LAMP1), LAMP2, integral lysosome membrane protein 2 (LIMP2), and CD63.

Endocytic pathways supply the cargo and lysosomal proteins to lysosomes for the degradation process. Endoplasmic reticulum (ER) takes the role of synthesizing the lysosomal proteins, and the proteins are transported to the trans-Golgi network (TGN) through the Golgi apparatus. After TGN, lysosomal proteins follow the con-

stitutive secretory pathway to the plasma membrane and consequently uptaken by lysosomes through endocytosis. Besides, lysosomal proteins may follow a direct intracellular path to the endo-lysosomal system. For instance, the clathrin-dependent transportation of lysosomal hydrolases by mannose-6-phosphate receptors (M6PRs) is a well-studied pathway as a direct intracellular path. Yet, there are multiple pathways for transporting both lysosomal hydrolases and lysosomal membrane protein, such as GCase transportation mediated with LIMP2.

Currently, there are several described models of lysosomal biogenesis in the literature [37]. Early endosomes' development from the plasma membrane and the subsequent maturation to late endosomes and lysosomes is one of the described models for lysosome biogenesis [38]. The subsequent model includes vesicular transport, where endosomal transporter vesicles move the cargo from early to late endosomes and to lysosomes; or directly from the late endosomes to lysosomes. The third model signifies late endosomes build a contact site with lysosomes moving the cargo, and resulting in the separation of lysosomes and the late endosomes. The last lysosome biogenesis model suggests a heterotypic combination of late endosomes-lysosomes to develop hybrid organelles, accompanied by re-arranging the lysosomes, as described in the GCase transportation to lysosomes (Figure1.2).

### 1.3.1 Lysosomal pH Regulation

The pH of the cellular elements is strictly regulated and essential for biological events. The proton transport into organelles is fundamentally organized with the proton pumps depending on ATP activation, and named as  $H^+$ -ATPases or V-ATPases, constituting the primary elements of pH homeostasis. In order to efficiently acidify an organelle with a V-ATPase, the operation should contain a mechanism to disseminate the generated luminal-positive transmembrane voltage, a process named as a counterion pathway. In general, two mechanistic pathways are responsible for the dissemination. Initially, cation permeability can transfer cations out of the specific organelle as the process continues with each translocated proton. Additionally, an anion is transferred to the lumen by anion permeability for each proton translocated. Yet, both of these mechanisms can be used in combination to increase acidification ratios further effectively. Since these activities separate electric charge while generating a transmembrane voltage, another ion needs to lead to disseminate this voltage to pump interior. It is additionally known that, ClC-7, a  $Cl(-)/H(+)$  antiporter, is involved in this process. V-ATPase and the counterion

transporter are essential factors in the mechanisms managing the pH of the lysosome. With these mechanisms in lysosomes, V-ATPases maintain a low pH around 4-4.5, while also critical for newly produced acid hydrolases of Golgi to transport to the lysosome.

VATPases are formed with different 14 subunits and arranged to ATP-hydrolytic domain (V1 ) and a proton-translocation domain (V0 ) that operate the rotary mechanism [38]. Intracellular events of VATases are rigidly coordinated with various mechanisms such as reversible separation of the the V1 and V0 domains. The V1 domain is a peripheral complex with a molecular weight of 650 kDa. The domain is found within the cytoplasmic section of the membrane, which provides ATP hydrolysis. The V0 domain is enclosed in the membrane and is a complex with a molecular weight of 260 kDa. The domain is effective for the transporting of protons originating from the cytoplasm to the lumen or extracellular area. V1 domain is formed of eight subunits listed as A, B, C, D, E, F, G and H, whereas V0 includes six subunits listed as a, d, e, c, c and c. The VATPases are similar to the F1 F0 ATP synthases with features of being large, multisubunit complexes formed of a peripheral ATP hydrolytic domain, and with an integral proton-translocation domain [39] . Unlike FATPases, they do not operate under physiological requirements to synthesize ATP from ADP and phosphate.

### 1.3.2 V-ATPase Regulation

Since VATPases are employed in numerous physiological pathways, the organization of VATPase activity is vital. Regulation is performed through multiple mechanisms such as reversible dissociation of the V1 V0 domains, restraining their cellular localization, and variations in the ability of proton bounding and transport accompanied by ATP hydrolysis. Many interactions of VATPases were known, which are essential for this regulation. Multiple stimuli can cause VATPase complexes to reversibly dissociate toward their constructors V1 and V0 domains, eventually closing down ATP-dependent proton transport [40].

Differences in the ability of proton bounding transport with ATP hydrolysis change acidification with different rates [41]. Initially, VATPases that accommodate several isoforms of subunit constitute varying proton coupling capabilities. For instance, the VATPases present in the Golgi hold a weaker proton coupling ratio compared to the VATPases present in the vacuole. Consequently, reversible function and isoform constituents resulting in the varying proton coupling rates are the major causes for

the efficacy of acidification.

#### **1.4 Aim of the Thesis**

In this study, we have presented a bioinformatics analysis of the Gaucher disease related single nucleotide variations in order to better understand their categorical properties, and categorized Turkish population-derived L296V mutation in the GD-linked SNPs spectrum. From this point forth, we have established a background to compare and contrast whether these variations can be subjected and respond to newly developing lysosomal re-acidification therapies through increasing V-ATPase expression. The obtained results in this study, establishes a basic research stage for re-acidification therapies on Gaucher Disease patients.

## 2. MATERIALS and METHODS

### 2.1 Wet-Lab Analysis

#### 2.1.1 Primary Cell Culture

Patient samples have been collected from three Turkish GD patients bearing mutations in the GBA gene. Originating from these samples, primary fibroblast cultures were set in Hacettepe University, Department of Medical Biology, by our collaborators Serap Dökmeci and brought to SUNUM (Sabanci University Nanotechnology Research and Application Center) laboratories for distinct molecular experimentation. The patient, as encoded as Patient 2 (P2), has homozygous L444P mutation and diagnosed as Type 3 with several symptoms such as bone defects, hepatosplenomegaly, and neuropathy. Patient 3 (P3) carries homozygous D409H mutation and diagnosed as Type 1 with symptoms, including cardiac valvular involvement, severe cardiac diseases, and common GD symptoms. Patient 4 (P4) carries heterozygous N370S and identified as Type 1 with hepatosplenomegaly symptoms. Control fibroblasts in our experiments have been received from the gingiva of healthy individuals. Unfortunately, wet-lab experimentation could not have been sustained with the culture of patient encoded as P1 (L296V), due to lack of patient-derived tissues.

| Amino Acid Substitution | Patient ID / GD Type | Phenotype                                                                 |
|-------------------------|----------------------|---------------------------------------------------------------------------|
| L296V                   | P1 / Severe Type 1   | Growth retardation<br>Bone defects<br>Hepatosplenomegaly<br>Hypersplenism |
| L444P                   | P2 / Type III        | Bone defects<br>Hepatosplenomegaly<br>Neurological symptoms               |
| D409H                   | P3 / Type I          | Severe cardiac disease<br>Bone defects<br>Hepatosplenomegaly              |
| N370S                   | P4 / Type I          | Hepatosplenomegaly                                                        |

Table 2.1 The relation between SNPs and the symptoms of GD-patients.

Primary cell lines were sustained in Dulbecco's modified Eagle's medium (DMEM, Pan Biotech, P04-03500) and supplemented with %10 (v/v) heat-inactivated fetal bovine serum (FBS, Pan Biotech, P30-3304), 100 U/ml penicillin/streptomycin (Biological Industries, 03-031-1B), 2 mM L-glutamine (Biological Industries, 03-020-1B) and 1x MEM non-essential amino acid solution (Pan Biotech, P08-32100). Cells were maintained at 37°C with 5% CO<sub>2</sub>.

For cryopreservation of cells, cells in 10cm culture plates were detached with trypsin (Pan Biotech, P10-019500) then subjected to counting. Each cryotube contained at least 2 million of cells for further utilization. Cryotubes were prepared with 900  $\mu$ L cell suspension and an addition of 100  $\mu$ L DMSO (Neofroxx GmbH, 67-68-5). Cryotubes were frozen in cryobox, and deposited at -80°C overnight. The next day, cryotubes were conveyed to liquid nitrogen tank for long term maintenance.

In order to thaw cells, cells stored in the nitrogen tank were taken out and thawed into a 37°C water bath for 10 seconds. Accordingly, cells were given into a 15 ml tube, including a 5 ml culture medium. The cell suspension was centrifuged at 300g for 5 minutes in order to remove DMSO, and the pellet was dissolved with 8 ml culture medium and seeded in a 10 cm culture plate.

### 2.1.2 Gene Transfection

The optimum protocol for transfecting primary fibroblasts requires electroporation. Accordingly, Gaucher patient cells are homogenized and counted with Trypan blue reagent. The required number of cells were transferred to 1.5 mL tubes and centrifuged at 3000 g for 5 min. Nucleofactor reagent (Nucleofector<sup>TM</sup> Kits for Human Dermal Fibroblast (NHDF), Lonza) were mixed with a ratio of 4.5:1 nucleofactor to supplement. Next, the supernatant is discarded, and cells were resuspended in 90  $\mu$ L nucleofactor mixture, plasmid HA-NBP1 (Addgene, no:14586) 1:10 pMAX GFP plasmid which is included in the transfection kit. Followingly, the mixture is taken to the electroporation cuvette and electroporated with the AMAXA electroporation device. After 10 minutes of room temperature incubation, cells were mixed with the medium and seeded to 10 cm. plates.

### 2.1.3 Lysosomal pH Measurement Assay

With Gaucher cells, pH curve buffer (prepared with MES hydrate) was performed with 3-3.5-4-4.5-5-5.5-6 and 7. Monensin and Nigericin were freshly prepared and supplemented to pH buffers before starting the reaction, for the application of pH stabilization of the pH curve buffers. LysoSensor<sup>TM</sup> Yellow/Blue DND-160 (PDMPO, Thermo Fisher Scientific, L7545) was used to measure the Ph values of the cells' lysosomes internal pH. Lysosensor was prepared according to the manufacturer's protocol, and 5-minute incubation at 37°C is performed. Consequently, cells were washed with sterile PBS (Pan Biotech, P04-36500), and fluorescence intensity of the cells was measured at the wavelengths 355-460 nm 355-510 nm (Thermo Fisher Scientific, Fluoroskan<sup>TM</sup> Microplate Fluorometer). Lysosomal pH was determined by using these ratios.

#### **2.1.4 B-glucocerebrosidase Enzyme Activity Assay**

After obtaining %80 confluency of the cells in a plate, the reaction was begun with an addition of 2.5 mM 4- methylumbelliferyl -D glucopyranoside (4-MuGlc, Sigma, M3633) substrate in 0.2 M 50  $\mu$ L acetate buffer (pH 4.0). The cells were maintained at 37°C incubator for six hours. The reaction was terminated with the addition of 0.2 M glycine buffer (pH 10.8). The fluorescent activity was measured at 355-460 nM using single measurement.

#### **2.1.5 In-direct Immunofluorescence Assay and Confocal Microscopy**

Amxa transfected cells were seeded on 6-well plates with mounted glass covers. At the peak 48th hour of the plasmid HA-NBP1, cells are rinsed with 1x PBS. Accordingly, cells are fixed with %4 PFA on ice for 20 minutes and 3 times of washing with PBS applied. Fixed cells are blocked with %0.1 BSA and %0.1 Saponin mixture in PBS for 30 minutes on ice. After removing blocking buffer, covers are inverted and mounted on a new 6-well plate where 60  $\mu$ L of 1:200 diluted GBA antibody (Cat. ABN1387, Millipore) is added in the center. Overnight incubation at 4°C is applied. The next day, mounted covers are recovered and washed with 1X PBS, two times. Cells are incubated with 300  $\mu$ L of 1:300 diluted secondary antibodies (Alexa Flour<sup>TM</sup> 568 (Thermo Fisher Scientific) in %0.1 BSA and %0.1 Saponin) at room temperature for one hour. Accordingly, cells are washed with 1x PBS three times, and 1:6000 DAPI staining of the nucleus is applied for 5 minutes at room temperature. Covers are mounted on the slides, and confocal microscopy images are taken with Carl Zeiss LSM 710.

#### **2.1.6 Site-directed Mutagenesis of L444P**

In order to provide clones for Gaucher Disease linked mutations for future studies, plasmid GBA cDNA ORF Clone, C-OFPSpark (Cat. HG12038-ACR, Sino Biologicals) has been retrieved. QuikChange II Site-directed mutagenesis kit (Cat. 200523, Agilent) ingredients has been used for the reaction and mutagenesis has been performed for L444P variant with the following oligonucleotide sequences to replicate amino acid substitution: L444P-Forward (CACTGCGTCCGGGTCGTT) and

L444P-Reverse (AACGACCCGGACGCAGTG). Mutagenesis reaction constituents and thermal cycles are demonstrated in the Table 2.2 and Table 2.3, respectively. Accordingly, the end product has been digested for 3 hours at 37°C with 0.5  $\mu$ L of DPN1 to get rid off the parental DNA. Remaining mutant DNA has been transformed to DH5-alpha with heat shock at 42°C. 10 of the colonies have been selected randomly, these sub-clones of bacteria have been cultivated and plasmids have been isolated. Sequencing has been performed on the isolated plasmids in order to validate L444P mutation.

| Reaction Component  | Volume & Weight        |
|---------------------|------------------------|
| 10X Reaction Buffer | 2.5 $\mu$ L            |
| Template dsDNA      | 150 ng                 |
| Forward Primer      | 1 ul (10pmol/ $\mu$ L) |
| Reverse Primer      | 1 ul (10pmol/ $\mu$ L) |
| dNTP mix            | 0.5 $\mu$ L            |
| DMSO                | 0.75 $\mu$ L           |
| PFUultra            | 0.5 $\mu$ L            |
| dd H2O              | Up to 25 $\mu$ L       |

Table 2.2 Site-directed mutagenesis reaction mixture.

| Step | Temperature ( $^{\circ}$ C) | Time (Minutes) |
|------|-----------------------------|----------------|
| 1    | 95                          | 1              |
| 2    | 95                          | 1.5            |
| 3    | 60                          | 1.5            |
| 4    | 72                          | 12             |
|      | Go to Step 2 -<br>30 Cycles |                |
| 5    | 72                          | 15             |
| 6    | 4                           | $\infty$       |

Table 2.3 Site-directed mutagenesis reaction thermal cyclers protocol.

### **2.1.7 Statistical analyses**

The statistical significance of differences between groups was assessed by two-tailed Student's t-test. Data were represented as means of  $\pm$  S.D. of 3 independent experiments. Values with  $p < 0.05$  were considered as significant.

## **2.2 Dry-Lab Analysis**

### **2.2.1 Sequence recovery**

The nucleotide sequence of the human GBA gene was obtained in FASTA format from the NCBI Gene database. The amino acid sequence of GCase protein (P04062) in the FASTA format was received from the UniProt database (<https://www.uniprot.org/uniprot/P04062.fasta>).

### **2.2.2 Analysis of SNPs in human GBA**

Several bioinformatics algorithms were employed in predicting the functional effects of the SNPs in human GBA protein. These can be listed as PolyPhen-2 (<http://genetics.bwh.harvard.edu/>), SIFT (<http://sift.jcvi.org/>), and PROVEAN (<http://provean.jcvi.org>). Also, PhD-SNP (<https://snps.biofold.org/phd-snp/phd-snp.html>) was used to predict SNPs that present pathogenic consequences in total, high-risk SNPs for further studies are elucidated.

### **2.2.3 Identification of mutant SNPs position in different domains**

The InterPro (<http://www.ebi.ac.uk/interpro/>) tool was utilized to distinguish different conserved regions of GCCase protein and map SNPs positions in these domains. The protein sequence was registered as the input query to identify domains and motifs.

### **2.2.4 Evolutionary phylogenetic analysis of human GBA protein**

The evolutionary conservation of individual amino acids in GCCase protein was inspected by applying the ConSurf web server ([consurf.tau.ac.il/](http://consurf.tau.ac.il/)). ConSurf algorithm employs a method with an empirical Bayesian approach to resolve evolutionary conservation and the classification of putative structural also functional residues of GCCase protein. The resulting conservation score ranging can be listed as; variable conservation between 1 to 4, intermediate conservation between 5 to 6, and highly conserved between 7 to 9.

### **2.2.5 Extrapolation of amino acid changes and disease phenotypes**

The online server applied for inspecting the molecular background of the disease-related variations of amino acids in a mutant protein is the MutPred server (<http://mutpred.mutdb.org/>). That utilizes various characteristics combining function and structure of the GCCase protein as well as its evolution. The prediction precision is enhanced by combining three different servers, SIFT, PSI-BLAST, and Pfam, profiles beside TMHMM and MARCOIL, that employ structural disorder estimation algorithms, which coupled by the information retrieved from DisProt.

### 2.2.6 Stability analysis of human GBA protein

I-Mutant 3 (<http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi>) is employed to evaluate SNP provoked changes in GCase's dependability. I-Mutant determines the unfolding Gibbs free energy changes value as a contrast between the wild type protein and the mutant protein. I-Mutant predicts the potential in free energy change rate while providing a reliability index for the estimations, in which the lowest to the highest range is 0 to 10.

### 2.2.7 Prediction of post translational modification sites for GBA

The post-translational modifications regions in the amino acid sequences of GCase protein and the mutants were inspected with the ModPred algorithm ([www.modpred.org/](http://www.modpred.org/)).

### 2.2.8 Prediction of structural effect of point mutation on human GBA protein

Project HOPE ([www.cmbi.ru.nl/hope/](http://www.cmbi.ru.nl/hope/)) was employed to distinguish the specific SNPs' structural effect in the GCase amino acid sequence. An alignment with BLAST was conducted by the algorithm corresponding to UniProt and PDB databases to obtain the related tertiary structure information and provide a homology model of the protein. Accordingly, the generated model is subjected to the estimation of the new protein features by utilizing the Distributed Annotation System.

### 2.2.9 Prediction of secondary structure

Protein secondary structure was obtained by utilizing the PSIPRED tool (<http://bioinf.cs.ucl.ac.uk/psipred/>), including the matrices generated by PSI-BLAST.

### 2.2.10 Protein-protein interactions prediction

Interactions between proteins are essential in the evaluation of functional interactions existing in the cell. Accordingly, the Search Tool for Retrieving Interacting Genes/Proteins (STRING) database (<https://string-db.org/>) was employed to emerge GCaSe protein interactions with different proteins. The database produced protein-protein interactions through direct or indirect connections within GCaSe protein and other proteins through employing its database of 24,584,628 proteins of 5090 organisms. For this study, GBA as gene and Homo sapiens as organism were entered as the input query.



### 3. RESULTS

#### 3.1 SNP dataset consists of most frequent Gaucher Disease linked GBA

##### mutations

In this research, we practiced comprehensive bioinformatics analyses to reveal the properties of functional genetic variants within the human GBA coding region. The most prevalent reported SNPs of the GBA gene, as L444P, N370S, D409H, and novel L296V, were recovered as in concordance with the available patient-derived fibroblasts. In the NCBI dbSNP database, there are 78 rsIDs with PubMed citation. Still, the present investigation focuses on SNPs that confer single amino acid changes into human GCase protein. Furthermore, to prove experimental consequences in-vitro with our available patient-derived Gaucher Disease cell lines, the SNPs with the most frequent phenotypes have been tested. The SNP list was organized depending on functional properties and is presented in Table 3.1. Additionally, L296V is a novel mutation and demonstrates a missense variation as chr1:g.155206257C>G (GRCh37).

The SNP list was controlled with a multiplicity of in silico SNP evaluation algorithms to determine if a suspected missense mutation affected GCase work. SNPs are commonly known to change the conformation of the proteins, interaction with other proteins, and protein function. Therefore, it is essential to explore the capacity of our SNP list in human GBA to distort the Lysosomal Acid Glucosylceramide protein's structure and function.

Table 3.1 Functional consequence types of the most common SNPs present in the human GBA gene principal isoforms.

| Amino acid substitution (SNP ID) | Isoform Stable ID <sup>1</sup> | Exon  | Reference Allele | Alternative Allele | Consequence Type <sup>2</sup>              |
|----------------------------------|--------------------------------|-------|------------------|--------------------|--------------------------------------------|
| L444P (rs421016)                 | ENST00000327247                | 11/12 | T                | C                  | Missense Variant                           |
|                                  | ENST00000327247                | 11/12 | T                | G                  | Missense Variant                           |
|                                  | ENST00000368373                | 10/11 | T                | G                  | Missense Variant                           |
|                                  | ENST00000368373                | 10/11 | T                | C                  | Missense Variant                           |
| N370S (rs76763715)               | ENST00000327247                | 10/12 | A                | G                  | Missense Variant,<br>Splice Region Variant |
|                                  | ENST00000327247                | 10/12 | A                | C                  | Missense Variant,<br>Splice Region Variant |
|                                  | ENST00000368373                | 9/11  | A                | G                  | Missense Variant,<br>Splice Region Variant |
|                                  | ENST00000368373                | 9/11  | A                | C                  | Missense Variant,<br>Splice Region Variant |
| D409H (rs1064651)                | ENST00000327247                | 10/12 | G                | C                  | Missense Variant                           |
|                                  | ENST00000368373                | 9/11  | G                | C                  | Missense Variant                           |
| L296V<br>(Novel Mutation)        | ENST00000327247                | 9/12  | C                | G                  | Missense Variant                           |
|                                  | ENST00000368373                | 8/11  | C                | G                  | Missense Variant                           |

<sup>1</sup> GBA gene principal isoforms, ENST00000327247, and ENST00000368373 have been taken from APPRIS Database which provides the primary isoform information. <sup>2</sup> Consequence type information has been retrieved through Ensembl VEP (Variant Effect Predictor).

### 3.2 Analysis of SNPs in human GBA

We further utilized four in silico prediction algorithms, as listed; SIFT, PolyPhen-2, PhD-SNP. Regarding the results of PolyPhen-2, L444P and N370S exist as ‘possibly damaging’ to GBA protein function, while D409H is benign. Our SIFT analysis projected a similar scenario as PolyPhen-2, where L444P and N370S exist as deleterious to GBA protein function, while D409H is tolerated.

In silico techniques employ distinct algorithms, usually rising considerable heterogeneity in results. These tools are reviewed in the literature and recommended the algorithms of SIFT and PolyPhen-2 produce a more reliable achievement in recognizing deleterious variations. Hicks et al. confirmed this view, and this makes these tools useful for our study. Nevertheless, SNPs with greater confidence were assumed to be genuinely harmful. Accordingly, high-risk carrying SNPs in our dataset were classified based on estimated pathogenicity with a disease-linked prediction through the tested algorithms. Our SNP group was then analyzed further with PROVEAN, ConSurf, and I-Mutant. The PROVEAN analysis identified D409H as accountable for harmful single nucleotide polymorphism, yet N370S for being a neutral amino acid substitution. Furthermore, L296V and L444P mutations showed exact deleterious scores with the SIFT algorithm, indicating their higher risk of disease potential. Depending on the SIFT and PolyPhen-2 predictions and the symptoms of patients carrying these mutations, we have interpreted L444P, L296V, and N370S substitutions as high-risk carrying mutations for future experiments.

Table 3.2 L444P, N370S, L296V, and D409H are the pathogenic amino acid substitutions with several dangerous consequences on structure, function, and stability.

|            |            | Amino acid substitution |                   |             |                   |  |
|------------|------------|-------------------------|-------------------|-------------|-------------------|--|
| Tool       | Parameter  | L444P                   | N370S             | D409H       | L296V             |  |
| PolyPhen-2 | Score      | 0.938                   | 0.607             | 0.204       | 0.937             |  |
|            | Prediction | Possibly Damaging       | Possibly Damaging | Benign      | Possibly Damaging |  |
| SIFT       | Score      | 0                       | 0.02              | 0.05        | 0                 |  |
|            | Prediction | Deleterious             | Deleterious       | Tolerated   | Deleterious       |  |
| PROVEAN    | Score      | -4.995                  | -2.128            | -3.572      | -2.553            |  |
|            | Prediction | Deleterious             | Neutral           | Deleterious | Deleterious       |  |
| PhD-SNP    | Score      | 5                       | 3                 | 5           | 5                 |  |
|            | Prediction | Disease                 | Disease           | Neutral     | Neutral           |  |

### 3.3 Identification of domains in GBA protein

InterPro was applied to determine domains in GCase protein and distinguish the positions of SNPs. InterPro performs a functional interpretation of proteins by grouping within protein families. It further estimates the appearance of domains and active sites. Accordingly, the GCase protein comprises Glycoside hydrolase family 30, which belongs to a widespread group of enzymes that hydrolyze the glycosidic bond between two or more carbohydrates. The domains found to be TIM-barrel domain (78-427) and beta-sandwich domain (430-492). In our SNP dataset, residues D409 and N370 locates in the TIM-barrel domain, and L444 finds in the beta-sandwich domain. Additionally, the secondary structure of GBA protein has been generated with PSIREN, and it has been observed that the mixed distribution of coil, helix, and strands are present in the secondary structure.

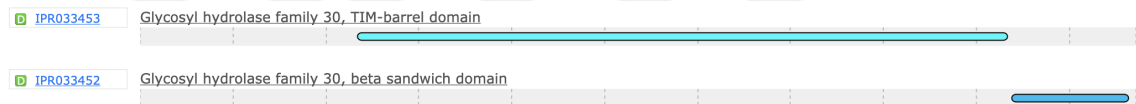


Figure 3.1 InterPro analysis revealed the GBA protein domains belonging to Glycoside hydrolase family 30.



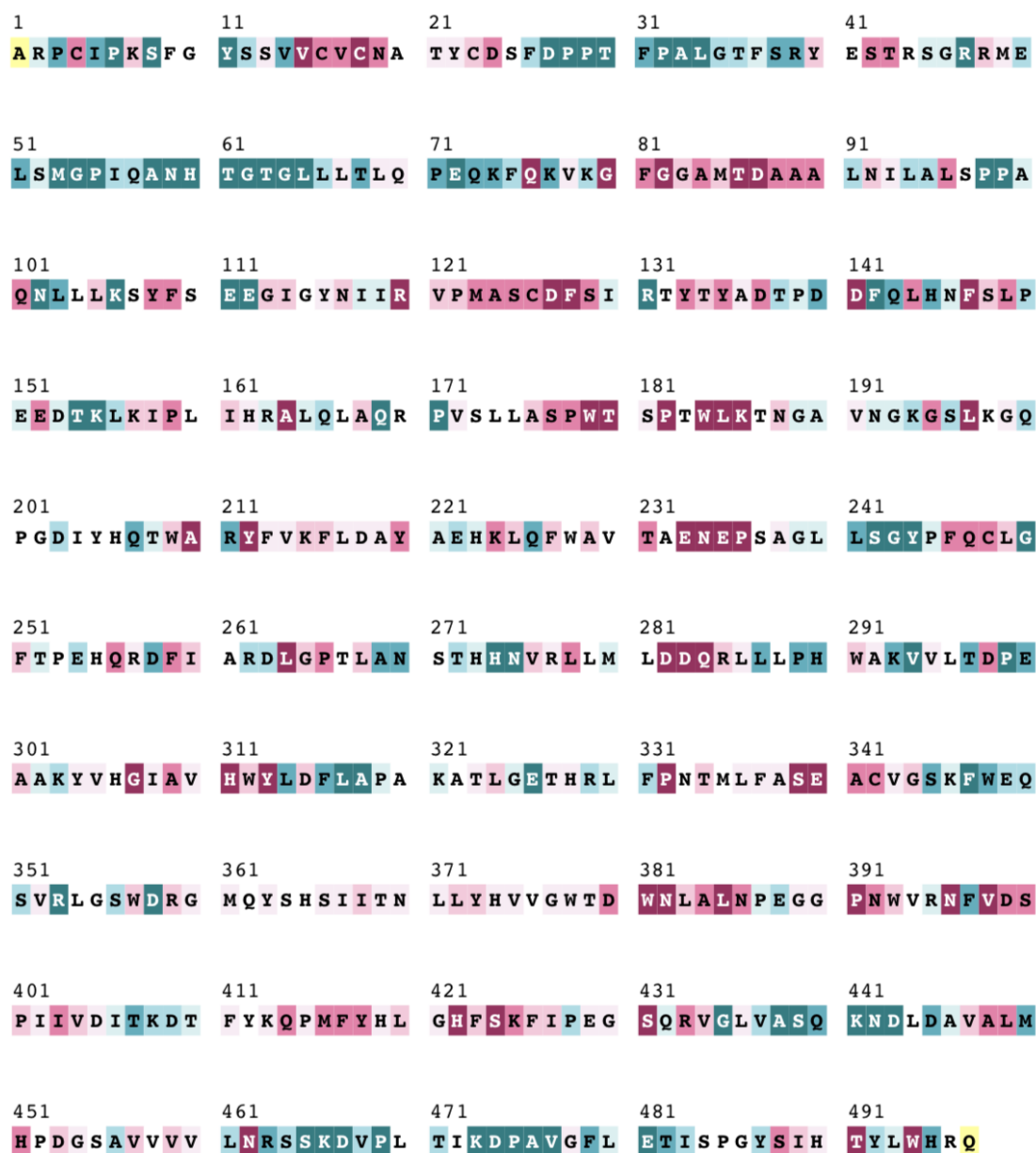
Figure 3.2 Secondary structure of GCase protein.

### 3.4 High-risk SNPs show variable conservation profile

Several critical biological processes regularly require the positional behavior of amino acids, either present in catalytic sites or other sites, as are essential for protein interactions. Accordingly, some amino acids become more distinct and conserved than different residues in a given protein sequence. It has been predicted that amino acid variations located at evolutionarily conserved positions emerge more harmful SNPs than the less conserved sites. Concerning the advanced analysis of the probable effects of the high-risk SNPs, the rate of conservation was determined by applying the ConSurf algorithm for amino acids of human GBA protein in an evolutionary aspect. Accordingly, amino acid positions that meet high-risk SNPs is a reasonable scenario; nevertheless, ConSurf also confirms additional residues that potentially hold useful significance. Consequently, mutations on the higher evolutionary conservation score are expected to hold more upper disease driving potentials.

The higher conserved sites of the protein surface or protein core assists in estimating their structural or functional base. As presented in Fig. 3.4, ConSurf analysis showed that the residue N370 is conserved with a score of 6, and L444 has a prediction of less conserved residues with an average conservation score of 5, with L444P being similar to our novel mutation of L296V.

Interestingly, three of our SNPs, as assumed to be harmful by employed SNP prediction algorithms (L444P, N370S, and L296V), were also structurally or functionally identified as important residues by ConSurf with higher conservation relative to D409H. Our data asserts the claim that, whether these SNPs are deleterious to GBA structure or function matches with relatively higher conservation scores.



**The conservation scale:**



Figure 3.3 Results from the Analysis of human GBA protein by ConSurf reports the divergence in the conservation spectrum of our SNP dataset.

### 3.5 Analysis of structural effects of high-risk SNPs in GBA

Moreover, the project HOPE was used to examine the structural consequences of the variant dataset. Its results showed that L444 and N370 are conserved, and their changes are presumably corrupting to the protein structure. For example, the N370S mutation residue is smaller than the wild-type residue. The N370S variant results in a serine residue in place of asparagine at the 370th amino acid located in the TIM-barrel domain. This domain region is also essential for interacting with Sap C and the phospholipid containing membranes. Substitution of the residue with polar uncharged serine may cause space present in the core structure of the domain and may lead to protein folding difficulties. A view of the variants and resulting structural consequences are shown in the figure. The L444P mutant forms a similar destabilizing condition. Replacing a leucine with a smaller proline residue can create a space in the core structure that potentially modify the signal transduction among the domains. Mutations in the Glycoside hydrolase family 30 domains are hypothesized to be accountable for interrupting the binding action of Sap C to GCase for activation enzyme activation and decreasing the enzymatic activity in general.

| Residue | Structure           | Properties                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|---------|---------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| L444P   | <p>Mutates Into</p> | <p>The wild-type and mutant amino acids differ in size.</p> <p>The mutant residue is smaller than the wild-type residue.</p> <p>The mutation will cause an empty space in the core of the protein.</p>                                                                                                                                                                                                                                                                                                          |
| N370S   | <p>Mutates Into</p> | <p>The wild-type and mutant amino acids differ in size.</p> <p>The mutant residue is smaller than the wild-type residue.</p> <p>The mutation will cause an empty space in the core of the protein.</p> <p>The hydrophobicity of the wild-type and mutant residue differs.</p> <p>The mutation will cause loss of hydrogen bonds in the core of the protein and as a result disturb correct folding.</p>                                                                                                         |
| D409H   | <p>Mutates Into</p> | <p>There is a difference in charge between the wild-type and mutant amino acid.</p> <p>The charge of the wild-type residue is lost by this mutation.</p> <p>This can cause loss of interactions with other molecules.</p> <p>The wild-type and mutant amino acids differ in size.</p> <p>The mutant residue is bigger than the wild-type residue.</p> <p>The residue is located on the surface of the protein.</p> <p>Mutation can disturb interactions with other molecules or other parts of the protein.</p> |
| L296V   | <p>Mutates Into</p> | <p>The wild-type and mutant amino acids differ in size.</p> <p>The mutant residue is smaller than the wild-type residue.</p> <p>The mutation will cause an empty space in the core of the protein.</p>                                                                                                                                                                                                                                                                                                          |

Table 3.3 Results from the Analysis of human GBA protein by Project HOPE reveals the structural effects of high-risk SNPs in GBA.

### 3.6 Identification of disease phenotype relation with our SNP dataset in

#### human GBA

MutPred acquired phenotypic examination of the amino acid variations estimated as pathogenic. MutPred algorithm predicts a diseased phenotype and recognizes the molecular mechanisms that occur from SNPs. Consequently, MutPred also presents relevant data between the classification of the SNP list, including destabilization or loss of solvent accessibility. The p-value scores of the amino acid substitution are shown in Figure 6. P-value scores of less than 0.05 are assigned to as acceptable and actionable hypotheses. The mutations of L444P, N370S, and D409H, are associated with altered stability, altered ordered interface, and modified metal binding. Even with the low prediction scores, L444P and L29V mutations demonstrated the potential to gain Ubiquitylation sites, which might eventually proceed to protein degradation.

| SNPs  | Actionable/Confident Hypothesis        | MutPred2 score | p-value |
|-------|----------------------------------------|----------------|---------|
| L444P | Altered Stability                      | 0.926          | 1.7e-03 |
|       | Altered Ordered interface              |                | 0.03    |
|       | Gain of B-factor                       |                | 0.02    |
|       | Altered Transmembrane protein          |                | 2.6e-03 |
|       | Gain of Ubiquitylation at K480         |                | 0.02    |
| N370S | Altered Ordered interface              | 0.693          | 0.03    |
|       | Loss of Relative solvent accessibility |                | 0.02    |
|       | Gain of Catalytic site at H413         |                | 0.05    |
|       | Altered Metal binding                  |                | 0.02    |
| D409H | Altered Metal binding                  | 0.746          | 8.3e-03 |
|       | Loss of Relative solvent accessibility |                | 2.8e-03 |
|       | Altered Ordered interface              |                | 0.01    |
|       | Altered Transmembrane protein          |                | 1.5e-04 |
|       | Gain of Strand                         |                | 0.02    |
| L296V | Altered Transmembrane protein          | 0.30           | 1.5e-04 |
|       | Gain of Ubiquitylation at K293         | 0.17           | 0.02    |

Table 3.4 Results from the Analysis of our SNP dataset by MutPred enlightens potential reasons for protein instability.

### **3.7 Prediction of post translational modification sites for human GBA**

#### **protein**

ModPred tool was used to examine the effects of SNPs on the post-translational modifications in GCase protein. Accordingly, ModPred estimated residues for proteolytic cleavage and O-linked glycosylation, as presented in Figure 7.

The removal or alteration of post-translational modification residues has harmful clinical assumptions. The variations present at the conserved positions change the post-translational modification of the GCase protein; mainly, D409H could result in the loss of proteolytic cleavage. Interestingly, serine substitution at residue 370 has introduced predicted glycosylation, which might provoke an increase in the protein stability relative to wild-type protein and help the reversion of stability with re-acidification. In contrast to MutPred, L296V and L444p substitution have not impacted post-translational modifications throughout the GCase protein.

### **3.8 Effects of mutations on protein stability and protein-protein**

#### **interaction of human GBA**

There is an agreement on most conditions connected with SNPs affects protein stability. Yet, the consequence of the SNP list on the GCase protein stability was examined by the I-Mutant algorithm. I-Mutant employs the structure-based interpretation of the mutant protein to deliver an assessment of the Gibbs free energy difference in mutant GCase with amino acid change at a particular residue. The I-Mutant results showed that almost all the amino acid substitution studied in this research intends to reduce the degree of free energy change value compared to wild-type GCase. Two significant decreases in mutant proteins' energy change have been seen in L444P and L296V mutations with neutral pH as in parallel with disease driving potentials predicted with SIFT and PolyPhen-2.

| Amino acid substitution | DDG (Kcal/mol) | Sign of DDG (Reliability Index) |
|-------------------------|----------------|---------------------------------|
| L444P                   | -2.39          | Decrease (7)                    |
| N370S                   | -0.91          | Decrease (6)                    |
| D409H                   | -0.34          | Decrease (7)                    |
| L296V                   | -1.54          | Decrease (7)                    |

Table 3.5 Results from the Analysis of our SNP dataset with I-Mutant shows a general decrease with introduced variants.

STRING database was applied to obtain functional interactions of GCase protein. STRING resulted the functional interaction partners of GCase with PSAP (Prosaposin), TCP1 (T-complex protein 1 subunit alpha), GLB1 (Beta-galactosidase), ASAH1 (Acid ceramidase), GALC (Galactocerebrosidase), UGCG (Ceramide glucosyltransferase), B4GALT6 (Beta-1,4-galactosyltransferase 6), GBA2 (Non-lysosomal glucosylceramidase), SGMS2 (Phosphatidylcholine:ceramide cholinephosphotransferase 2), DEGS1 (Sphingolipid delta(4)-desaturase DES1).

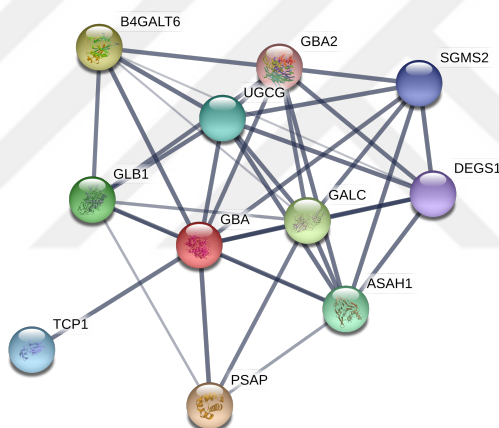


Figure 3.4 Protein-protein interaction network of GBA protein shown by STRING.

### 3.9 Lysosomal re-acidification increases GBA enzymatic activity in-vitro

Within the hypothesis of lysosomal re-acidification corrects the GCase enzymatic activity, patient-derived fibroblasts bearing D409H, L444P, and N370S single-nucleotide mutations were evaluated under V-ATPase over-expression. These groups were monitored for the GCase functionality for the hydrolysis of water-soluble synthetic  $\beta$ -glucosides; 4-methylumbelliferyl- $\beta$ -D-glucoside (MuGlc). Accordingly, the fluorescence intensity, which corresponds to enzymatic activity, has been captured. Yet, lysosomal pH measurement assay has been employed to demonstrate the related pH changes in the lysosomes due to V-ATPase over-expression. Accordingly, decreases in the lysosomal environment's pH provided an overall increase for the enzymatic activity of GCase within all SNPs.

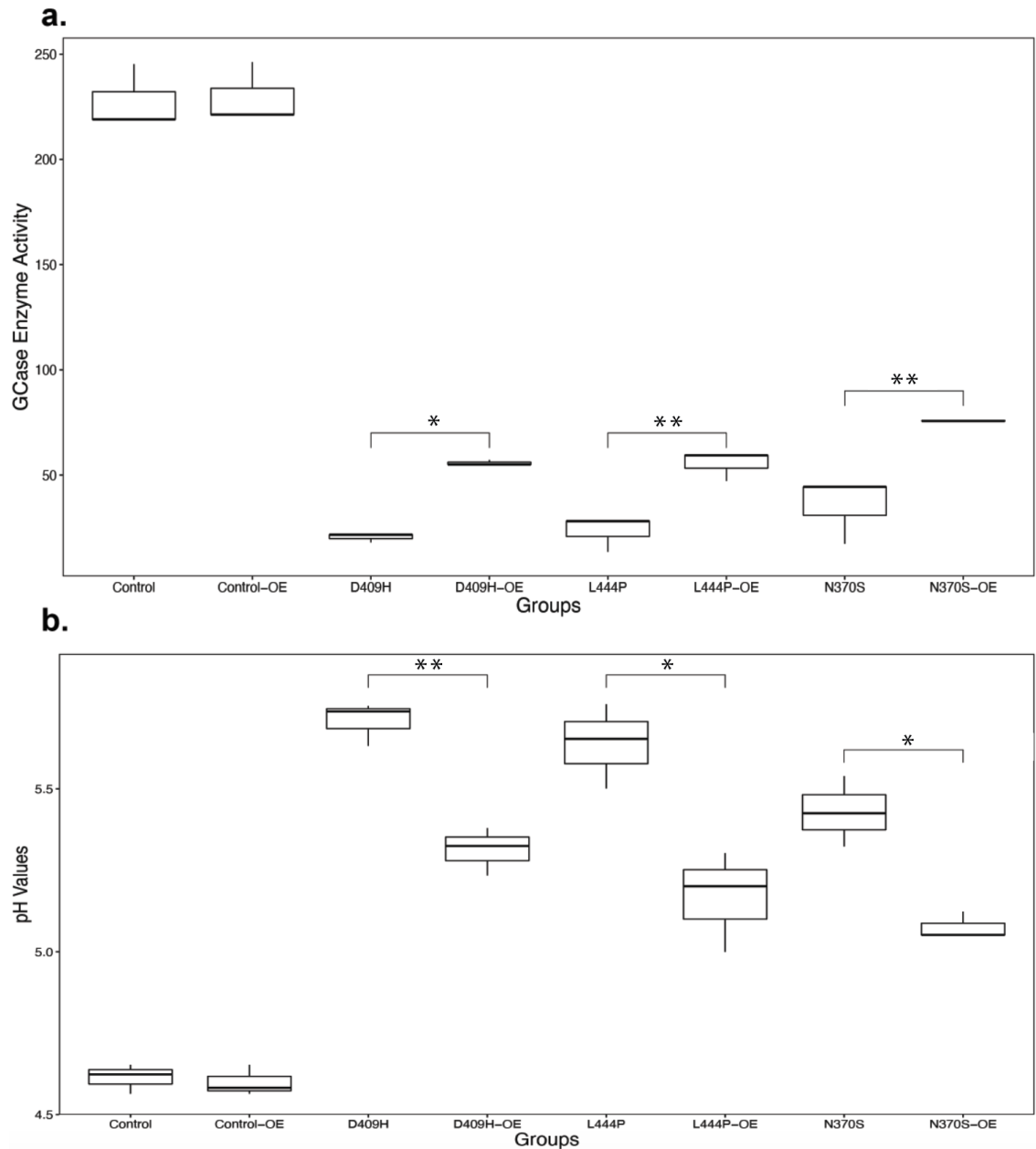


Figure 3.5 Lysosomal re-acidification increases GCase enzyme activity.

a. Results of the GCase enzyme activity assay show overall increase on the hydrolysis functionality for SNPs b. Re-acidification with over-expressed V-ATPase decreases the lysosomal pH and creates an acidic environment for GCase enzyme. The values with-in groups are p-values and demonstrate the significance. (mean  $\pm$  S.D. of independent experiments, n=3, \* p<0,05, \*\* p<0,005)

In order to contrast the efficiency of the correction for GCase activity, differences in the enzymatic activity have been normalized with the decreasing pH values. Within the SNP groups of normal and over-expression, the percent change in the enzymatic activity has been divided into percent change in the lysosomal pH. The resulting absolute values are considered to be normalized  $\Delta$  Enzyme Activity (Fig3.12). It has been found that the most significant change for GCase correction has occurred with the D409H mutation group. L444P and N370S mutation groups showed relatively lower but similar recovered enzyme activity.

| <b>Groups</b> | <b>Normalized <math>\Delta</math>Enzyme Activity</b> |
|---------------|------------------------------------------------------|
| Control       | 2.709469                                             |
| D409H         | 25.00582                                             |
| L444P         | 16.54424                                             |
| N370S         | 17.58011                                             |

Table 3.6 Normalization of the changes in the enzyme activity suits the results for appropriate interpretation between SNP groups.

### 3.10 Lysosomal re-acidification alters GBA sub-localization

To check the status of GBA gene localization and relative localization of the lysosomes, we analyzed the localization of the stained GBA proteins under V-ATPase over-expression. It is clear that control fibroblasts experience GBA density in close proximity to the nucleus. Yet, V-ATPase over-expression has influenced GBA proteins with a wide-spread range throughout the cytosol (Fig. 3.11) with increased expression of GBA proteins.



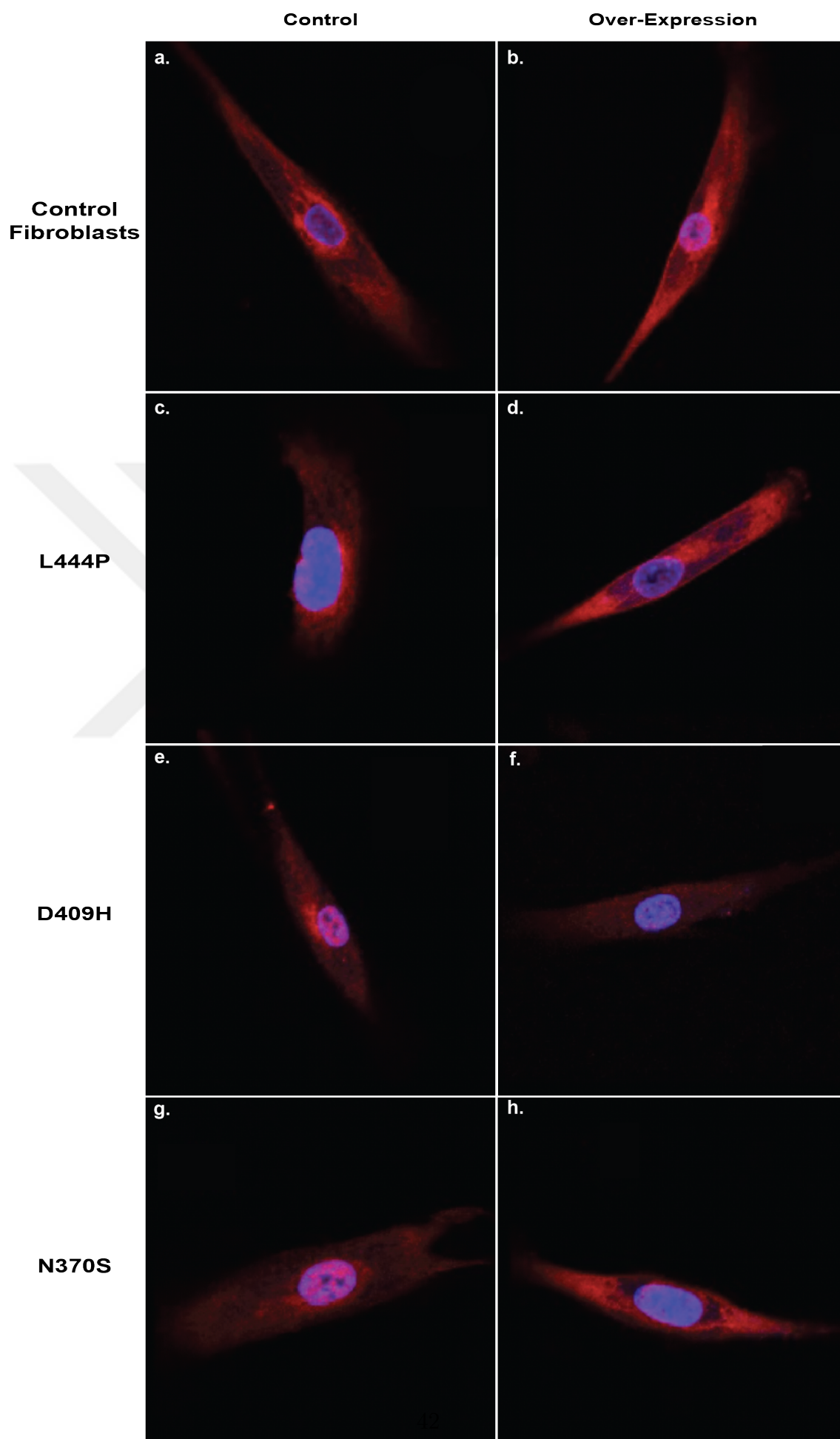


Figure 3.6 Confocal microscopy results of patient derived Gaucher fibroblasts reveals the change of GBA sub-localization after lysosomal re-acidification.

Accordingly, this finding has shaped our perspective to seek any distinctive effect in patient cells under extreme V-ATPase expression. Yet, a similar trend can be clearly seen with the patient cells containing L444P mutation where GBA proteins move off from the nucleus and the increase in the concentration after V-ATPase over-expression. Furthermore, comparable inclination occurs with the patient cells containing N370S variants of GBA protein. Even though a similar migration trend is apparent with the D409H mutation, the level of GBA protein decreases after V-ATPase over-expression. These findings affirmed that, V-ATPase over-expression in fibroblasts presented variations when contrasted to control fibroblasts of all groups in terms of expression levels and localization of -glucocerebrosidase. Yet, each mutation displayed similar features of migration patterns, whereas V-ATPase over-expression showed different patterns with D409H mutation. GBA coupling to the lysosomes with the observed migration pattern might reveal the tendency of increased cellular activity under V-ATPase expression.

## 4. DISCUSSION

In this study, we have employed various in-silico analysis methods to characterize Gaucher disease-related single nucleotide polymorphisms (SNPs) to understand their categorical properties and the reasons for their reaction orders to V-ATPase over-expression. The available patient-derived samples bearing L444P, N370S, and D409H mutations and additively, novel L296V mutation, were investigated in a broad spectrum. Accordingly, we have arranged our SNP datasets with their predicted deleterious effects on GCase function. Our initial reasoning based on different tools, including SIFT and PolyPhen-2, set D409H aside from the most deleterious SNPs. Yet, significant clinical symptoms were present with Patient 3 (D409H), such as severe cardiac disease and the common symptoms, including bone defects and hepatosplenomegaly. Hence, depending on the SIFT and PolyPhen-2 predictions, we have interpreted L444P, L296V, and N370S substitutions as high-risk carrying mutations for future experiments as a basis, and their effects on GCase as deleterious.

Amino acid variations located at evolutionarily conserved positions emerge more harmful SNPs than the less conserved sites, and L444P, L296V, and N370S were also identified as important functional residues with relatively higher conservation scores than D409H. It not surprising that the conservation score scores have followed a similar pattern as the SNPs possess the harmful effects on GCase folding. The structural analysis also demonstrated folding problems in the core of the GCase mutants L444P, L296V, and N370S, while surface interaction disturbance for D409H. It is agreed that most diseases linked with SNPs alter protein stability. Therefore, we have evaluated the degree of free energy change value of the GCase protein at neutral pH. As a concurrent result, L444P, L296V, and N370S mutations presented a higher difference in free energy change ( $\Delta DDG$ ) than D409H. Accordingly, we have built a consensus on clustering L444P, L296V, and N370S mutations aside from D409H. Therefore, when we have mapped the SNPs on o the domains of the GCase enzyme, it is more clear to reason our consensus on subgrouping to assess the effects on the locations of the variations on the GCase structure. It is known that the

catalytic site consists of a domain 3 TIM barrel-helix6 and helix7 (residues 76–381 and 416–430) [42] as a member of the glycosidase hydrolase family. [43] L296V, L444P, and N370 variations, as interpreted as the most deleterious variants, are present in the catalytic region, significantly lowering the GCase efficiency. Hence, the L296V mutation holds its position in the GD-linked SNPs spectrum with L444P and N370S, while possessing harmful effects on the catalytic function of GCase.

Besides the harmful effects of SNPs on GCase, molecular transportation complications of GCase enzyme towards lysosomes might emerge indirect GlcCer accumulation. LIMP2 protein is in charge of targeting GCase to lysosomes. Acidic lysosomal environment detaches GCase from the intracellular receptor LIMP2,[20,44], and GCase's binding region for LIMP2 is held between amino acids 399 to 409. [45] D409H variation poses an obstacle to LIMP2 binding to a considerable extent. Even though lysosomal re-acidification seems to recover the most efficiency with the mutant D409H, it is purposefully possible to be related to an increase in LIMP2 detachment of GCase into the lysosomes.

The effects on protein stability at neutral pH contrasted with the lysosomal re-acidification effects on the mutant GCases at lower pH, whereas the most stable mutant D409H, showed the most recovered enzyme activity. Thus, GlcCer accumulation of D409H variation is in-direct and associated with the mutant GCase enzyme's absence in the lysosomes. L444P and N370 variations showing more significant deleterious impacts on the GCase activity is related to the decrease in the catalytic activity, rather than the molecular transportation to the lysosomes, as in the case of D409H. Accordingly, our consensus on subgrouping the SNPs showed the exact pattern on the mutant GCases' enzymatic recovery with lysosomal re-acidification. Even though the normalized  $\Delta$ Enzyme Activity values demonstrate a considerable and similar increase in the recovery of protein stability with L444P and N370 variations, V-ATPase over-expression should be further subjected to more specific studies in the context of lysosomal biogenesis.

In conclusion, we have suggested a molecular background to compare and contrast whether the SNPs can be subjected to respond to newly developing lysosomal re-acidification therapies through increasing V-ATPase expression. The obtained results in this study organize a necessary research stage for V-ATPase re-acidification therapies on Gaucher Disease patients.

## REFERENCES

1. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NST, Abeyasinghe S, Krawczak M, Cooper DN. Human gene mutation database: 2003 update. *Hum Mutat* 2003;21:577–581.
2. Alaei, M. R., Tabrizi, A., Jafari, N., & Mozafari, H. (2019). Gaucher Disease: New Expanded Classification Emphasizing Neurological Features. *Iranian journal of child neurology*, 13(1), 7–24.
3. Hruska, K.S.; LaMarca, M.E.; Scott, C.R.; Sidransky, E. Gaucher disease: Mutation and polymorphism spectrum in the glucocerebrosidase gene (GBA). *Hum. Mutat.* 2008, 29, 567–583.
4. Goker-Alpan, O.; Hruska, K.S.; Orvisky, E.; Kishnani, P.S.; Stubblefield, B.K.; Schiffmann, R.; Sidransky, E. Divergent phenotypes in Gaucher disease implicate the role of modifiers. *J. Med. Genet.* 2005, 42, e37.
5. Cindik, N.; Ozcay, F.; Suren, D.; Akkoyun, I.; Gokdemir, M.; Varan, B.; Alehan, F.; Ozbek, N.; Tokel, K. Gaucher disease with communicating hydrocephalus and cardiac involvement. *Clin. Cardiol* 2010, 33, E26–E30.
6. Ben Bdira, F., Kallemeijn, W. W., Oussoren, S. V., Scheij, S., Bleijlevens, B., Florea, B. I., van Roomen, C., Ottenhoff, R., van Kooten, M., Walvoort, M., Witte, M. D., Boot, R. G., Ubbink, M., Overkleeft, H. S., & Aerts, J. (2017). Stabilization of Glucocerebrosidase by Active Site Occupancy. *ACS chemical biology*, 12(7), 1830–1841.
7. Parenti, G., Andria, G. & Ballabio, A. Lysosomal storage diseases: from pathophysiology to therapy. *Annu. Rev. Med.* 66, 471–486 (2015).
8. Kolter T, Sandhoff K. Sphingolipids – their metabolic pathways and the pathobiochemistry of neurodegenerative diseases. *Angew. Chem. Int. Ed. Engl.* 38(11), 1532–1568 (1999).
9. Kolter T, Sandhoff K. Principles of lysosomal membrane digestion: stimulation of sphingolipid degradation by sphingolipid activator proteins and anionic lysosomal lipids. *Annu. Rev. Cell. Dev. Biol.* 21, 81–103 (2005)
10. Vaccaro A.M., Motta M., Tatti M., Scarpa S., Masuelli L., Bhat M., Vanier M.T., Tylki-Szymanska A., Salvioli R. Saposin C mutations in Gaucher disease patients resulting in lysosomal lipid accumulation, saposin C deficiency, but normal prosaposin processing and sorting. *Hum. Mol. Genet.* 2010;19:2987–2997. doi: 10.1093/hmg/ddq204.
11. Auton, A., Abecasis, G., Altshuler, D. et al. A global reference for human genetic variation. *Nature* 526, 68–74 (2015). doi.org/10.1038/nature15393
12. Konrad J. Karczewski, Daniel G. MacArthur. 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. *bioRxiv* doi.org/10.1101/531210
13. Mikosch P., Hughes D. An overview on bone manifestations in Gaucher disease. *Wiener Med. Wochenschr.* 2010;160:609–624. doi: 10.1007/s10354-010-0841-y.

14. Orvisky E., Park J.K., LaMarca M.E., Ginns E.I., Martin B.M., Tayebi N., Sidransky E.  
Glucosylsphingosine accumulation in tissues from patients with Gaucher disease: Correlation with phenotype and genotype. *Mol. Genet. Metab.* 2002;76:262–270. doi: 10.1016/S1096-7192(02)00117-8.
15. Kinghorn K.J., Gronke S., Castillo-Quan J.I., Woodling N.S., Li L., Sirka E., Gegg M., Mills K., Hardy J., Bjedov I., et al. A *Drosophila* model of neuronopathic Gaucher disease demonstrates lysosomal-autophagic defects and altered mTOR signalling and is functionally rescued by rapamycin. *J. Neurosci.* 2016;36:11654–11670. doi: 10.1523/JNEUROSCI.4527-15.2016.
16. Dekker N., van Dussen L., Hollak C.E., Overkleeft H., Scheij S., Ghauharali K., van Breemen M.J., Ferraz M.J., Groener J.E., Maas M., et al. Elevated plasma glucosylsphingosine in Gaucher disease: Relation to phenotype, storage cell markers, and therapeutic response. *Blood.* 2011;118:e118–e127. doi: 10.1182/blood-2011-05-352971.
17. Hong Y.B., Kim E.Y., Jung S.C. Down-regulation of Bcl-2 in the fetal brain of the Gaucher disease mouse model: A possible role in the neuronal loss. *J. Hum. Genet.* 2004;49:349–354. doi: 10.1007/s10038-004-0155-2.
18. Rolfs A., Giese A.K., Grittner U., Mascher D., Elstein D., Zimran A., Bottcher T., Lukas J., Hubner R., Golnitz U., et al. Glucosylsphingosine is a highly sensitive and specific biomarker for primary diagnostic and follow-up monitoring in Gaucher disease in a non-Jewish, Caucasian cohort of Gaucher disease patients. *PLoS ONE.* 2013;8:e79732. doi: 10.1371/journal.pone.0079732.
19. Ron I., Horowitz M. ER retention and degradation as the molecular basis underlying Gaucher disease heterogeneity. *Hum. Mol. Genet.* 2005;14:2387–2398. doi: 10.1093/hmg/ddi240.
20. Reczek D., Schwake M., Schroder J., Hughes H., Blanz J., Jin X., Brondyk W., Van Patten S., Edmunds T., Saftig P. LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of  $\beta$ -glucocerebrosidase. *Cell.* 2007;131:770–783.
21. Velayati A., DePaolo J., Gupta N., Choi J.H., Moaven N., Westbroek W., Goker-Alpan O., Goldin E., Stubblefield B.K., Kolodny E., et al. A mutation in SCARB2 is a modifier in Gaucher disease. *Hum. Mutat.* 2011;32:1232–1238. doi: 10.1002/humu.21566.
22. Jian J., Tian Q.Y., Hettinghouse A., Zhao S., Liu H., Wei J., Grunig G., Zhang W., Setchell K.D., Sun Y., et al. Progranulin recruits HSP70 to  $\beta$ -Glucocerebrosidase and is therapeutic against Gaucher disease. *EBioMedicine.* 2016;13:212–224. doi: 10.1016/j.ebiom.2016.10.010.
23. Sidransky E. Gaucher disease: Complexity in a "simple" disorder. *Mol. Genet. Metab.* 2004;83:6–15. doi: 10.1016/j.ymgme.2004.08.015.
24. Stirnemann J., Vigan M., Hamroun D., Heraoui D., Rossi-Semerano L., Berger M.G., Rose C., Camou F., de Roux-Serratrice C., Grosbois B., et al. The French Gaucher's disease registry: Clinical characteristics, complications and treatment of 562 patients. *Orphanet J. Rare Dis.* 2012;7:77. doi: 10.1186/1750-1172-7-77.

25. Kaplan P., Andersson H.C., Kacena K.A., Yee J.D. The clinical and demographic characteristics of nonneuronopathic Gaucher disease in 887 children at diagn
26. Mignot C., Doummar D., Maire I., De Villemeur T.B., French Type 2 Gaucher Disease Study Group Type 2 Gaucher disease: 15 new cases and review of the literature. *Brain Dev.* 2006;28:39–48. doi: 10.1016/j.braindev.2005.04.005.
27. Tajima A., Yokoi T., Ariga M., Ito T., Kaneshiro E., Eto Y., Ida H. Clinical and genetic study of Japanese patients with type 3 Gaucher disease. *Mol. Genet. Metab.* 2009;97:272–277. doi: 10.1016/j.ymgme.2009.05.001.
28. Tylki-Szymanska A., Vellodi A., El-Beshlawy A., Cole J.A., Kolodny E. Neuronopathic Gaucher disease: Demographic and clinical features of 131 patients enrolled in the International Collaborative Gaucher Group Neurological Outcomes Subregistry. *J. Inherit. Metab. Dis.* 2010;33:339–346. doi: 10.1007/s10545-009-9009-6.
29. Weiss K., Gonzalez A.N., Lopez G., Pedoeim L., Groden C., Sidransky E. The clinical management of Type 2 Gaucher disease. *Mol. Genet. Metab.* 2015;114:110–122. doi: 10.1016/j.ymgme.2014.11.008.
30. Aerts J.M., Hollak C.E., Boot R.G., Groener J.E., Maas M. Substrate reduction therapy of glycosphingolipid storage disorders. *J. Inher. Metab. Dis.* 2006;29:449–456. doi: 10.1007/s10545006-0272-5.
31. Lukina E., Watman N., Dragosky M., Pastores G.M., Arreguin E.A., Rosenbaum H., Zimran A., Angell J., Ross L., Puga A.C., et al. Eliglustat, an investigational oral therapy for Gaucher disease type 1: Phase 2 trial results after 4 years of treatment. *Blood Cells Mol. Dis.* 2014;53:274–276. doi: 10.1016/j.bcmd.2014.04.002.
32. Dunbar C.E., Kohn D.B., Schiffmann R., Barton N.W., Nolta J.A., Esplin J.A., Pensiero M., Long Z., Lockety C., Emmons R.V., et al. Retroviral transfer of the glucocerebrosidase gene into CD34+ cells from patients with Gaucher disease: In vivo detection of transduced cells without myeloablation. *Hum. Gene Ther.* 1998;9:2629–2640. doi: 10.1089/hum.1998.9.17-2629.
33. Dahl M., Doyle A., Olsson K., Mansson J.E., Marques A.R., Mirzaian M., Aerts J.M., Ehinger M., Rothe M., Modlich U., et al. Lentiviral gene therapy using cellular promoters cures type 1 Gaucher disease in mice. *Mol. Ther. J. Am. Soc. Gene Ther.* 2015;23:835–844. doi: 10.1038/mt.2015.16.
34. Dell'Angelica, E. C., Mullins, C., Caplan, S. & Bonifacino, J. S. Lysosome-related organelles. *FASEB J.* 14, 1265–1278 (2000).
35. Conus, S. & Simon, H. U. Cathepsins: key modulators of cell death and inflammatory responses. *Biochem. Pharmacol.* 76, 1374–1382 (2008).
36. Eskelinen, E. L., Tanaka, Y. & Saftig, P. At the acidic edge: emerging functions for lysosomal membrane proteins. *Trends Cell Biol.* 13, 137–145 (2003).

37. Trivedi PC, Bartlett JJ, Pulinilkunnil T. Lysosomal Biology and Function: Modern View of Cellular Debris Bin. *Cells*. 2020;9(5)
38. Luzio, J.P.; Pryor, P.R.; Gray, S.R.; Gratian, M.J.; Piper, R.C.; Bright, N.A. Membrane traffic to and from lysosomes. *Biochem. Soc. Symp.* 2005, 77–86.
39. Nishi, T. & Forgac, M. The vacuolar (H<sup>+</sup>)-ATPases: Nature's most versatile proton pumps. *Nature Rev. Mol. Cell Biol.* 3, 94–103 (2002)
40. Yoshida, M., Muneyuki, E. & Hisabori, T. ATP synthase: a marvellous rotary engine of the cell. *Nature. Rev. Mol. Cell Biol.* 2, 669–677 (2001)
41. Kane, P. M. The where, when and how of organelle acidification by the yeast vacuolar H<sup>+</sup>-ATPase. *Microbiol. Mol. Biol. Rev.* 70, 177–191 (2006)
42. Kawasaki-Nishi, S., Nishi, T. & Forgac, M. Yeast V-ATPase complexes containing different isoforms of the 100-kDa a-subunit differ in coupling efficiency and in vivo dissociation. *J. Biol. Chem.* 276, 17941–17948 (2001)
43. Liou, B. et al. Analyses of variant acid beta-glucosidases: effects of Gaucher disease mutations. *J Biol Chem* 281, 4242–4253, <https://doi.org/10.1074/jbc.M511110200> (2006).
44. Liou, B., Haffey, W. D., Greis, K. D. & Grabowski, G. A. The LIMP-2/SCARB2 binding motif on acid beta-glucosidase: basic and applied implications for Gaucher disease and associated neurodegenerative diseases. *J Biol Chem* 289, 30063–30074, <https://doi.org/10.1074/jbc.M114.593616> (2014).
45. Dvir H, Harel M, McCarthy AA, et al. X-ray structure of human acid-beta-glucosidase, the defective enzyme in Gaucher disease. *EMBO Rep.* 2003;4(7):704-9.
46. Emre S, Gürakan F, Yüce A, Rolf A, Scott R, Ozen H. Molecular analysis of Turkish Gaucher disease patients: identification of novel mutations in glucocerebrosidase (GBA) gene. *Eur J Med Genet.* 2008;51(4):315-21.