

December 2020

A Rational Design Approach to Developing Second Generation Fabry Disease Treatments

Matthew Metcalf

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2



Part of the [Biochemistry Commons](#), and the [Structural Biology Commons](#)

Recommended Citation

Metcalf, Matthew, "A Rational Design Approach to Developing Second Generation Fabry Disease Treatments" (2020). *Doctoral Dissertations*. 2057.
https://scholarworks.umass.edu/dissertations_2/2057

This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

**A Rational Design Approach to Developing Second Generation Fabry Disease
Treatments**

A Dissertation Presented

by

Matthew Metcalf

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2020

Molecular and Cellular Biology

© Copyright by Matthew Metcalf 2020

All Rights Reserved

**A Rational Design Approach to Developing Second Generation Fabry Disease
Treatments**

A Dissertation Presented

by

Matthew Metcalf

Approved as to style and content by:

Scott Garman, Chair

Peter Chien, Member

Jeanne Hardy, Member

Daniel Hebert, Member

Scott Garman
Graduate Program Director
Molecular and Cellular Biology

DEDICATION

I dedicate this work to my wonderful family. I really could not have completed this program without the support of my mother, Marianne, who was always there to support me throughout my entire life. She is a truly wonderful mother/person who has sacrificed so much for my brother and me.

I could not have survived graduate school without my brother, Nate, who was my roommate and confidant throughout my graduate career.

I would also like to thank my father, Matthew, my grandparents, Robert, Marlene, Joan, and Dean, as well as my extended family: Jeff, Jules, Emerald, Eden, Christian, Cameron, Lucas, and Amanda.

ABSTRACT

A Rational Design Approach to Developing Second Generation Fabry Disease Treatments

September 2020

MATTHEW C. METCALF, B.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Scott Garman

Fabry disease is an X-linked lysosomal storage disorder that affects approximately 1 in 40,000 males in its classical form and as many as 1:4,600 in its late-onset form [1]. The disease is caused by mutations in the gene encoding α -galactosidase (α -GAL), which results in deficient levels of α -GAL activity in the lysosomes of patients [2, 3]. This lack of enzymatic activity causes macromolecular substrates to accumulate in tissues, and can result in a wide range of symptoms such as impaired renal and cardiac function [4]. The severity of disease is linked to the amount of residual enzyme activity [5, 6]. Mutations resulting in little to no residual activity lead to the more severe classical form of the disease, whereas those that retain a fraction of wild-type enzyme levels lead to the less severe late-onset form of the disease. The FDA has approved both enzyme replacement therapy (ERT) and pharmacological chaperone therapy (PCT) for the treatment of Fabry disease. ERT involves the intravenous injection of recombinant α -GAL. The injected enzyme is delivered to the lysosome through the binding of extracellular mannose-6-phosphate receptors. ERT has been shown to clear accumulated substrate in the majority of tissue types, and has shown to slow the impairment of organ function

typically associated with the disease [7, 8]. Despite being broadly efficacious for the treatment of both forms of Fabry disease, additional treatment options were developed.

PCT recently received approval for the treatment of a select number of Fabry patients. The treatment involves the oral administration of Galafold™, a small molecule also known as 1-deoxygalactonojirimycin (DGJ). DGJ is a potent competitive inhibitor of α -GAL, and acts by stabilizing mutant forms of α -GAL as they traffic to the lysosome [9, 10]. Only a portion of Fabry patients are approved to receive this treatment [11]. These patients possess one of the 348 genetic mutations for which this treatment has been found applicable. The majority of these patients have the late-onset form of the disease and possess some level of residual α -GAL activity. Despite the success of PCT and ERT at treating a broad number of patients there still exists a number of shortcomings associated with either treatment.

One such shortcoming is the immunogenicity associated with the recombinant ERT enzymes. 88% of patients receiving ERT develop immune responses including both IgG and IgE based reactions [12-14]. It is our belief that the immune response is present because classical Fabry patients produce no correctly folded α -GAL, and when correctly folded α -GAL is presented to the immune system it is treated as foreign. This results in immune response and is likely to trigger the formation of neutralizing antibodies. In order to bypass this issue, we interconverted the active sites of α -GAL and a highly homologous human enzyme, α -NAGAL. The engineered enzymes possess lower catalytic efficiency than the wild-type counterparts, but have acquired the selectivity of their counterparts. We confirmed

this through enzymatic characterization and through x-ray crystallography. Most importantly we showed that enzymes retain their native antigenicity. By engineering novel functionality into previously existing protein scaffolds we have highlighted a rational approach to engineering less immunogenic therapeutics.

In addition to possessing undesirable immunogenic properties, the efficiency of ERT in the clearance of substrate in podocyte cells has come into question [15, 16]. These cells play a critical role in the function of the glomerular filtration barrier and ultimately kidney function. Due to the function of these cells, they form the third layer of filtration as well as performing vital role in the structure of glomerular capillaries, they are naturally a problematic cell type for serum-circulated ERT molecules to reach. Permeability of this barrier has been shown to be dependent on charge and molecular diameter, with permeability having an inverse relationship with molecular diameter. By engineering mutations that disrupt the dimer interface of α -GAL we have successfully expressed and purified a monomeric form of the enzyme, which possesses slightly lower than wild-type levels of activity and stability. The monomeric α -GAL may provide an excellent point of entry towards engineering a more potent ERT molecule for the treatment of renal variants of Fabry disease.

While investigating the selectivity and affinity of α -Gal and α -NAGAL towards galactose and α -N-acetylgalactosamine analogs, we observed that at concentrations similar to those used in clinical trials DGJ significantly inhibits human β -GAL. This observation led us to further investigate the effect of α -linked functional groups attached to the position corresponding to C1 in a galactose scaffold. In DGJ and the

2-acetamido variant, DGJNAc, this position is occupied by a hydrogen. We utilized inhibition assays and x-ray crystallography to probe the effect of certain functional groups at this position on specificity and affinity towards α -GAL or α -NAGAL. We determined that these groups provided little to no change in affinity, but provided increased specificity towards α -selecting active sites. These findings coupled with previous studies performed by members of the Garman lab provide a clear set of guidelines towards developing ligands specific towards either α -GAL α -NAGAL.

TABLE OF CONTENTS

| | Page |
|------------------------------------------------------------------------------------------------------------------------------------|-------------|
| ABSTRACT..... | v |
| LIST OF TABLES | xiii |
| LIST OF FIGURES | xvi |
| 1. BACKGROUND | 1 |
| Lysosomes..... | 1 |
| Lysosomal trafficking | 1 |
| Lysosomal storage disorders..... | 2 |
| Fabry disease..... | 4 |
| Human α -galactosidase..... | 6 |
| Treatments..... | 7 |
| Enzyme replacement therapy | 8 |
| Pharmacological chaperone therapy | 11 |
| 2. INTERCONVERSION OF THE SPECIFICITIES OF HUMAN LYSOSOMAL ENZYMES ASSOCIATED WITH FABRY AND SCHINDLER DISEASE..... | 15 |
| Abstract..... | 15 |
| Introduction..... | 15 |
| Results..... | 18 |
| Biochemical Characterization of α -GAL ^{SA} and α -NAGAL ^{EL} | 18 |
| Enzymatic activity α -GAL ^{SA} and α -NAGAL ^{EL} | 19 |
| Crystal structure of α -GAL ^{SA} | 21 |

| | |
|--------------------------------------------------------------------------------------------------------------------------|-----------|
| Discussion | 22 |
| Enzyme replacement therapy | 22 |
| Properties of the engineered enzymes..... | 23 |
| Experimental Procedures | 26 |
| Molecular biology | 26 |
| Cell transfection | 26 |
| Protein expression and purification | 27 |
| Kinetic assays..... | 27 |
| Crystallization and x-ray data collection | 28 |
| 3. THE RATIONAL DESIGN OF A POTENTIAL SECOND GENERATION | |
| ERT THERAPEUTIC WITH INCREASED RENAL EFFICACY | 34 |
| Introduction..... | 34 |
| Results..... | 36 |
| α-GAL_{F273G+W277G} retains significant enzymatic activity | 36 |
| α-GAL_{F273G+W277G} shows increased membrane permeability | 37 |
| α-GAL^{F273G+W277G} displays altered mobility on size-exclusion columns | 38 |
| Discussion | 40 |
| Experimental Procedures | 42 |
| Molecular biology | 42 |
| Cell transfection | 42 |
| Protein expression and purification | 42 |
| Size exclusion chromatography | 43 |
| Membrane permeability assay | 43 |

| | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| Enzyme kinetics | 44 |
| 4. A-HOMO VARIANTS OF DEOXYGALACTONOJIRIMYCIN AND 2-ACETAMIDO-1,2-DEOXY-D-GALACTONOJIRIMYCIN SHOW INCREASED SPECIFICITY TOWARDS ENZYMES ASSOCIATED WITH FABRY AND SCHINDLER DISEASE | 50 |
| Introduction..... | 50 |
| Results..... | 54 |
| Inhibitory properties of GJ, α -HGJ, and α -HGJNAc | 54 |
| Specificity of α -HGJ..... | 55 |
| Atomic basis for iminosugar binding..... | 55 |
| Discussion..... | 57 |
| Experimental Procedures | 59 |
| Crystallization and x-ray data collection | 59 |
| Protein expression and purification | 60 |
| Enzyme inhibition assays..... | 61 |
| 5. CONCLUSIONS AND FUTURE DIRECTIONS..... | 70 |
| SUMMARY | 75 |
| APPENDIX..... | 77 |
| PHARMACOLOGICAL CHAPERONES FOR HUMAN A-N-ACETYL GALACTOSAMINIDASE | 77 |
| HUMAN ACID SPHINGOMYELINASE STRUCTURE PROVIDE INSIGHT TO MOLECULAR BASIS OF NIEMANN-PICK DISEASE | 84 |
| LIGAND-PROMOTED PROTEIN FOLDING BY BIASED KINETIC PARTITIONING | 95 |

BIBLIOGRAPHY.....107

LIST OF TABLES

| Table | Page |
|----------------------------------------------------------------------------------------------------------------------------|------|
| 1. Enzymatic parameters for enzyme interconversion study..... | 31 |
| 2. Structural statistics for interconverted α -galactosidase..... | 33 |
| 3. Summary of iminosugar inhibition constants against α -galactosidase and α -N-acetylgalactosaminidase..... | 65 |
| 4. Structural statistics for iminosugar bound α -galactosidase..... | 69 |

LIST OF FIGURES

| Figure | Page |
|----------------------------------------------------------------------------------------------------------|------|
| 1. Cartoon representation of α -galactosidase structure..... | 13 |
| 2. Model of α -galactosidase secretion pathway | 14 |
| 3. Interconverted α -GAL and α -NAGAL structural and biochemical analysis..... | 30 |
| 4. α -GAL ^{SA} crystal structures..... | 32 |
| 5. Model of the α -GAL dimer interface..... | 45 |
| 6. Cross-species α -GAL amino acid sequence alignment..... | 46 |
| 7. Michaelis-Menten kinetics of α -GAL ^{SA} | 47 |
| 8. Membrane permeability assay of α -GAL ^{SA} | 48 |
| 9. Chromatogram of α -GAL ^{SA} SEC elution..... | 49 |
| 10. Structure of α -galactose and α -N-acetylgalactosamine and imino analogs..... | 62 |
| 11. Enzymatic inhibition of α -GAL by GJ..... | 63 |
| 12. Enzymatic inhibition of α -GAL and α -NAGAL by α -HGJ and α -HGJNAc..... | 64 |
| 13. Crystal structure of α -GAL bound to GJ..... | 66 |
| 14. Crystal structure of α -GAL bound to α -HGJ..... | 67 |
| 15. Comparison of iminosugar binding to α -GAL..... | 68 |

CHAPTER 1: Background

Lysosomes

Lysosomes are dynamic organelles responsible for several key functions of eukaryotic cells. In addition to playing a crucial role in plasma membrane repair, nutrient sensing, and cell signaling the organelle is the primary engine of macromolecular degradation. Degradation is the function most commonly associated with the organelle, as it is both the most thoroughly researched and provides the namesake of the organelle (lysosome translates to digestive body in Greek). In order to degrade the vast amount of substrates from both intracellular and extracellular sources the lysosome is equipped with approximately 60 different soluble hydrolases including glycosidases, lipases, sulfatases, peptidases, nucleases, and phosphatases [17]. When applied to a substrate this array of hydrolases act in sequence to catabolize the macromolecule into basic macromolecular/molecular subunits, which can then be utilized or secreted by the cell.

The lysosomal hydrolases do not act alone in their function to degrade intracellular and extracellular substrate. The lysosome also contains numerous integral lysosomal membrane proteins (LMPs), which act in a supportive role towards the hydrolases. The approximate 25 LMPs present in the lysosome of mammals are responsible for the acidification of the lysosomal lumen, substrate import, and the trafficking of several lysosomal hydrolases into the lysosome.

Lysosomal trafficking

Proper lysosomal function can only be achieved if the required lysosomal hydrolases and LMPs are successfully trafficked to the lysosome. Both classes of

lysosomal protein utilize the secretory pathway and are translocated into the endoplasmic reticulum using the Sec61 translocation channel. The proteins are then trafficked through the trans-Golgi network towards lysosomes. The exact mechanism of this trafficking varies between proteins and certain proteins have been shown to utilize multiple receptor pathways. The most commonly utilized receptor pathway for the soluble hydrolases is the mannose 6-phosphate pathway. In this pathway, N-linked glycans of the cargo molecule are phosphorylated as the protein traverses the trans-Golgi network, and the phosphorylated glycans are recognized by either the cation-dependent mannose 6-phosphate receptor or the cation-independent receptor. The receptor/cargo complex then associates with AP1 and/or GGA proteins to form clathrin-coated vesicles that traffic through early and late endosomes towards mature lysosomes [18, 19]. Other known pathway receptors utilized by the soluble hydrolases include the VPS10 domain-containing receptors sortilin, SorLA, and SOR1-3, and the LMP LIMP-2 [20, 21]. Except for LIMP-2, little is known of the trafficking pathways used by LMPs.

Lysosomal storage disorders

When a gene encoding a lysosomal enzyme is mutated in a fashion that the resulting enzyme is functionally defective, the substrate of defective enzyme accumulates in the lysosome. Accumulated substrates in critical cell types result in disease pathology. The deficiencies as a group are referred to as lysosomal storage diseases, and together they have an estimated incidence of ~1:7,000 live births [22]. The most common lysosomal storage disorder is Gaucher disease. Other examples include Pompe disease, Niemann-Pick diseases, Tay-Sachs disease, and Fabry disease.

There are several scenarios that result in deficient levels of lysosomal enzyme activity. In one scenario the patient possesses the gene for a wild-type lysosomal enzyme, but possess defective genes encoding proteins involved in the trafficking machinery required to deliver the enzyme to the lysosome. An example of this scenario, in which deficient lysosomal enzyme activity results not from deficiencies in the enzyme, but instead defects in a trafficking pathway, is inclusion-cell disease (I-cell disease, also known as mucopolipidosis II). I-cell disease results from a loss of N-acetylglucosamine-1-phosphotransferase, the enzyme responsible for the transfer of GlcNAc-phosphate to N-linked glycans of proteins trafficking through the secretory pathway [23-25]. Deficiencies of the transferase prevent the formation of the mannose 6-phosphate moiety, normally required for utilization of the mannose 6-phosphate receptor pathway, and results in numerous lysosomal enzymes being secreted from the cell instead of being properly targeted to the lysosome. In patients suffering from I-cell disease substrates from many enzymes accumulate, and this disease has a severe phenotype as there many enzymes that utilize mannose 6-phosphate receptor-dependent pathways [24]. An example of a defect in a mannose 6-phosphate receptor-independent pathway leading to disease symptoms can be observed in individuals with mutant forms of LIMP-2. Mutations in LIMP-2 lead to action myoclonus-renal failure [26]. LIMP-2 has been shown to be critical for the lysosomal targeting of β -glucocerebrosidase, the enzyme associated with Gaucher's disease [21, 27]. A subset of β -glucocerebrosidase mutants are localized to the ER, and these mutants also lead to action myoclonus-renal failure [28]. The sub-cellular location of the β -glucocerebrosidase mutants, shared phenotype, and the fact that the mutations

are not located at the enzyme active-site imply that these Gaucher's disease mutants are the result of a loss of complex formation between β -glucocerebrosidase and LIMP-2.

Another scenario that leads to a deficiency in lysosomal enzyme is the presence of mutations that directly alter the catalytic ability of the enzyme. This type of mutation results in an enzyme that folds correctly, and is correctly trafficked to the lysosome, but is unable to catalyze the hydrolysis of its substrate. These mutations are present in the pools of reported disease-causing mutations for the majority of lysosomal storage diseases, but only represent a small portion of these pools. This is illustrated by the fact that most lysosomal enzymes have active sites comprising of only a dozen or so active site residues, but often have hundreds of disease-causing mutations associated with the associated disease.

The most common scenario resulting in a lysosomal storage disease is one in which the disease mutations result in a destabilized protein. These variants of the enzymes possess mutations that negatively impact the stability of the enzyme. In the case of Fabry disease over 60% of mutations occur in the hydrophobic core of the enzyme, and result in the enzyme being unable to evade the endoplasmic reticulum associated degradation pathway (ERAD). A consequence of being recognized by ERAD folding sensors is that the mutant enzyme is targeted for degradation.

Fabry disease

Fabry disease is the second most common lysosomal storage disorder. It is an X-linked dominant disorder that affects ~1:40,000 men; however, studies screening newborn children indicate that the disease may be underdiagnosed and may have an

incidence rate as high as 1:4,600 [1]. This incidence rate is higher than that of the commonly referenced incidence of lysosomal storage disorders as a whole, 1:7000. This discrepancy hints that lysosomal storage diseases are heavily misdiagnosed as a group. Fabry disease is caused by mutations in *GLA*, which encodes α -galactosidase (α -GAL) [2, 3]. A consequence of this deficiency is the accumulation of substrate, primarily globotriaosylceramide (Gb3) and lyso-globotriaosylceramide (lyso-Gb3), which manifests disease symptoms [29, 30]. The severity of the disease is directly related to amount of residual lysosomal α -GAL enzyme levels, and the disease has been partitioned into two forms depending on the amount residual lysosomal enzyme levels in males.

Classical Fabry disease is the most severe form of the disease. Patients suffering from the classical form of the disease possess little to no residual α -GAL activity, 5-15% of normal levels [31]. Classical Fabry disease is characterized by chronic neuropathic pain, skin lesions, lenticular and corneal opacities, angiokeratoma, hypohidrosis, abdominal pain or bloating, as well as cardiac, renal, and cerebrovascular complications [4]. The average lifespan of a person with Fabry disease is 58 years, and many of these individuals require a kidney transplant during their life [32].

Late-onset Fabry disease is the less severe form of Fabry disease. Patients with this form of the disease retain low levels of lysosomal α -GAL enzyme activity, 5-35% of normal levels [6]. Late-onset Fabry disease can further be divided into renal variants and cardiac variants. Both variants do not display the symptoms associated with the classical form such as skin lesions, angiokeratomas, lenticular and corneal opacities, and neuropathic pain. Cardiac variants typically present cardiac symptoms between the fifth and eighth decades of life. Symptoms associated with the cardiac variant include left

ventricular hypertrophy, arrhythmias, and cardiomyopathy [30, 33, 34]. The renal variant typically presents with proteinuria and later onset end-stage renal disease [1, 35, 36].

Female Fabry patients typically display less severe disease manifestations compared to males due to *GLA* being located on the X chromosome [37]. The severity of symptoms in heterozygous females is more difficult to predict based on *GLA* sequencing due to lyonization also known as X-chromosome silencing, which randomly silences gene expression on one X chromosome. In addition, the phenomenon of X-inactivation skewing creates further X chromosome gene expression variation which differs from cell to cell and can result in varying levels of residual enzyme activity between organs within an individual [38, 39].

Human α -galactosidase

Human α -GAL exists exclusively as a homodimer, with each monomer consisting of a β -sandwich domain and a catalytic TIM barrel (Fig 1) [40]. α -GAL is responsible for catalyzing a double-displacement reaction resulting in the hydrolysis of terminal α -linked galactose from galactosides. It accomplishes this function using 13 residues in the active site where aspartic acid 170 acts as the catalytic nucleophile, and aspartic acid 231 acts as both an acid and a base throughout the reaction [41]. The remaining active site residues are responsible for substrate specificity. The bulk of mutations that cause Fabry disease result in single amino acid substitutions, 60% of which occur in the hydrophobic core of the enzyme. Due to this observation, Fabry disease can be considered a protein folding-disease.

Treatments

There are a number of proposed treatments for Fabry disease and other lysosomal storage diseases: pharmacological chaperone therapy (PCT), gene therapy, substrate reduction therapy, hematopoietic cell transplantation, and enzyme replacement therapy (ERT) [13, 42-44]. Presently the only approved treatments for Fabry disease are ERT and PCT, with gene therapy and substrate reduction therapy in clinical trials.

Hematopoietic cell transplantation has been evaluated for a number of lysosomal storage disorders. This treatment involves the introduction of enzyme-expressing cells into the bone marrow, blood, and the central nervous system. This strategy has the advantage of being able to deliver functional enzyme across the blood-brain barrier, and is an appealing option for the treatment of lysosomal storage disorders with associated neurological complications. Additionally, the infrastructure required to produce material for this treatment is less demanding than those required for more traditional treatments, and as such appears as an attractive option for diseases with very small patient populations. The most extensive testing has been performed on α -mannosidosis patients, and patients receiving this treatment in general showed less developmental retardation than those receiving no treatment, although there was significant variation within the patient pool [45]. Despite the promising results from these initial studies it should be noted that there is considerable risk of serious complications associated with this treatment as the procedure is invasive.

Substrate reduction therapy makes use of small molecules to limit production of substrates associated with lysosomal storage disorders. This strategy has been applied to the treatment of type 1 Gaucher disease in the form of Miglustat. Miglustat, also known

as Zavesca®, acts by inhibiting glucosylceramide synthase, which catalyzes the first step in glycosphingolipid synthesis [46]. Inhibition at this step in the synthesis pathway prevents the production of any derivative glycosphingolipids, including the accumulated substrate associated with Gaucher's disease, glucocerebroside. In addition to lower levels of glucocerebroside, all other glycosphingolipids are affected by this compound, and as a result, miglustat has been suggested for the treatment of other lysosomal storage diseases that accumulate glycosphingolipids, and has been approved for the treatment of Niemann-Pick disease type C. It should be noted that initial proof-of-concept studies of miglustat were carried out in a Fabry mouse model, due to limitations of Gaucher's disease mouse models of the time [47, 48]. Since the approval of miglustat, several second generation substrate reduction therapy compounds possessing superior pharmacokinetics and specificity and have been synthesized. Included in this second generation are two compounds that have entered clinical trials for Fabry disease. One of these compounds, venglustat, received FDA fast-track designation in 2015, and is currently in phase 2 clinical trials. One obvious issue associated with the current crop of substrate reduction therapies for lysosomal storage disorders is a lack of specificity, and this is reflected in the fact that miglustat is only approved for Gaucher's patients unable to receive enzyme replacement therapy.

Enzyme replacement therapy

ERT was the initial treatment for Fabry disease (approved in the US in 2003 and Europe in 2001). This treatment involves the intravenous injection of recombinantly expressed α -GAL. The enzyme binds either the mannose 6-phosphate receptor, sortilin,

or megalin, at which point the cell undergoes endocytosis and the enzyme traffics to the lysosome. This strategy can clear Gb3 and lyso-Gb3 deposits in the majority of tissues [8, 13, 49]. The two therapeutic products that have been developed for ERT of Fabry disease are agalsidase alfa (Replagal®) and agalsidase beta (Fabrazyme®). Fabrazyme® and Replagal® are nearly identical except for the N-linked carbohydrates, resulting from the different expression systems used to express the two recombinant enzymes (Fabrazyme® is expressed in CHO cell lines and Replagal® is expressed in human cell lines). Detailed analysis of the oligosaccharide structures of the Replagal® and Fabrazyme® ERT proteins revealed that Fabrazyme® possessed higher degrees of sialylation and phosphorylation [7]. Fabrazyme® was also shown to have a greater amount of enzyme cellular uptake when applied to Fabry mice and cultured human Fabry fibroblasts, due to the higher amount of mannose 6-phosphate glycans [50]. Additionally, it has been shown that high sialic acid content results in higher serum half-life of glycoproteins [51]. This discrepancy highlights the role glycosylation plays in ERT, and has spawned numerous studies aimed towards glycol-engineered ERT molecules with greater efficacy.

Numerous approaches have been developed to produce recombinant enzymes with higher mannose 6-phosphate content. Genzyme/Sanofi has pioneered methods to produce mammalian expressed recombinant acid α -glucosidase and β -glucocerebrosidase with exposed terminal mannose residues, through the use of in vitro exoglycosidase treatments or kifunensine treatment during protein expression. Subsequent enzymatic glycan modifications or chemical conjugations were then used to add mannose 6-phosphate glycans [52-55]. In addition to chemical and enzymatic modifications of mammalian expressed enzymes, alternate cell lines have been used to facilitate the

production of recombinant enzymes with higher mannose 6-phosphate content. These cultured carrot root cells produce recombinant proteins with plant pauci-mannose type glycans, which do not require additional processes to expose terminal mannose residues [56]. Unfortunately, the pauci-mannose glycans contain fucose and xylose residues normally not found in mammals, and the use of these enzymes as ERT poses significant concerns regarding immunogenicity.

A common issue associated with enzyme replacement therapy is immune responses caused by the administration of the ERT drug. In the case of Fabrazyme® 50-55% of patients receiving the therapy experience infusion reactions, and 88% of patients develop IgG or IgE immune responses [8, 13, 14]. There is currently a pegylated form of α -GAL, Pegunigalsidase alfa, which is undergoing clinical trials as a Fabry treatment [57]. This modified enzyme was designed with the goal of increasing plasma half-life, and decreasing immunogenicity. Results from a 1 year phase 1/2 clinical trial show the treatment maintained kidney function as measured by glomerular filtration rate and showed the clearance of Gb3 inclusion bodies from most tissue [57]. Though the number of patients in this study was limited, immune reactions to this treatment appear to be mild, with only a small portion of patients developing IgG responses, which decreased over the course of the trial. This approach represents one strategy to address serum stability and immunogenicity, but it is likely that alternate strategies will also be needed to address additional shortcomings identified in the current ERT drugs.

One of these additional shortcomings of ERT for Fabry disease that has come to light in recent years is a deficiency in the ability of the drug to preserve podocyte health. Unlike the near complete clearance of Gb3 inclusions in endothelial and mesangial cells

after brief periods of ERT, podocyte Gb3 inclusions linger, although at levels lower than those present prior to treatment [15]. In fact, podocyte foot effacement, a sign of podocyte damage, has been shown to progress over time in patients receiving ERT [16]. Considering the crucial role of podocytes in kidney function, this observation highlights a critical shortcoming of the current ERT drugs, one that should be prioritized during the design of the next generation of Fabry ERT drugs.

Pharmacological chaperone therapy

The second approved treatment for Fabry disease is pharmacological chaperone therapy (PCT). The principle guiding this strategy is that the addition of a ligand into the folding environment; (the endoplasmic reticulum), encountered by a disease mutant will bias the folding equilibrium towards the folded state (Fig 2). This principle was first tested in Fabry patients when Frustaci et al, began a study in which a patient with a cardiac variant of the late-onset form of the disease received infusions of galactose (1g galactose / 1kg body mass) every other day for two years. The patient experienced significant improvements in cardiac health and quality of life [58]. This initial success showed proof-of-concept, but was far from an ideal treatment option, as it required 4-hour infusions of large amounts of galactose every other day. However, this success led to pre-clinical trials with compounds possessing more desirable clinical properties. One of these compounds, 1-deoxygalactonojirimycin (DGJ), had previously been shown to enhance enzyme levels in Fabry lymphoblasts [59]. DGJ, also known as Migalastat, Galafold®, and Amigal, is an iminosugar analog of galactose that acts as a competitive inhibitor of α -GAL [60]. At near-neutral pH, like that of the ER, DGJ binds α -GAL with

a 190,000 fold increase in affinity compared to galactose [10]. This increased affinity is pH dependent allowing for dissociation of the complex upon entry into the acidic environment of the lysosome [10]. The high affinity of DGJ towards α -GAL coupled with the pH sensitive of this affinity as well as the pre-clinical results led to it being entered into clinical trials. After completion of multiple phase III clinical trials of PCT for Fabry disease, 1-deoxygalactonojirimycin (DGJ) was approved in the US in 2018 and was approved for the European Union in May 2016.

The studies outlined in this document aim to further the development of new more efficacious forms of both enzyme replacement therapy and pharmacological chaperone therapy. Outlined are strategies for the development of more specific pharmacological chaperones, a more renally efficacious enzyme replacement molecule, and a potentially novel mechanism for pharmacological chaperone therapy.

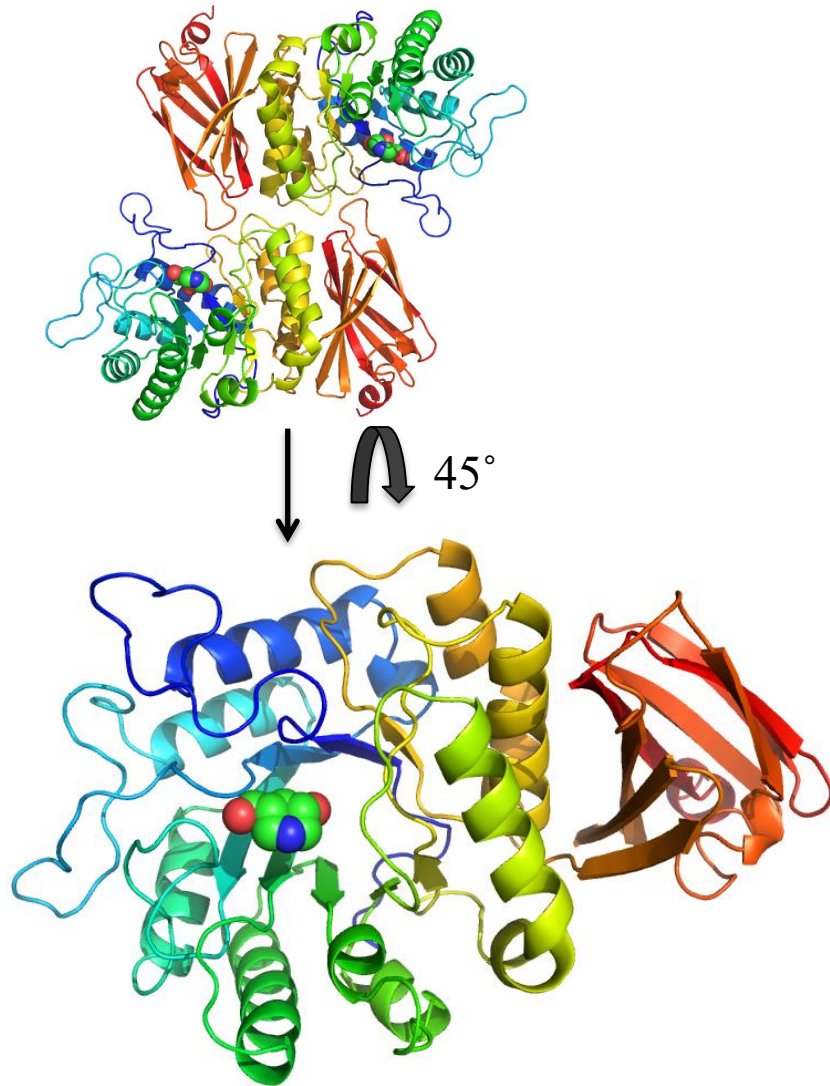


Figure 1. The structure of human α -GAL

Human α -GAL exists as a homodimer (shown in the top portion of the panel). One monomer (shown in the bottom of the panel) consists of a β -sandwich domain (shown in orange), and a catalytic TIM barrel domain. The active site of the monomer is shown bound to a ligand.

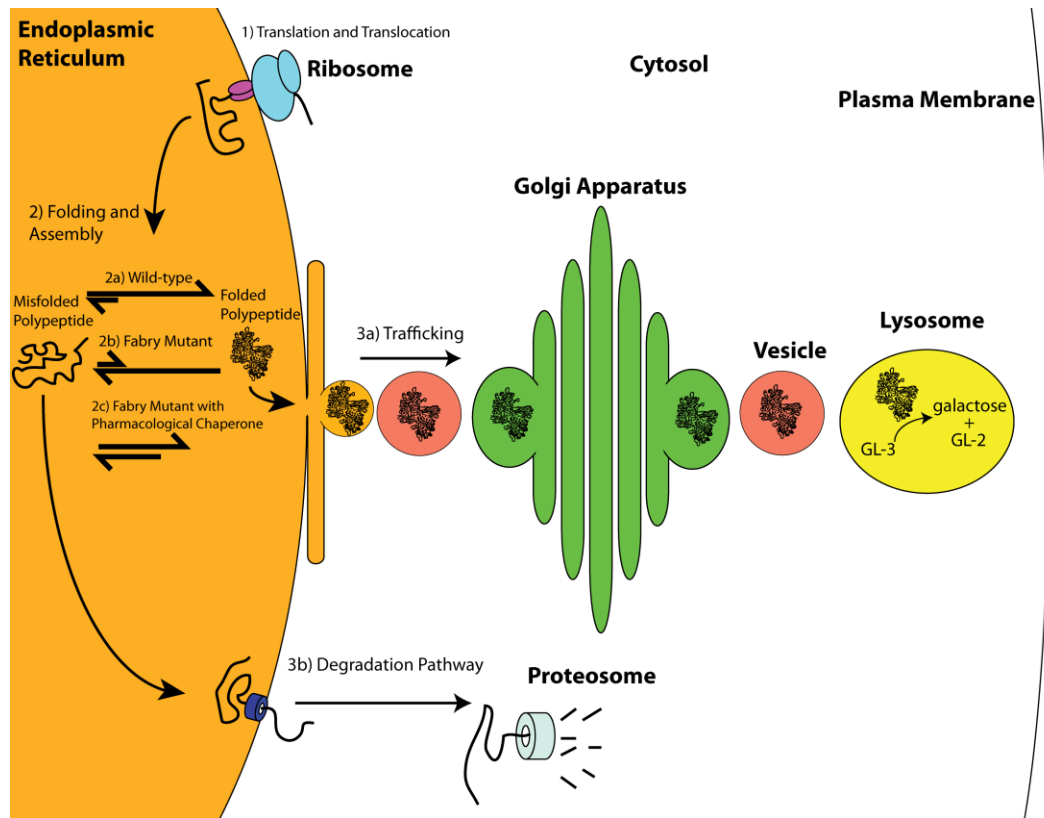


Figure 2. Overview of α -GAL trafficking

1. Translation initiates in the cytosol and translocation begins as the N-terminal signal sequence is exposed to translocation machinery.
2. The protein begins to fold as it emerges from the translocation pore and can either reach its native state or misfold.
3. The protein is trafficked out of the ER towards the lysosome if it has adopted its native fold or degraded if it has not.

CHAPTER 2: Interconversion of the Specificities of Human Lysosomal Enzymes Associated with Fabry and Schindler Disease

Abstract

The human lysosomal enzymes α -galactosidase (α -GAL, EC 3.2.1.22) and α -N-acetylgalactosaminidase (α -NAGAL, EC 3.2.1.49) share 46% amino acid sequence identity and have similar folds. The active sites of the two enzymes share 11 of 13 amino acids, differing only where they interact with the 2-position of the substrates. Using a rational protein engineering approach, we interconverted the enzymatic specificity of α -GAL and α -NAGAL. The engineered α -GAL (which we call α -GAL^{SA}) retains the antigenicity of α -GAL but has acquired the enzymatic specificity of α -NAGAL. Conversely, the engineered α -NAGAL (which we call α -NAGAL^{EL}) retains the antigenicity of α -NAGAL but has acquired the enzymatic specificity of the α -GAL enzyme. Comparison of the crystal structures of the designed enzyme α -GAL^{SA} to the wild-type enzymes shows that active sites of α -GAL^{SA} and α -NAGAL superimpose well, indicating success of the rational design. The designed enzymes might be useful as non-immunogenic alternatives in enzyme replacement therapy for treatment of lysosomal storage disorders such as Fabry disease.

Introduction

Lysosomal enzymes are responsible for the catabolism of metabolic products in the cell. Deficiencies in lysosomal enzymes lead to lysosomal storage diseases,

characterized by an accumulation of undegraded substrates in the lysosome. In humans, there are at least 40 different lysosomal storage diseases (including, for example, Tay-Sachs, Sandhoff, Gaucher, and Fabry diseases), each of which is caused by a lack of a specific enzymatic activity. Fabry disease is caused by a defect in the *GLA* gene, leading to loss of activity in the enzyme α -galactosidase (α -GAL,³ EC 3.2.1.22, also known as α -GAL A)[4]. The α -GAL enzyme cleaves substrates containing terminal α -galactosides, including glycoproteins, glycolipids, and polysaccharides. Defects in the *GLA* gene in Fabry patients lead to the accumulation of unprocessed neutral substrates (primarily globotriaosylceramide (Gb3)), which then leads to the progressive deterioration of multiple organ systems and premature death. Fabry disease is an X-linked inherited disorder with an estimated prevalence of 1 in 40,000 male births but may be highly underdiagnosed[1, 4].

The human gene most closely related to the *GLA* gene is the *NAGA* gene, encoding the enzyme α -N-acetylgalactosaminidase (α -NAGAL, EC 3.2.1.49, also known as NAGA and α -GAL B)[4]. The two genes are derived from a common ancestor[61], encoding proteins that share 46% amino acid sequence identity (Fig. 3A); they belong to glycoside hydrolase family 27 and clan D[62]. The α -NAGAL enzyme recognizes and cleaves substrates containing terminal α -N-acetylgalactosaminide (α -GalNAc) sugars, and (less efficiently) substrates containing terminal α -galactosides. Defects in the *NAGA* gene lead to the lysosomal storage disease Schindler disease, characterized by the accumulation of glycolipids and glycopeptides, resulting in a wide range of symptoms, including neurological, skin, and cardiac anomalies[4].

The only approved therapy for the treatment of Fabry disease is enzyme replacement therapy, where patients are injected with recombinant enzyme purified from mammalian cell expression systems[8, 13]. One problem with this treatment is that up to 88% of patients develop IgG antibodies against the injected recombinant enzyme[13, 14]. In patients who make no functional α -GAL enzyme, the recombinant α -GAL used in enzyme replacement therapy can be treated as a foreign antigen. Because the *GLA* gene is located on the X-chromosome, hemizygous males who inherit a non-functional copy of the gene have no second copy to establish immunotolerance.

Previously, we determined the three-dimensional crystallographic structures of the human α -GAL and α -NAGAL enzymes[40, 63]. As expected for two structures that share 46% sequence identity, the overall folds of the two enzymes are similar (Fig 3B and 3C). Each enzyme is a dimer where each monomer contains an N-terminal (β / α)₈ barrel with the active site and a C-terminal antiparallel β domain. The monomers of the two enzymes superimpose with an r.m.s.d. of 0.90 Å for α 378 carbons. The active sites of the two structures are highly similar, as 11 of the 13 active site residues are conserved. The only differences in the active sites of the two structures correspond to where the substrates are different: in α -GAL, the residues near the 2-OH on the substrate include larger glutamate and leucine, whereas in α -NAGAL, the larger 2-N-acetyl on the substrate interacts with the smaller serine and alanine residues on the enzyme (Fig 3B and 3C). As we and others have noted, in glycoside hydrolase family 27, the two residues primarily responsible for recognition of the 2-position of the ligand are both located on the same loop in the structure, the β 5– α 5 loop in the (β / α)₈ barrel domain[40, 63-67].

The similarities between the active sites of the enzymes in the family and the proximity of the two residues responsible for the recognition of the 2-position of the ligand led us and others to hypothesize that interconversion of the enzyme specificities would be possible[40, 63, 66-68]. We replaced two residues in the active site of α -GAL (E203S and L206A), leading to a new protein (α -GAL^{SA}) with the enzymatic specificity of an α -NAGAL enzyme. Additionally, we replaced two residues in the active site of α -NAGAL (S188E and A191L), leading to a new protein (α -NAGAL^{EL}) with the enzymatic specificity of an α -GAL enzyme. In this report, we show that the designed enzymes maintain the antigenicity of the original protein, but have the specificity of the other enzyme. X-ray crystallographic studies of the α -GAL^{SA} protein provide a structural basis for ligand specificity in the family of proteins.

Results

Biochemical Characterization of α -GAL^{SA} and α -NAGAL^{EL}

We expressed the four enzymes (α -GAL, α -GAL^{SA}, α -NAGAL, and α -NAGAL^{EL}) in stably transfected Tn5 insect cells and purified protein from the supernatant. SDS-PAGE analysis shows that the purified variant proteins migrate at the same size as their wild-type equivalents, 50 kDa for α -GAL and α -GAL^{SA} and 52 kDa for α -NAGAL and α -NAGAL^{EL} (Fig 3D). To test the antigenicity of the wild-type and variant proteins, we performed Western blots on all four proteins using antibodies against human α -GAL and human α -NAGAL. α -GAL and α -GAL^{SA} cross-react only with the anti- α -GAL antibody, whereas α -NAGAL and α -NAGAL^{EL} cross-react only with the anti- α -

NAGAL antibody (Fig 3E), indicating that the variant proteins retain their original antigenicity.

Enzymatic activity α -GAL^{SA} and α -NAGAL^{EL}

We tested the four enzymes against two synthetic substrates, pNP- α -Gal and pNP- α -GalNAc. In the wild-type enzymes, the larger active site of α -NAGAL allows it to bind and cleave both substrates, although the pNP- α -Gal less efficiently. The smaller active site of α -GAL allows for efficient catalysis of pNP- α -Gal, but no detectable activity on pNP- α -GalNAc due to steric clashes between the N-acetyl group on the 2-position of the sugar and the larger Glu-203 and Leu-206 side chains of α -GAL. Table 1 summarizes the enzyme kinetic data.

The variant enzymes α -GAL and α -NAGAL show the opposite substrate specificity compared with their starting wild-type enzymes. The α -NAGAL^{EL} enzyme shows the catalytic properties of wild-type α -GAL: it has no activity against the pNP- α -GalNAc substrate, but shows high activity against pNP- α -Gal. The α -GAL^{SA} enzyme has the catalytic properties of wild-type α -NAGAL: it has activity against pNP- α -GalNAc and reduced activity against pNP- α -Gal.

Because the enzymes in this family have overlapping substrate specificity, we used the ratio of the specificity constant k_{cat}/K_M for the two substrates as a measure of the ability of the enzymes to discriminate between the two related substrates. The substrate specificity ratios $(k_{cat}/K_M)_{pNP-\alpha-galNAc} / (k_{cat}/K_M)_{pNP-\alpha-gal}$ for α -GAL and α -NAGAL^{EL} are zero, because those enzymes show no activity toward the pNP- α -GalNAc substrate. The

substrate specificity ratios for α -NAGAL and α -GAL^{SA} are similar: 57.3 ± 10.3 for α -NAGAL and 42.9 ± 9.0 for α -GAL^{SA}, showing they have comparable ability to distinguish between the two substrates.

Comparison of the enzymatic parameters of α -GAL and of α -NAGAL^{EL} (the enzyme engineered to have α -galactosidase activity) shows that the K_m values of the two enzymes are similar (6.9 and 7.6 mM, respectively, for the pNP- α -Gal substrate). The engineered enzyme has a turnover number k_{cat} that is ~one-third that of the native α -GAL enzyme (13.7 s^{-1} versus 37.8 s^{-1} , respectively).

Comparison of the enzymatic parameters of α -NAGAL and of α -GAL^{SA} reveals that the K_m value of the engineered enzyme against the pNP- α -GalNAc substrate is 30-fold larger than that of the wild-type enzyme (21.0 and 0.68 mM, respectively) and 2-fold larger against the pNP- α -Gal substrate (49.1 and 27.5 mM, respectively). The turnover numbers k_{cat} of α -NAGAL and α -GAL^{SA} are similar against the pNP- α -GalNAc substrate (15.1 and 21.5 s^{-1} , respectively) and differ 9-fold against the pNP- α -Gal substrate (10.7 versus 1.2 s^{-1}). Overall the kinetic parameters show that, although the engineered enzymes are not as efficient as their wild-type counterparts, they have the same ability to discriminate among different substrates as their wild-type equivalents. In particular, the α -GAL^{SA} engineered enzyme is somewhat less effective at catalyzing the turnover of substrate compared with its wild-type equivalent, α -NAGAL.

Crystal structure of α -GAL^{SA}

To examine the structural basis for the reduced catalytic efficiency of the α -GAL three different ligands, N-acetylgalactosamine (GalNAc), galactose, and glycerol (Table 2 and Fig 4A–C). Superposition of α -GAL^{SA} and α -NAGAL by their $(\beta/\alpha)_8$ barrel domains results in an r.m.s.d. of 0.58 Å for 290 C α atoms in the domain. Remarkably, the superposition of the entire $(\beta/\alpha)_8$ domains shows that the GalNAc ligands and the active site residues superimpose nearly exactly (Fig 4D). Although 54% of the residues differ between α -NAGAL and α -GAL^{SA}, the ligand superimpose with an r.m.s.d. of 0.38 Å for the 15 atoms in the ligand.

Superposition of the $(\beta/\alpha)_8$ barrel domains of α -GAL^{SA} and α -GAL shows that the active site residues and ligands superimpose closely, except for Glu-203 and Leu-206 in α -GAL, which are replaced with Ser and Ala in α -GAL^{SA} (Fig 4E). The structure of α -GAL^{SA} with galactose bound shows a shift in the location of the catalytic nucleophile Asp-170 relative to its location in other structures in the family. The Asp-170 nucleophile shifts into the empty space produced by the reduction in size of the side chain in the E203S substitution. This shift affects the hydrogen bonding of Asp-170 to Tyr-134 and Tyr-207, and likely contributes to the reduced catalytic efficiency of the α -GAL^{SA} variant protein.

The crystallographic experiments used the cryoprotectant glycerol, which appeared in the active site of α -GAL^{SA}. We determined two structures of glycerol-soaked α -GAL^{SA} in space groups C222₁ and P2₁2₁2₁. Because each crystal has two monomers of the protein in the asymmetric unit, we have four crystallographically

independent active site complexes with glycerol. One of the four glycerol-soaked monomers shows significant changes in the active site: the glycerol binds in a different orientation, the Arg-227 side-chain rotamer changes, and the $\beta 6$ - $\alpha 6$ loop containing the catalytic acid/base Asp-231 shifts as well (Fig 4F). This large rearrangement in the active site is unique to α -GAL^{SA} among the glycosidase family 27 structures and may also contribute to the reduced catalytic efficiency of α -GAL^{SA}.

Discussion

The active site of human α -GAL is unable to accommodate 2-N-acetylated ligands due to steric clashes between the protein and the N-acetyl group on the ligand. Here, we have engineered α -GAL^{SA}, the first α -GAL enzyme capable of binding α -GalNAc ligands. In the reciprocal experiment, the engineered α -NAGAL^{EL} enzyme has lost all of its activity against α -GalNAc ligands and has enhanced activity against α -galactosides.

Enzyme replacement therapy

Enzyme replacement therapy can successfully treat Fabry disease. However, up to 88% of male patients develop an immune response to the injected recombinant enzyme, including both IgG- and IgE-based reactions[12-14]. The antigenicity of the glycoproteins used in enzyme replacement therapy for Fabry disease patients might limit the effectiveness of the treatment by reducing the amount of enzyme effectively delivered to the lysosomes. We envision the α -NAGAL^{EL} molecule would have little

immunogenicity in Fabry disease patients (who make typical amounts of α -NAGAL glycoprotein and are thus immunologically tolerant toward α -NAGAL). Consistent with this, heterozygous female Fabry disease patients (with one wild-type copy of the *GLA* gene) do not make the comparable immune responses as their hemizygous male counterparts against injected enzyme during enzyme replacement therapy[69, 70].

Although patients with the severe form of Fabry disease have little or no α -GAL enzyme activity, patients with the variant forms of Fabry disease can have from 5 to 35% of wild-type enzyme activity[5, 6]. This suggests that the threshold level of α -GAL enzymatic activity necessary to prevent Fabry disease symptoms is 100% wild-type activity. Although the enzymatic activity of the engineered proteins is lower than their wild-type equivalents, in enzyme replacement therapy, the reduced immunogenicity of the designed proteins might compensate for their reduced activity.

Sakuraba and colleagues have also tested this hypothesis by reporting a protein similar in design to α -NAGAL^{EL}, but expressed in Chinese hamster ovary cells, leading to a different glycosylation pattern[68]. When injected in a mouse model of Fabry disease, the mammalian-expressed protein led to a reduction in the amount of Gb3 in the tissues of the mouse. This further suggests that the designed enzyme might act as a useful tool for the treatment of Fabry disease.

Properties of the engineered enzymes

Western blotting with anti- α -GAL and anti- α -NAGAL antibodies shows that the former reacts only with α -GAL and α -GAL^{SA}, whereas the latter reacts only with α -NAGAL and α -NAGAL^{EL}. The sequence divergence between α -GAL and α -NAGAL is

sufficient to show no immunological cross-reactivity. Thus, engineering the active sites of α -GAL and α -NAGAL to have novel substrate specificities leads to new enzyme activities without altering antigenicity. This approach may be useful as a general strategy for protein-based therapeutics where reducing immunogenicity is an issue, such as Gaucher disease, where 15% of patients on enzyme replacement therapy develop IgG antibodies to the recombinant enzyme[71].

The diverse sequences in glycoside hydrolase family 27 (which includes human α -GAL and α -NAGAL) show high conservation of the active site residues, indicating strong evolutionary pressure on the active site. One modular component of ligand binding in the family is the recognition of the 2-position of the sugar ring. In this family, a single loop on the protein, the β 5– α 5 loop in the N-terminal domain, interacts with the substituent on the 2-position of the sugar. The similarity across the members of the family allowed us to interconvert the ligand recognition through substitution of two residues in the loop. The modularity of the loop is also seen in the structures of other members of the family, including the rice and *Hypocrea jecorina* α -GAL structures, where one turn of helix in the β 5– α 5 loop is replaced with a shorter loop and a longer β 5 strand[65, 72]. In those structures, Cys and Trp residues fill the space of the Glu-203 and Leu-206 residue of α -GAL or the Ser-188 and Ala-191 residues in α -NAGAL.

The newly engineered enzymes α -GAL^{SA} and α -NAGAL^{EL} are not as catalytically efficient as their wild-type equivalents. The structures of α -GAL^{SA} suggest that there is considerably more flexibility in the active site of the α -GAL^{SA} enzyme when compared with α -NAGAL, the enzyme with an identical active site constellation.

The glycerol-soaked structures of α -GAL^{SA} provide a structural explanation for the reduced catalytic efficiency of the engineered enzyme. When the larger Glu-203 and Leu-206 residues of α -GAL are replaced with smaller Ser and Ala residues in α -GAL^{SA}, the active site has more open space in it. This allows the Arg-227 side chain to move toward the space vacated by the shortening of the Glu-203 side chain, and the β_6 - α_6 loop containing Asp-231 moves toward the space vacated by the shortening of the Leu-206 side chain. The crystal structures indicate that the active site of α -GAL^{SA} is more dynamic than the active sites of the wild-type enzymes α -GAL and α -NAGAL, which might explain the reduced catalytic efficiency of the designed α -GAL^{SA} enzyme. Another indication of the higher mobility of the active site of α -GAL^{SA} appears in the glycerol-soaked α -GAL^{SA} structure, which has higher atomic B-factors in the rearranged β_6 - α_6 loop.

In this report, we describe the first instance of bidirectional interconversion of the enzymatic activities of two human lysosomal enzymes. Rational design of enzymatic function is a challenging task and generally requires large changes in the active site: for example, the classic case of conversion of trypsin activity into that of chymotrypsin required 11 substitutions in 4 sites on the protein[73]. In general, changing substrate specificity is easier than changing the reaction mechanism of an enzyme[74]. The family of lysosomal glycosidases might prove to be a fruitful target for further enzyme engineering, because many glycosidases use a similar mechanism with an arrangement of two carboxylates located on opposite sides of the glycosidic linkage to be cleaved.

In conclusion, we have shown the viability of a rational design approach to engineering new functionality into human lysosomal enzymes. This approach allows for encoding new enzymatic functions into existing protein scaffolds. By reusing existing proteins in new ways, our approach avoids the immunogenicity problems that are frequently seen in enzyme replacement therapies. This approach might also be used for a wide range of protein-based therapeutics when immunogenicity problems exist.

Experimental Procedures

Molecular biology

Human α -GAL and human α -NAGAL were expressed in stably transfected *Trichoplusia ni* (Tn5) insect cells as described previously[41, 63]. The α -GAL^{SA} variant was generated from the wild-type α -GAL construct using PCR-based site-directed mutagenesis (forward primer, 5'-G TAC TCC TGT **TCG TGG CCT GCT** TAT ATG TGG-3 (substitutions in bold), and reverse primer, 5'-CAC AAT GCT TCT GCC AGT CCT ATT CAG GGC-3), ligated, transformed into bacteria, and confirmed by sequencing. The α -NAGAL^{EL} variant was derived from the wild-type α -NAGAL construct by PCR (forward primer, 5'-CGG CCT CCC CCC AAG GGT GAA CTA-3 , and reverse primer, 5'-CC TTC ATA **GAG TGG CCA CTC** GCA GGA GAA-3), ligated, transformed into bacteria, and sequenced.

Cell transfection

Adherent Tn5 cells in SFX-Insect medium (HyClone) were transfected with plasmid DNA, and selection for stably transfected cells using 100 μ g/ml blasticidin (added after

48 h and every 48 h for 10 days) was carried out. Stable adherent cells were re-suspended in SFX medium for larger scale suspension cultures.

Protein expression and purification

One-liter cultures of stable cells secreting α -GAL^{SA} and α -NAGAL^{EL} proteins were grown to 3×10^6 cells/ml. The supernatant was clarified and concentrated by tangential flow filtration (Millipore Prep/ Scale) and exchanged into Ni²⁺ binding buffer (50 mM Na₃PO₄, pH 7.0, 250 mM NaCl, 20 mM imidazole, and 0.01% NaN₃). The retentate was loaded onto a Ni²⁺-Sepharose 6 Fast Flow column (Amersham Biosciences) and eluted with a gradient of 0 – 60% elution buffer (50 mM Na₃PO₄, pH 7.0, 250 mM NaCl, 250 mM imidazole, and 0.01% NaN₃). Eluate fractions were pooled, desalted, and concentrated before loading onto a Source 15Q anion-exchange column[63]. Fractions eluted by a linear salt gradient were screened by activity assays and by Western blots. Fractions containing pure protein were pooled and concentrated to 1.0 mg/ml for storage.

Kinetic assays

Hydrolysis of the synthetic substrates *para*-nitrophenyl- α -galactose (pNP- α -Gal) and *para*-nitrophenyl- α -N-acetylgalactosamine (pNP- α -GalNAc) (Toronto Research Chemicals) at 37°C were monitored by absorbance at 400 nm using an extinction coefficient of $18.1 \text{ mM}^{-1} \text{ cm}^{-1}$. 0.25–1.2 μg of enzyme in 10 μL of 100 mM citrate/phosphate buffer (pH 4.5) was added to 12 substrate concentrations (pNP- α -Gal from 0.1 to 50 mM, and pNP- α -GalNAc from 0.01 to 10 mM). Each minute for 10 min,

the sample absorbance was measured after adding 200 mM Na₃BO₃ buffer, pH 9.8. Error bars were determined from triplicate measurements by two experimenters for each data point. K_M , V_{max} , k_{cat} , and error bars were determined from a weighted fit of Michaelis-Menten hyperbola in KaleidaGraph. Substrate solubility limits prevented saturation in some experiments. Substrate specificity ratios for each enzyme were defined as $(k_{cat}/K_M)_{pNP-\alpha-GalNAc} / (k_{cat}/K_M)_{pNP-\alpha-Gal}$, indicating the preference of an enzyme for galactosaminide substrates.

Crystallization and x-ray data collection

Crystals of α -GAL^{SA} were grown as described for the D170A α -GAL variant[41]. Crystals were obtained from a 1:1 mixture of reservoir solution (12% polyethylene glycol 8,000, 0.1 M sodium cacodylate, pH 6.5, and 22 mM Mg(CH₃COO)₂) and 7.0 mg/ml protein in 10 mM Na₃PO₄, pH 6.5. Crystals were transferred stepwise into reservoir solution containing 200 mM ligand (GalNAc or galactose) and then into cryoprotectant solution (15% polyethylene glycol 8,000, 0.1 M sodium cacodylate, pH 6.5, 22 mM Mg(CH₃COO)₂, 20% glycerol, and 200 mM ligand). Crystals were flash-cooled in liquid nitrogen, and x-ray data were collected at 100 K at beamline X6A at the Brookhaven National Laboratory or at the microfocus beamline NECAT 24-ID-C at Argonne National Laboratory. X-ray images were processed using HKL2000 [75] and phased by molecular replacement in AMoRe [76] or by Fourier synthesis using human α -GAL coordinates (PDB: 3HG3)[41]. We selected the overall resolution limits based upon I/σ_1 criteria. Atomic models were built using the program O [77], with refinement in REFMAC5[76]. Ramachandran plots were computed using PROCHECK [78]. Coordinates were

superimposed using LSQMAN [79]. Figures were made in MolScript [80] and POVScript [81]. Coordinates and structure factors are deposited in the Protein Data Bank under accession codes 3LX9, 3LXA, 3LXB, and 3LXC.

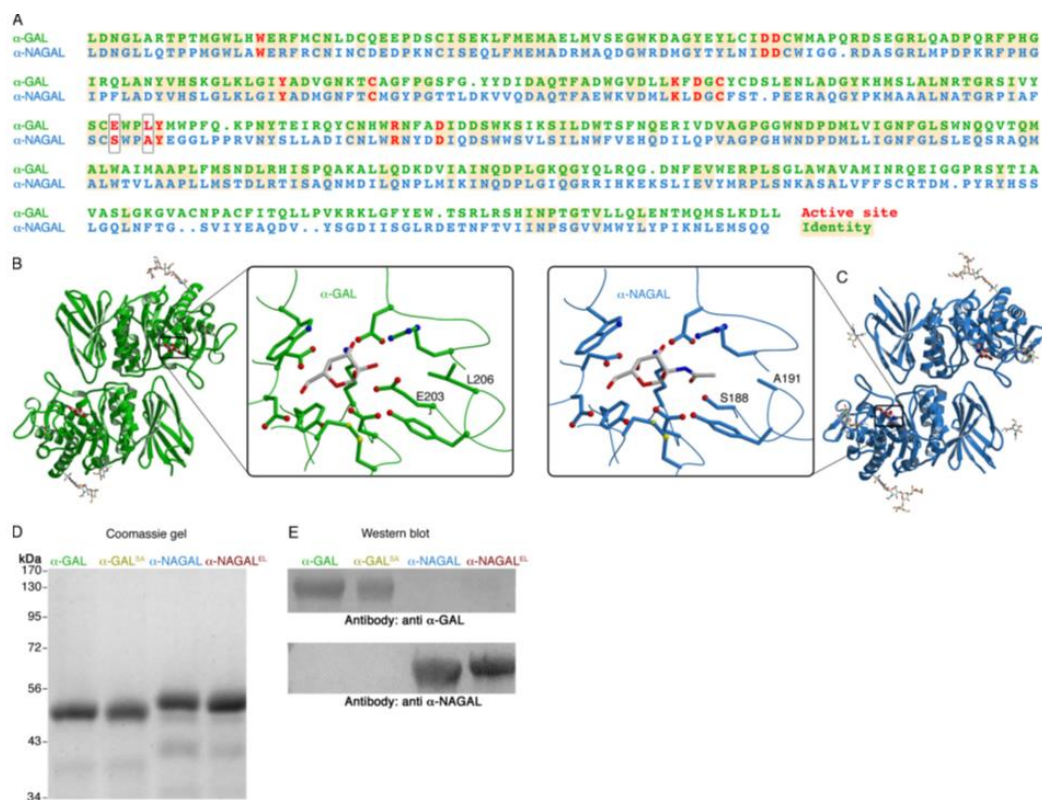


FIGURE 3. α -GAL and α -NAGAL structural and biochemical analyses

(A) Sequence alignment of the α -GAL and α -NAGAL proteins. Active site residues are red, and identities have yellow backgrounds. The two active site residues that differ are boxed. (B) and (C), ribbon diagrams of α -GAL (green) and α -NAGAL (cyan) with attached carbohydrates. Insets show the active sites of α -GAL and α -NAGAL with their catalytic products -galactose and α -GalNAc, respectively (gray). 11 of the 13 active site residues are conserved between the enzymes, although the overall sequence identity is 46%. The two residues that differ (Glu-203 and Leu-206 in α -GAL; Ser-188 and Ala-191 in α -NAGAL) select for the substituent on the 2-position of the ligand. (D), the four purified proteins are shown on a Coomassie-stained SDS gel. -GAL and -GALSA (with three *N*-linked glycosylation sites each) run smaller on the SDS gel than α -NAGAL and α -NAGALEL (with five *N*-linked glycosylation sites each). (E) Western blots of the four proteins, detected with polyclonal anti- α -GAL (top) and polyclonal anti- α -NAGAL antibodies (bottom). The variant proteins retain the antigenicity of the original proteins.

TABLE 1
Enzymatic parameters

| Enzyme | pNP- α -Gal | | | pNP- α -GalNAc | | |
|-------------------------------|--------------------|-----------------------|-----------------------------------|-----------------------|-----------------------------------|-----------------------------------|
| | K_m <i>mM</i> | k_{cat} s^{-1} | k_{cat}/K_m $mM^{-1} s^{-1}$ | K_m <i>mM</i> | k_{cat} s^{-1} | k_{cat}/K_m $mM^{-1} s^{-1}$ |
| α -GAL | 6.88 \pm 0.07 | 37.8 \pm 0.2 | 5.49 \pm 0.06 | | No activity detected ^a | |
| α -NAGAL ^{EL} | 7.58 \pm 0.07 | 13.7 \pm 0.1 | 1.81 \pm 0.02 | | No activity detected ^a | |
| α -NAGAL | 27.5 \pm 4.7 | 10.7 \pm 0.9 | 0.39 \pm 0.07 | 0.68 \pm 0.01 | 15.1 \pm 0.1 | 22.4 \pm 0.1 |
| α -GAL ^{SA} | 49.1 \pm 7.2 | 1.20 \pm 0.14 | 0.024 \pm 0.005 | 21.0 \pm 0.8 | 21.5 \pm 0.7 | 1.03 \pm 0.03 |

^a $k_{cat} < 0.01 s^{-1}$.

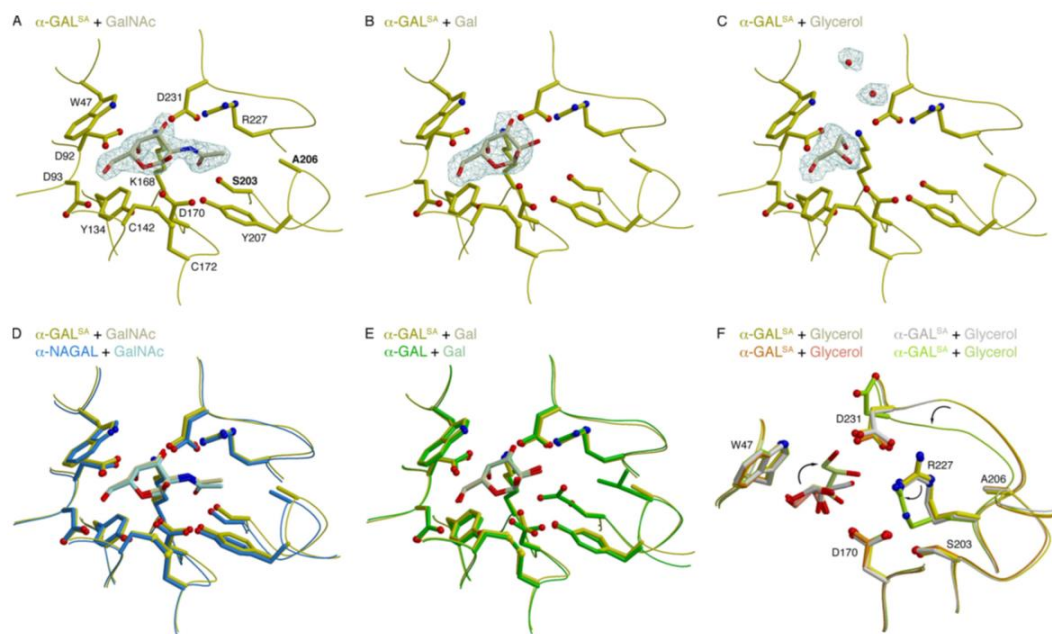


FIGURE 4. α -GAL^{SA} crystal structures.

(A–C) σ_A -weighted $2Fo-Fc$ total omit electron density maps of α -GAL^{SA} calculated in SFCHECK [36]. (A) GalNAc-soaked crystal contoured at 2.0σ . (B) galactose-soaked crystal contoured at 1.8σ . (C) glycerol-soaked crystal soaked at 1.0σ . Maps have a cover radius drawn around ligands and/or waters in the active site. Active site residues are labeled in A. (D) a superposition of crystal structures of the active sites of α -GAL^{SA} and α -NAGAL, each with α -GalNAc bound in the active site. When the structures are superimposed by their $(\beta/\alpha)_8$ barrels, the ligands superimpose nearly exactly. (E) a superposition of crystal structures of the active sites of α -GAL^{SA} and α -GAL, each with -galactose bound in the active site. (F) a superposition of the four monomers of glycerol-soaked α -GAL^{SA}, with glycerol bound in the active site. In one of the four structures (*green*), the glycerol binds in a vertical orientation, and shows differences in Arg-227 and the loop containing Asp-231 (*arrows*).

TABLE 2
 α -GAL^{5A} data collection and refinement statistics

| PDB ID | Ligand | | | |
|-------------------------------------------------------|-----------------------|-----------------------------------------------|-----------------------------------------------|-----------------------|
| | GalNAc 3LX9 | Galactose 3LXA | Glycerol 3LXB | Glycerol 3LXC |
| Data collection | | | | |
| Beamline | APS 24-ID-C | NLSL X6A | NLSL X6A | APS 24-ID-C |
| Wavelength (Å) | 1.07188 | 0.98010 | 0.98010 | 1.07188 |
| Space group | C22 ₁ | P2 ₁ 2 ₁ 2 ₁ | P2 ₁ 2 ₁ 2 ₁ | C22 ₁ |
| Resolution (Å) | 50-2.05 | 50-3.0 | 50-2.85 | 50-2.35 |
| (last shell) | (2.09-2.05) | (3.11-3.0) | (2.90-2.85) | (2.39-2.35) |
| Cell parameters <i>a</i> , <i>b</i> , <i>c</i> (Å) | 89.95, 139.49, 182.58 | 59.50, 105.85, 181.85 | 59.57, 106.92, 181.51 | 89.75, 139.77, 182.45 |
| No. of observations | 481,042 | 172,299 | 136,330 | 193,566 |
| No. of unique observations | 72,009 | 23,638 | 27,897 | 47,535 |
| (last shell) | (3,354) | (2,309) | (1,334) | (2,335) |
| Multiplicity | 6.7 | 7.3 | 4.9 | 4.1 |
| (last shell) | (3.6) | (7.2) | (4.9) | (4.0) |
| Completeness (%) | 99.4 | 100.0 | 99.3 | 98.6 |
| (last shell) | (93.3) | (99.9) | (99.9) | (98.7) |
| R_{sym}^a | 0.120 | 0.271 | 0.177 | 0.153 |
| (last shell) | (0.701) | (0.851) | (0.616) | (0.862) |
| I/σ_I | 18.6 | 7.6 | 9.1 | 12.3 |
| (last shell) | (1.6) | (1.8) | (2.2) | (2.0) |
| Refinement | | | | |
| $R_{\text{work}}/R_{\text{free}}^b$ (%) | 17.62/21.82 | 21.45/24.39 | 22.50/26.38 | 18.35/23.68 |
| No. of atoms | 7,057 | 6,557 | 6,657 | 7,032 |
| Protein | 6,241 | 6,303 | 6,320 | 6,241 |
| Carbohydrate | 257 | 218 | 170 | 227 |
| Water | 559 | 36 | 131 | 552 |
| Other | 0 | 0 | 36 | 12 |
| Average <i>B</i> , Å ² | 37.5 | 34.5 | 32.5 | 45.3 |
| Protein average <i>B</i> , Å ² | 36.3 | 33.9 | 31.8 | 44.4 |
| Ramachandran plot^c | | | | |
| Favored (%) | 91.4 | 89.5 | 89.5 | 89.7 |
| Allowed (%) | 8.0 | 9.6 | 9.6 | 9.5 |
| Generous (%) | 0.3 | 0.6 | 0.9 | 0.7 |
| Forbidden (%) | 0.3 | 0.3 | 0.0 | 0.1 |
| r.m.s.d. | | | | |
| Bonds (Å) | 0.0053 | 0.0150 | 0.0088 | 0.0081 |
| Angles (°) | 1.056 | 1.502 | 1.222 | 1.172 |

^a $R_{\text{sym}} = \sum_i \sum_h |I_{h,i} - \langle I_h \rangle| / \sum_i \sum_h I_{h,i}$, where $I_{h,i}$ is the *i*th intensity measurement of reflection *h* and $\langle I_h \rangle$ is the average intensity of that reflection.

^b $R_{\text{work}}/R_{\text{free}} = \sum_h |F_o - F_c| / \sum_h |F_o|$, where F_c is the calculated and F_o is the observed structure factor amplitude of reflection *h* for the working or free set, respectively.

^c Ramachandran statistics were calculated in PROCHECK.

CHAPTER 3: The Rational Design of a Potential Second Generation ERT Therapeutic with Increased Renal Efficacy

Introduction

Fabry disease is an X-linked disorder resulting from mutations in *GLA*, the gene encoding α -galactosidase (α -GAL)[2]. Deficient lysosomal α -GAL activity leads to the accumulation of substrates containing terminal α -galactosides, the most prominent being globotriaosylceramide (Gb3). This accumulation leads to functional impairment of numerous organ systems including the cardiac and renal systems. Fabry disease is characterized by chronic neuropathic pain, skin lesions, lenticular and corneal opacities, angiokeratoma, hypohidrosis, abdominal pain or bloating, as well as cardiac, renal, and cerebrovascular complications. The average lifespan of a person with the classical form of Fabry disease is 58 years, and the majority of these individuals receive a kidney transplant during their life [32]. The incidence rate of classical Fabry disease in males is estimated to be between ~1:40,000-1:60,000, but based on the screening of newborns it may be as high as 1:4,600 [1, 4, 82].

One current treatment for Fabry disease is enzyme replacement therapy (ERT). This treatment involves the intravenous injection of mammalian expressed α -GAL, which is taken into cells through extracellular receptors such as the mannose-6-phosphate receptor. ERT has been shown to be efficacious for the clearance of Gb3 deposits from the majority of afflicted tissues [8, 13, 49, 83]. However, the efficacy of ERT towards the stabilization of renal function appears to be limited. It has been demonstrated that ERT allows for the clearance of Gb3 deposits from some renal cells, including mesangial cells and renal endothelial cells, but the effectiveness of the treatment on podocyte depositions

has recently come into question. In a recent study evaluating long-term ERT on young patients, podocyte foot process effacement was observed in the majority of patients despite early administration of ERT [16]. One interpretation of these results is that despite being treated with ERT, these patients still progress towards more severe kidney complications.

The podocyte is a highly differentiated epithelial cell type that plays a vital role in the function of the kidney. Podocytes make up the most distal layer of the glomerular filtration barrier, which is responsible for filtering blood. The permeability of a molecule across the glomerular filtration barrier is dependent on its molecular radius and overall charge [84, 85]. Large molecules and negatively charged molecules are less able to permeate the barrier. α -GAL due to its highly negative charge and dimeric structure is likely not able to readily permeate the glomerular filtration barrier.

The human α -GAL homodimer possesses an overall negative charge due to a high number of negatively charged side-chains. Additionally, each monomer contains three modified N-linked glycans with varying amounts of sialylation and phosphorylation. The dimer measures approximately $75\text{\AA} \times 75\text{\AA} \times 50\text{\AA}$. The interface between monomers spans approximately 75\AA and buries 2200\AA^2 of surface area. Along this interface, amino acid F273 accounts for 12% of the total surface area (130\AA^2) (Fig 5). Sequence alignments of human α -GAL with other glycoside family 27 members reveal that this residue is conserved among dimeric members (Fig 6). The side-chain of F273 packs against multiple hydrophobic side-chains in the β -sandwich domain of the opposite monomer. In monomeric α -galactosidases, such as rice, this position is occupied by a glycine. The sequence alignment also reveals that there are conserved differences

between monomers and dimers at the position corresponding to 277 in the human form. In human and mice α -GAL this position is occupied by a tryptophan, and in monomers this position contains amino acids with smaller side chains. Both of these residues are on the α 7- β 7 loop of the (α / β)₈ domain.

Due to differential conservation of the positions F273 and W277 and the structural importance of these two positions in the dimer interface, we hypothesized that substituting these residues would result in a change of oligomeric state. By substituting these residues with glycines, we have engineered a monomeric enzyme. In this report, we show through the use of size-exclusion chromatography and in vitro membrane permeability studies, that the engineered enzyme retains a large fraction of the wild-type enzyme activity in vitro while existing in a monomeric state.

Results

α -GAL_{F273G+W277G} retains significant enzymatic activity

To assess the enzymatic properties of α -GAL_{F273G+W277G}, we used the synthetic substrate *para*-nitrophenyl- α -galactopyranoside (pNP-Gal) to determine K_M and k_{cat} values (Fig 7A). We then used these values to calculate the kinetic efficiency constant, also known as the specificity constant, of α -GAL_{F273G+W277G} for comparison to previously reported values (Fig 7B). We calculated the kinetic efficiency constant of α -GAL_{F273G+W277G} was determined to be $0.375 \text{ mM}^{-1} \text{ s}^{-1}$. This value, when compared to the previously reported value of $5.49 \text{ mM}^{-1} \text{ s}^{-1}$ for the wild-type enzyme, reveals an approximate 15-fold decrease in efficiency.

The observed enzymatic deficit of α -GAL_{F273G+W277G} when compared to the wild-type enzyme, is predominantly driven by an increased K_M . The calculated K_M of α -GAL_{F273G+W277G} of 64.14mM is approximately 10-fold higher than the wild-type enzyme, 6.88mM, whereas the k_{cat} of α -GAL_{F273G+W277G}, 24.05 s⁻¹ was only slightly lower than that previously reported for the wild-type enzyme towards the same substrate, 37.8 s⁻¹. It must be noted that due to the solubility limit of the substrate used for this assay, V_{max} and k_{cat} values cannot be reported with a high degree of certainty.

Overall, these findings indicate that the active site of the engineered enzyme is less readily saturated, but substrate, once bound, is turned over at a rate comparable to the wild-type enzyme. It reasons that these findings are the result of increased flexibility of active site components due to the removal of steric restraints imposed by the packing of the dimer.

α -GAL_{F273G+W277G} shows increased membrane permeability

In order to assess the ability of α -GAL to permeate a membrane we used the synthetic membranes in Amicon spin columns. Unfortunately, in vivo experiments would require an amount of enzyme exceeding our production capabilities, so we opted to perform in vitro experiments using a synthetic membrane. We applied the wild-type and mutant α -GAL to Amicon spin columns of varying molecular weight cut-offs (3kDa, 30kDa, 50kDa, 100kDa), then spun the columns. The enzyme activity and volume of both the retentate and filtrate fractions were measured and used to determine the enzyme units per fraction. The enzymatic units of each fraction were then compared to the total enzymatic units calculated per sample prior to centrifugation in order to determine

percent total enzymatic units of each fraction (Figure 8A). The % total enzymatic units of the filtrate and retentate fractions of the wild-type enzyme remained consistent across the 3 kDa, 30 kDa, and 50 kDa MWCO fractions. With wild-type protein and the 100 kDa membrane, 36% of the activity appeared in the filtrate, while a significantly smaller amount, 17.6%, appeared in the filtrate of the 50 kDa. The inverse result was seen in the retentate samples, with the 50 kDa sample retaining 75.5%, and 63.0% being retained in the 100 kDa sample. The samples from α -GAL^{F273G+W277G} show increasing activity in the filtrate and decreasing activity in the retentate across the 3 kDa, 30 kDa, 50 kDa, and 100 kDa membranes.

We then compared the ratio of % enzyme units in the filtrate to that of the retentate fraction for each sample (Fig 8B). As expected, the ratio of the two enzymes was similar for the 3kDa samples. Major differences between the two enzymes in the filtrate/retentate ratio were observed as the molecular weight cut-off of the membranes were increased. α -GAL^{F273G+W277G} showed significant permeation when applied to 30kDa membrane, this value increased when applied to 50kDa, at which point the amount of permeation appears to have reached a plateau as the 100kDa samples show a similar ratio. The wild-type enzyme did not show any change in permeation until the 100kDa membrane. The ratio of this sample was roughly equivalent to those of α -GAL^{F273G+W277G} applied to 50kDa and 100kDa membranes.

α -GAL^{F273G+W277G} displays altered mobility on size-exclusion columns

In order to determine the molecular weight of α -GAL^{F273G+W277G} we used size exclusion chromatography (Fig 9). Sizing standards were used to generate a molecular

weight standard curve. The elution volume of α -GAL_{F273G+W277G} maps to a molecular weight of 39.6 kDa. The elution volume of the wild-type enzyme was calculated to be 81.9 kDa. The estimated molecular weight of α -GAL_{F273G+W277G} is in good agreement with the molecular weight of an individual α -GAL monomer (43 kDa, based on the primary amino-acid and three N-linked pauci-mannose carbohydrates). Most importantly, the approximate 2-fold difference in molecular weight between the wild-type enzyme and α -GAL_{F273G+W277G} matches the expected difference between dimeric and monomeric forms of α -GAL.

Discussion

Human α -GAL exists exclusively as a dimer in nature. The extensive dimer interface buries 2200Å² of surface area. We have engineered a monomeric form of α -GAL by introducing 2 amino-acid substitutions and confirmed the quaternary state by size-exclusion chromatography. α -GAL_{F273G+W277G} was also shown to more readily permeate size-selective synthetic membranes. The engineered enzyme retains a significant portion of the wild-type enzyme activity.

α -GAL_{F273G+W277G} shows an approximately 15-fold decrease in enzymatic kinetic efficiency. The primary source of this decreased efficiency appears to be an increased K_M . This decreased ability to achieve active site saturation is possibly the result of increased active site flexibility. In particular, the β 1- α 1 loop is a likely culprit for this effect. In the dimeric form of the enzyme, this loop packs extensively against the β -domain of the opposing monomer and represents a significant portion of the dimer interface. This loop feeds directly into the active site cavity and is in close proximity to active site residue W47. Reinforcing this explanation is the fact that this loop exists only in the dimeric members of glycoside hydrolase family 27, as shown in figure 1. It should also be noted that this loop contains numerous negatively charged side-chains, and its removal may facilitate greater permeability across the glomerular filtration barrier. The β 1- α 1 loop is a prime target for future research aimed at enhancing the enzymatic efficiency of α -GAL_{F273G+W277G} derivatives.

Although the enzyme was shown to have a reduced molecular radius, the effect these mutations have on the ability of the enzyme to be trafficked to podocytes remains unanswered as that question exceeds the scope of this study. The two predominant factors limiting a protein's permeability across the glomerular filtration barrier are charge and molecular radius. We have effectively targeted one of these variables, and future studies may couple the α -GAL_{F273G+W277G} scaffold with mutations to modify the charge of the enzyme to generate a superior renal ERT molecule.

Despite the success of ERT in removing substrate deposits from the majority of afflicted tissue types, the renal efficacy of the treatment has come into question [15]. Patients receiving ERT to treat Fabry disease have been shown to present progressive renal impairment (in the form of podocyte foot effacement) despite receiving treatment [16]. Due to the nature of podocytes and their relative position in the glomerular filtration barrier, deficient ERT delivery is a possible explanation for the progression of renal complications. α -GAL_{F273G+W277G} may present an advantageous scaffold for an improved treatment for the renal complications of Fabry disease. Despite the lower catalytic efficiency of the enzyme, it has previously been established that 100 percent wild-type activity is not required to prevent the onset of disease manifestations. A combination of the current ERT treatment supplemented with an α -GAL_{F273G+W277G} enzyme may be a highly effective cocktail as the treatment would retain the efficacy of the current drugs with the added benefit of an increase in renal delivery.

Experimental Procedures

Molecular biology

Human α -GAL was expressed in stably transfected *Trichoplusia ni* (Tn5) insect cells as described previously. The α -GAL_{F273G+W27G} mutant was generated using the Phusion site-directed mutagenesis strategy, kinase-treated, ligated, transformed into DH5 α cells, then confirmed by sequencing. The primers used to generate the mutations were as follows; Forward: 5'-GAT TGG CAA CGG TGG CCT CAG CGG GAA TCA GCA AG-3' Reverse: 5'-ACT AAC ATA TCT GGG TCA TTC CAA CCC CCT GGT CC-3'.

Cell transfection

Adherent Tn5 cells grown in t-25 flasks with serum free IPI-41 media were transfected with 1.8 μ g of plasmid using Cellfectin 2. The cells were selected with 100 μ g/mL blasticidin for 14 days (fresh selective media was applied every 48 hrs). The cells were allowed to recover in non-selective media for 48 hours and were then suspended in order to seed larger cultures.

Protein expression and purification

Stably transfected cell cultures were grown in 1L cultures and were harvested once a high density was reached (~4 X 10⁶ cells/mL). The supernatant from these cultures was clarified using centrifugation. The clarified supernatant was buffer exchanged into nickel binding buffer (50 mM sodium phosphate pH 7.0, 250 mM NaCl, .01% azide) using a tangential flow cartridge. The buffer exchanged supernatant was applied to a 5mL

Nickel FF column and eluted using a 0-30% linear gradient of elution buffer (50 mM sodium phosphate pH 7.0, 250 mM NaCl, 400mM imidazole, .01% azide). The fractions containing α -GAL/ α -GAL_{F273G+W277G} were identified using activity assays and coomassie blue staining. The fractions containing α -GAL/ α -GAL_{F273G+W277G} were pooled, concentrated, desalted, and applied to an Uno Q1 column. The protein was eluted using a linear salt gradient and the fractions containing pure α -GAL/ α -GAL_{F273G+W277G} were concentrated and stored in 50mM Bis-tris pH 6.0, 100mM NaCl at a concentration of ~1mg/mL.

Size exclusion chromatography

All sizing exclusion experiments were performed using a Superdex 75 5/150 GL column. The column was pre-equilibrated with a 50mM Bis-tris pH 6.0, 150mM NaCl, .01% azide buffer prior to each run. A standard curve was generated using molecular weight standards from Sigma.

Membrane permeability assay

Amicon filters were pre-washed with a 1% bovine serum albumin in 20mM bis-tris pH 6.0, and spun for 20 minutes at 13100 g. The filters were then loaded with 400 μ L of a solution of α -GAL, 1% BSA, 20mM bis-tris pH 6.0. The sample was then spun for 15 minutes at 13100 g. The volume of the retentate and filtrate were measured, and 10 μ L of each component were applied to 20 μ L of 20mM pNP- α -Gal in 100mM citrate/phosphate buffer at pH 4.5. The reaction was incubated at 37°C for 10 minutes. 10 μ L of the reaction was then applied to 190 μ L of quenching buffer (200mM sodium

borate pH 9.8). Absorbance readings of the quenched reactions were taken measuring absorbance at 400nm.

Enzyme kinetics

Enzymatic characterization of α -GAL_{F273G+W277G} was performed by measuring the rate of hydrolysis of the synthetic substrate *para*-nitrophenyl- α -galactopyranoside (pNP- α -Gal). The enzyme was mixed with pNP- α -Gal at concentrations ranging from 25 μ M-50mM and incubated at 37°C for ten minutes with samples being withdrawn every minute and quenched in 200mM sodium borate pH 9.8. The samples were applied to a plate reader measuring absorbance at a wavelength of 400nm. Using an extinction coefficient of 18.1mM⁻¹ cm⁻¹ the concentration of cleaved *para*-nitrophenyl was determined, and error bars were calculated using triplicate measurements for each data point. The K_M , V_{max} , and error bars were calculated by a weighted Michaelis-Menten hyperbola using Kaleida-Graph. Due to substrate solubility limitations substrate saturation was not achieved for the mutant, and as a result substrate specificities, k_{cat}/K_M were used as a comparison to the wild-type enzyme.

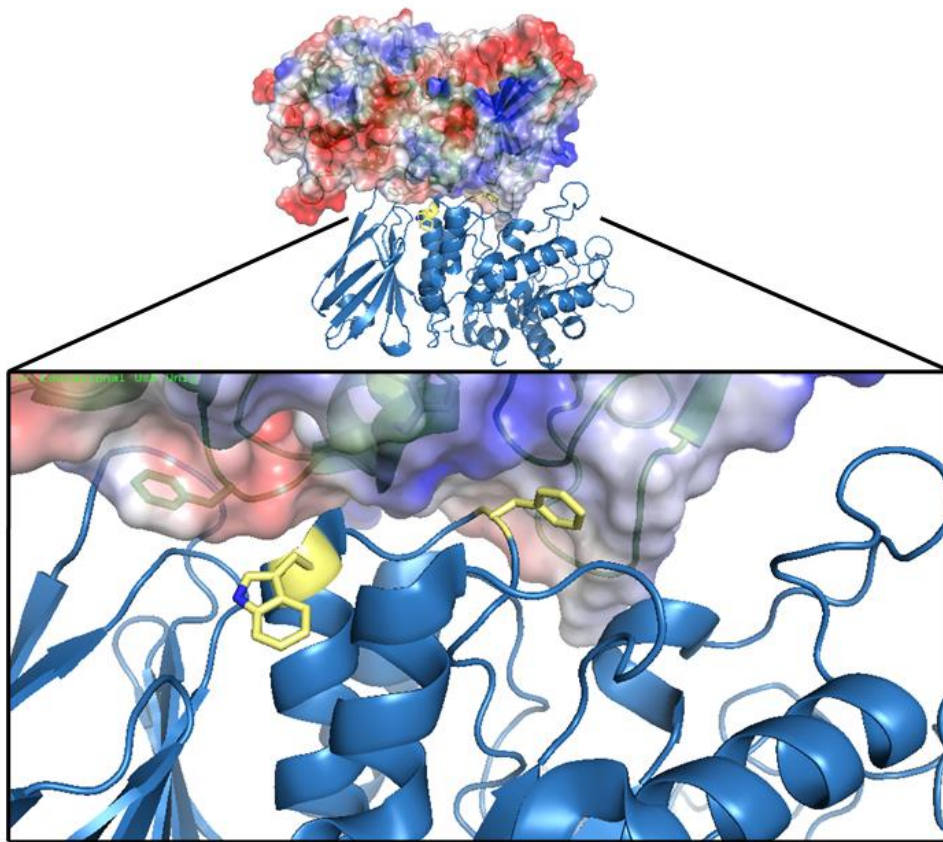


FIGURE 5. Dimer interface of human α -GAL

One monomer is presented a surface representation while the other is shown as a cartoon. F273 and W277 are shown in yellow. F273 packs extensively against hydrophobic residues of β -domain of the opposing monomer. W277 has the indole side-chain solvent exposed.

```

H.sapiens
M.musculus
A. Niger
Coffee
Rice
S.cerevisiae
1 -- MQLRNPEHLHGCLALRFLALVSMIDIPGARALDNG
1MAHKLLSRDRLVLCCLALCPALFVMSILGKRALDNG
1-----MRWLLTSSALLVPAALVVRP--DG
1-----MKSPGTEBYTRRSLLANG
1-----FENG
1-----MFAFYFTIACISLKGVFVSPSYNG
1-----LGLTPQMGWMSWHFYCG
1-----VSEQLLLDTRADRIIDLGLKDMGVK67

H.sapiens
M.musculus
A. Niger
Coffee
Rice
S.cerevisiae
88 YL CIDDCWMAAP--QRDSEGRLOADPQRFPHGIRQLANYVHSGKGLKLGIVADVGNKTCAGF--PSPFGYYDIDAOQTFADWGVDLLKFDGICYD-----1275
90 YL CIDDCWMAAP--EDSKRGLRLOADPQRFPSGIKKHLANYVHSGKGLKGIYADVGNKTCAGF--PSPFGSYDIDAOQTFADWGVDLLKFDGCHD-----1277
65 YINIDDCWMSYKSGRRITTRIIIDPPDFPNRISGVADQVHALGLKLGITYSSAGLTTCSAGT--PALSGLYEEIDAQSF AEWGIDYIKYDNCGVPINLTDQYTYCVDPDSTDGSNYPN1276
62 YINIDDCWALE--NDSQGNLVPKSGTFPSGIKALADYVHAKGLKLGITYSAGSQTCSNK--MPGSLDHEEQDVKTFASWGVDTLKYDNCNN-----1280
47 YVNIIDCWALE--SDSQGNLVPVNRQTFPSGIKALADYVHAKGLKLGITYSAGSQTCSNK--MPGSLDHEEQDVKTFASWGVDTLKYDNCND A-----1285
68 YIILDDCWS--GRSDGFLVDEQKFPNGMGHVADHLHNSFLFGHYSSAGETYCAGY--PQSLGREEDAOFFANNRVDYLYKYDNC-----YAKR124

H.sapiens
M.musculus
A. Niger
Coffee
Rice
S.cerevisiae
126 -----SLENLADGGYKHMSLALNRTGRSIVYS--CEWPLVYMWPFKQKPNYTEIRQYCNHWNRNFADIDDSWKS-----IKSLDWT SFNQ250
128 -----SVVSLNENGYKYMALALNRTGRSIVYS--CEWPLVLRPFHKPNYTDIQYYCNHWNRNFADIDDSWKS-----IKNILSWT VVYQ252
127 GT CVNLTDAAPQGVOWATSTTAKRYQRHRDALLSVNRITLIVSLCDWQG-----ADVNAMNGNATGN SWRMSGDIDTWSR-----IAEIANE-----N28
151 -----NISPKE--YPIMSKALLNSGRSIFFSICEWGG-----EDP--ATWAKKEVNGNSWRITGDIDANNWS-----MTSRADH-----N215
136 -----GRSVMER--YTRMSNLMKTYGKNIFFSICEWGG-----ENP--ATWAKGRNGNSWRITGDIDANNWS-----MTSRADH-----N200
155 GQF G-----TPEISYHRYKAMSDALNKTGRPIFYISLCNWGO-----DLTFWMSGIANSWRMSGDVYAEFTRPDSRCPDGD EYDCKYAGFHCSIMN1LNK----A246

H.sapiens
M.musculus
A. Niger
Coffee
Rice
S.cerevisiae
251 ERIVDVVYAGP GGWMDPDMLVIGNFGLSNNQVYVQMALWAIMAAPL FMSNDLRHISQAKALLQDKDVI AINDPP-LGRQGYQLRQ-----GDHFEVWERPLSGLAWAVAN353
253 KEIVVYAGP GGWMDPDMLVIGNFGLSNDQVYVQMALWAIMAAPL LMSNDLRQISSQAKALLQDKDVI AINDPP-LGRQGYCFRR-----ENHIEVWERPLSNLAWAVAN355
259 SF LMNYANFVWGYPPDMLVGNGNLTL PENRAHFALWAMHKAPL IIGTPLSDIDTSHLTL SNKPLLFHQDAVIGRPAAYKMGVNPDPWTFDP EHPAEVWYS GPTSSGEVFLM372
216 DKWASVYAGP GGWMDPDMLEVGN GMITTE YRSHFSI WALAKAPL IIGDIRSMGATFQLLSNAEVI AVNQDK-LGVQGNKVKT-----YGDLEWVAGPLSGKRVAVAL318
201 DQWAAVYAGP GGWMDPDMLEVGN GMSAEYRSHFSI WALAKAPL IIGDIVRSMGQTKNILLSNEVI AVNQDS-LGVQGNKVKQ-----DNGL EWVAGPLSNNRKAVAL303
247 APHGNVYAGP GGWMDPDMLEVGNLITDDE EKAHFSMWAVKSPLL IIGD VNNLKASSYSIYSQASVIAINDOS-NGIPATRVWRYVSDTDEYGOGEIQWMSGPLDNGDQVVAL359

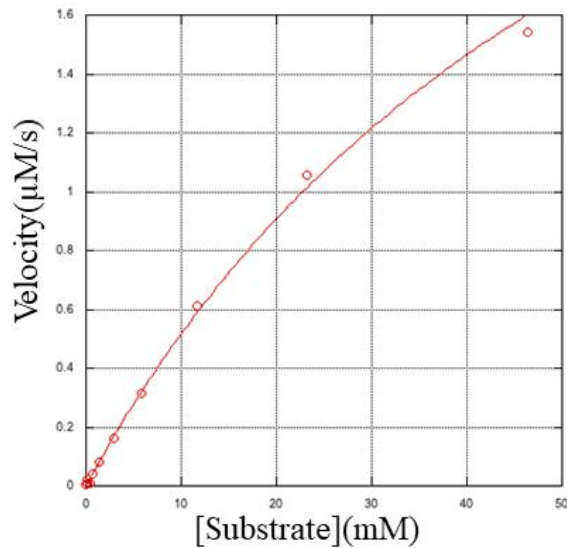
H.sapiens
M.musculus
A. Niger
Coffee
Rice
S.cerevisiae
354 INRQEI GGPRSYTIAVASI GKGVACNP-----ACFIITQLLVYKRRKLGFEYWTSHLRSHINPTGTVLLOLENTIMQMSLDLL-----429
356 RNLQEI GGPCPTTIOISSI GRKACNP-----GCIITQLLPEKVVHGFYEWTLTKLRTVNPSPGTVLFRLE R-----421
373 LNSEGEVKTRSAAVMEVPELKDQGTK-----KNSKKEKRGKVTYAWTGDGLGCVKQDYEYVQLQANDVAALVVGQGC-----443
319 VNR-----GSSYATITAYMSDVGLPSTA-----VYNARDLWMSHT-----ENSKVQGISAVVDADSDKMTVLLTPQ-----378
304 VNR-----QSYQATITAHWSNIGLAGSY-----AVTADLWMSH-----SFAAQGDISAVVADPDKMKTVLLTPN-----362
360 LMGGSVSRPMTLEEIFDMSLGSKRLTSTWIDLWNRVDNSTASAILGRNKATGILYNATEOSYKXDGLSKNDTRLFGOKIGSLSPNALINTTVP AHGIAFYRLRPPSS-----471

```

Figure 6. Alignment of Representative dimeric or monomeric α -galactosidase amino sequences

Amino acid sequences of α -galactosidases from various species were chosen to represent either monomeric and dimeric enzymes. The sequences of dimeric enzymes are shown in blue with monomers being shown in green.

A

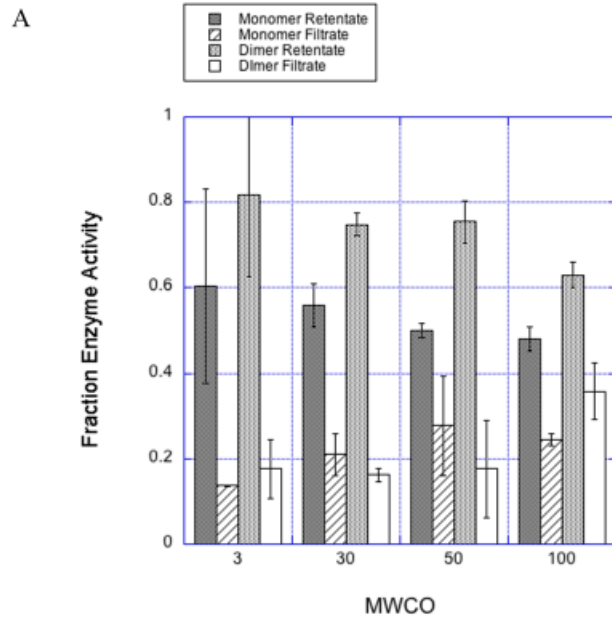


B

| Enzyme | K_M (mM) | k_{cat} (s^{-1}) | k_{cat}/K_M |
|--------------------------------------|------------------|------------------------|-----------------|
| WT α -GAL | $6.88 \pm .07$ | $37.8 \pm .2$ | $5.49 \pm .06$ |
| α -GAL ^{F273G+W277G} | 64.14 ± 1.89 | $24.05 \pm .59$ | $0.375 \pm .01$ |

FIGURE 7. Michaelis-menten kinetics of α -GAL^{F273 + W277}

(A) Michaelis-Menten plot of α -GAL^{F273 + W277}. pNP-Gal was used as a substrate and product, pNP, was measured by OD at 400nm. The results of this assay were used to determine K_M and k_{cat} , these values are shown in (B) as well as those previously published for the wild-type enzyme.



B

| Ratio of % Enzyme activity of Filtrate to Retentate | | |
|-----------------------------------------------------|--------------------------------------|-----------|
| MWCO | α -GAL ^{F273 + W277} | Wild-type |
| 3 | .23 | .21 |
| 30 | .38 | .22 |
| 50 | .55 | .23 |
| 100 | .51 | .57 |

FIGURE 8. Enzyme permeability assay using synthetic size-selective membranes

Wild-type and α -GAL^{F273 + W277} enzymes were applied to Amicon Ultra .5mL spin cartridges of varying molecular weight cut-offs. Total enzyme units were determined prior to concentrating and the fraction enzyme activity was calculated for the retentate and filtrate post concentration as shown in (A). The ratio of percent total enzyme units in the filtrate to retentate was calculated for both the wild-type and α -GAL^{F273 + W277} and is summarized in the table (B).

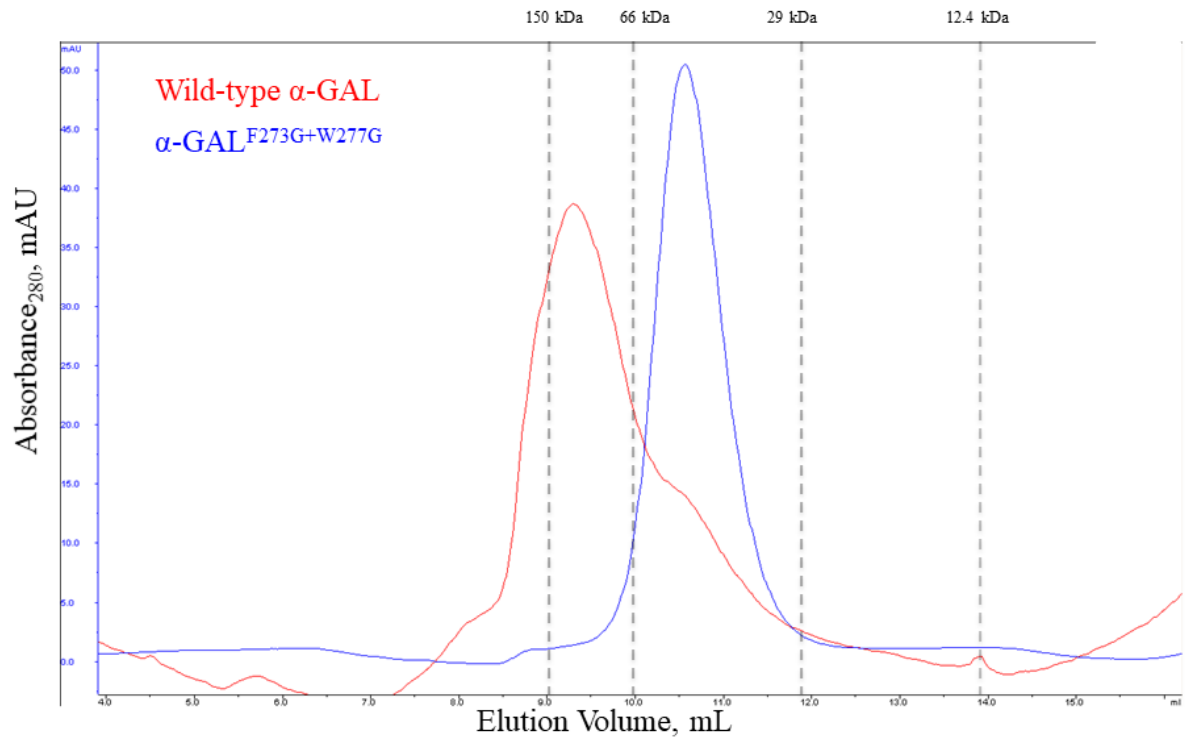


FIGURE 9. Size-exclusion chromatography elution profiles of wild-type α -GAL and α -GAL^{F273 + W277}
 Overlay of the elution chromatogram of wild-type α -GAL (red) and α -GAL^{F273 + W277} (blue). Elution peaks of sizing standards are represented by dashed lines.

CHAPTER 4: α -Homo Variants of Deoxygalactonojirimycin and 2-acetamido-1,2-deoxy-D-galactonojirimycin Show Increased Specificity Towards Enzymes Associated with Fabry and Schindler Disease

Introduction

Fabry disease is an X-linked disorder resulting from mutations in the *GLA*, the gene encoding α -galactosidase (α -gal) [4, 86]. Deficient levels of α -gal activity results in accumulation of glycosphingolipids, primarily globotriaosylceramide (GL-3), in the cells of a wide array of tissues, and as these deposits accumulate, normal tissue function becomes impaired [4, 86, 87]. Currently, the two FDA-approved strategies for treatment of Fabry disease are enzyme replacement therapy (ERT) and pharmacological chaperone therapy (PCT).

ERT involves the infusion of recombinant human α -gal into patients [8, 83]. First approved in 2001 for use in Europe then in 2003 in the United States, enzyme replacement therapy has been shown to clear GL-3 deposits from critical organ tissue [8, 13, 49]. Despite the demonstrated efficacy of ERT, there are a number of issues associated with this treatment. Most obviously, the need for biweekly intravenous infusion causes significant disruption in a patient's lifestyle. The treatment is also extremely expensive, costing approximately \$200,000-\$300,000 per year [88]. Additionally, the drugs are also unstable at neutral pH, resulting in a short circulating serum half-life [15, 83]. Limited cellular uptake has also been reported for a number of cell types such as cardiomyocytes, glomerular podocytes, and the central nervous system as a whole [15, 83]. When considering all of these issues the prospect of an alternative treatment led to the advancement of PCT.

PCT for Fabry disease involves the stabilization of α -gal by a small molecule, 1-deoxygalactonojirimycin (DGJ), also known as migalastat, AT1001, GR181413A, and Galafold™, that binds and stabilizes wild-type and mutant forms of the enzyme [43, 59]. PCT as a monotherapy received FDA approval in 2018 for a subset of amenable patients (348 of the 841 known Fabry genotypes as of 2018) [11]. This limited patient pool is the result of the mutations possessed by non-amenable patients. The *GLA* genes carried by these patients either have frameshifts, large deletions or insertion, truncations, or other mutations that result in a product that is unable to bind DGJ or is too unstable to clear the endoplasmic reticulum-associated degradation pathway even after chaperone binding [89]. PCT has been shown to increase α -gal activity in cultured cells from amenable Fabry patients, and this carried over to clinical trials with patients treated showing increased α -gal activity, and the reduction of GL-3 deposits in multiple cell types, tissues, and organs. [90-92]. A major advantage of DGJ is that it is administered orally every other day, and does not require intravenous infusions. In addition to its use in the FDA-approved monotherapy, combination therapy using DGJ; and an enzyme-replacement therapeutic has been suggested [93]. In combination therapy, DGJ would act to stabilize the ERT drug, resulting in an increased circulating serum half-life. The recent success of DGJ as disease treatment has led to the search for new targets for pharmacological chaperones, as well as the design of new more effective iminosugar derivatives.

DGJ is an iminosugar analog of galactose, the product of the reaction catalyzed by α -GAL (Figure 10). The compound is a potent reversible competitive inhibitor of α -GAL. It has been shown that an interaction between the heterocyclic nitrogen of DGJ and

the carboxylate of active site D170 is responsible for the 190,000 fold increase in affinity compared to α -galactose, the natural product α -GAL catalyzed reactions [10].

Additionally, it has been shown that this affinity is pH-dependent with a significant decrease in affinity at acidic pH. These characteristics define DGJ as a successful PCT therapeutic, as the compound binds with the highest affinity in the endoplasmic reticulum and serum; and a lower affinity in the acidic environment of the lysosome, where it is required to cleave substrate.

DGJ has also been suggested as a potential treatment option not only for Fabry disease, but also for the treatment of Schindler/Kanzaki disease [94]. Schindler/Kanzaki disease results from deficient lysosomal α -N-acetylgalactosaminidase (α -NAGAL) activity [95, 96]. α -NAGAL and α -GAL are evolutionarily linked, having evolved from the same ancestral gene [40]. As a result, the two enzymes share the same fold and have highly similar active sites (11 of 13 active residues are conserved). The two different active site residues are responsible for substrate recognition of groups attached to the C2 position of the substrate sugar. The active site of α -NAGAL is capable of binding a larger N-acetyl group at this position, whereas in the case of α -GAL, steric constraints prevent these compounds from binding. As a result, α -NAGAL has less specificity towards substrates/ligands than α -GAL, as illustrated by the ability of α -NAGAL to catalyze the hydrolysis of both terminal α -galactosides and α -N-acetylgalactosaminides, as well as its high affinity for DGJ, with a 1.6 μ M K_i [63]. It has been shown that substituting an N-acetyl group for a 2-hydroxyl group of DGJ was sufficient to obtain α -NAGAL specificity, as demonstrated by 2-acetamido-1,2-deoxy-D-galactonojirimycin

(DGJNAc) [94]. This modularity allows for a simple approach to convert prospective Fabry PCT molecules to those specific towards Schindler/Kanzaki disease treatment.

DGJ, at concentrations similar to those used in the clinical trials, has been shown to inhibit human β -galactosidase (GLB1). In fact, DGJ is a potent inhibitor of GLB1 with a measured K_i of 15 μ M [94]. The low specificity of DGJ results from the lack of an axial group on the one position of the iminosugar rings. The active sites of both β - and α -selecting glycosidases have evolved the capability to differentiate between β - and α -linked galactosides. With no 1-hydroxyl group, DGJ suffers from lack of specificity. It is also possible that some binding affinity is sacrificed by using the 1-deoxy forms of the iminosugars. For example, crystal structures of both α -GAL and α -NAGAL with natural products bound (α -galactose and α -N-acetylgalactosamine respectively); show hydrogen bonding between the 1-hydroxyl groups of the ligands and the carboxyl of active site side-chains D231 in α -GAL or D217 in α -NAGAL that are not possible within the deoxy iminosugar analogs.

In this study, we investigated the effect axially linked groups on the C1 position of galactose or N-acetylgalactosamine analogs have on binding characteristics to human α -GAL or α -NAGAL. Using galactonojirimycin (GJ), we investigated the importance of hydrogen bonding between the 1-hydroxyl of a galactose analog and the carboxylate of α -GAL active site residue D231 in terms of binding affinity. Additionally, we investigated the effect of α linked groups to C-1 on either Gal-derived or GalNAc-derived iminosugars on relative affinity, binding mode, and specificity. Using structural and enzyme inhibition assays, we determined that additions to the C-1 position of

iminoalditols have a slightly negative effect on binding affinity towards either α -GAL or α -NAGAL, but allow for greater specificity towards their intended targets.

Results

Inhibitory properties of GJ, α -HGJ, and α -HGJNAc

In order to assay the various compounds in this study, we measured their binding capabilities using enzyme inhibition assays. The measured K_i of GJ for α -GAL was approximately 2-fold higher than that of DGJ (Fig 11). The slightly weaker binding of GJ to α -GAL may result from the GJ sample consisting of two pseudoanomers. The difference in binding affinity is small and corresponds to an approximate $\Delta\Delta G$ of 0.5 kcal/mol compared to DGJ. This increase in ΔG compared to DGJ indicates that the hydrogen bond gained between α -GJ and D231 contributes little to binding affinity.

More significant changes were observed when comparing the binding of α -homo-2-acetamido-1,2-dideoxy-galactonojirimycin(α -HGJNAc) and α -homo-1-deoxy-galactonojirimycin (α -HGJ) to that of DGJ and DGJNAc. α -HGJ binds both α -GAL and α -NAGAL weaker than DGJ with K_i values 686nM and 2.4 μ M respectively (Fig 12A and 12B). These compounds were generously provided by the Bleriot lab [97]. These values, summarized in table 3, represent an 18-fold increase in K_i against α -GAL and a 2-fold increase against α -NAGAL compared to DGJ. When comparing the binding of α -HGJNAc to α -NAGAL to that of DGJNAc, we observed a 9-fold increase in K_i (440nM compared to 50nM) (Fig 12C). Taken as a whole, the presence of α -linked groups on either scaffold decreases the ligand's affinity to their target enzymes.

Specificity of α -HGJ

Specificity is a crucial criterion for the evaluation of a pharmacological chaperone. DGJ, despite possessing high affinity towards α -GAL, has been shown to be deficient in this regard[94]. We sought to assay the specificity of α -HGJ, and we tested this by comparing activity levels of either α -GAL or GLB1 in the presence or absence of 20 μ M α -HGJ. What we observed was that GLB1 retained 98% of its activity when incubate with high concentrations of α -HGJ. This represents a notable increase in specificity over DGJ because under similar conditions, DGJ reduced GLB1 levels to ~70% of those observed in the absence of DGJ.

Atomic basis for iminosugar binding

The structures of α -GAL + α -GJ and α -GAL + α -HGJ complexes were solved to further investigate binding of the ligands to α -GAL (Table 4). Analysis of both structures showed little difference in binding compared to that of DGJ (Fig 13 and Fig 14). The α -GAL + α -GJ structure shows well-defined density for α -GJ (Fig 13b). Unsurprisingly, the structure revealed that α -GJ utilizes the same binding mode as DGJ as shown in the superposition of the α -GAL + α -GJ with α -GAL + DGJ (Fig 13a). One clear difference between the two structures is an additional interaction present in the α -GAL + α -GJ structure. This interaction is between the axial hydroxyl group from position 1 of the sugar ring and the carboxylate of the active site D231. This hydrogen bond has a measured length of 2.8 Å and corresponds to a bond of moderate strength. Since DGJ does not possess a hydroxyl at this position, this observation is unsurprising. An

additional difference between the binding of α -HGJ and α -GJ to DGJ is an increased bond length between the carboxylate of D170 and the heterocyclic nitrogen of the ligand in the respective structures, 2.7 Å in the DGJ structure and 2.8 Å in the α -GJ structure. Slight rotations of α -GJ results in a minor shift of the ligand heterocyclic nitrogen relative to the DGJ structure. It appears this rotation is necessary to accommodate the 1-hydroxyl present in α -GJ.

The structure of the complex of α -HGJ bound to α -GAL shows clear density for α -HGJ (Fig 14b). The structure of α -GAL + α -HGJ contains an interaction not found in that of the DGJ structure. This hydrogen bond is between the hydroxyl of the hydroxymethyl at position 1 of the sugar ring and the carboxylate of the active site D231. This interaction has a measured distance of 3.5 Å, corresponding to a weak hydrogen bond. As in the case of α -GJ and DGJ structures, the α -GAL + α -HGJ has a preserved hydrogen bond between the carboxylate of D231 and the hydroxyl at the 2 position of the iminosugars with a measured distance of 2.6 Å. Unlike what was observed in the α -GAL + α -GJ structure, the ionic interaction between the carboxylate of D170 and the heterocyclic nitrogen of the sugar ring of α -HGJ is conserved with what is observed in the α -GAL + DGJ and maintains a distance of 2.7 Å across the two structures. When aligning the α -GJ and α -HGJ structures to the active site of the α -GAL + DGJ structure, we observed a more pronounced displacement of the α -HGJ relative to the two ligands (Fig 15). This greater relative displacement may explain the fact that α -HGJ has a lower affinity towards α -GAL than GJ and DGJ.

Discussion

Since it was first reported that infusions of galactose could be used to ameliorate disease symptoms in a patient suffering from the cardiac variant of Fabry disease, the application of pharmacological chaperones towards the treatment of Fabry disease was launched [58]. With the success of DGJ as a pharmacological chaperone as a treatment for Fabry disease, the application of this strategy towards other lysosomal storage disorders has become an expanded focus of research. The search for superior PCT molecules for Fabry disease treatment options has also led to new avenues of research. One such avenue is the development of more membrane permeable drugs. Initial attempts to increase permeability involved the design of DGJ derivatives with increased relative octanol-water partition coefficients through N-alkylation [43]. These initial studies resulted in a less efficacious molecule, although the field remains open. Other avenues include the design of drugs that extend the structure of DGJ to bind additional portions of the active site. These studies found that when treated with 1-deoxygalactonojirimycin arylthioureas, Fabry fibroblast cells possessed considerably higher α -GAL activity than those treated with DGJ [98]. Another area of research has sought to produce compounds with a more pH-sensitive mechanism of action as described by Mena-Barragan et al [99]. In this study, we sought to increase the specificity of the DGJ scaffold towards the active site of α -selecting enzymes.

In order for a competitive inhibitor to be a high-quality for PCT, it needs to meet a number of criteria: it must possess reasonable affinity towards its target ($K_i < 10\mu\text{M}$), bind reversibly, permeate the ER lumen, and possess high specificity (adapted from Fan et al. [100]). In the cases of DGJ and DGJNac, the first two criteria are met through the

interaction between the heterocyclic nitrogen of the compound and the active site carboxylate of D170. At neutral pH, the carboxylate of D170 in α -GAL or D156 in α -NAGAL is deprotonated and forms an ionic interaction with the protonated heterocyclic nitrogen of DGJ or DGJNAc. Upon entering the acidic environment of the lysosome, the carboxylate becomes protonated, and the ionic interaction is lost, resulting in a significant decrease in affinity. Through this mechanism, the compounds are able to stabilize their targets in the neutral pH of the endoplasmic reticulum and release upon entering the acidic pH of the lysosome, where they are able to act upon substrate. Despite the pH-dependent affinity of DGJ and DGJNAc, both compounds lack specificity, because they bind enzymes that cleave both α - and β - glycosidic linkages. Both compounds lack the 1-hydroxyl of galactose or N-acetylgalactosamine and as a result do not possess specificity towards α - and β - selecting hydrolases.

α -HGJNAc and α -HGJ retain the pH-dependent affinity due to the heterocyclic nitrogen of the sugar ring. However, α -HGJNAc and α -HGJ contain a stable α -linked hydroxymethyl group attached to C-1 of the sugar rings. The presence of this group provides specificity towards α -selecting galactosides, as shown by the ability of α -HGJ to inhibit α -GAL while showing little inhibitory effect on human β -GAL. This finding suggests that the addition of an α -linked group to C-1 of either a α -Gal or α -GalNAc-derived iminosugar is sufficient to gain specificity towards α -GAL or α -NAGAL. However, the α -linked group produces a small penalty on binding affinity. The structural analysis of α -GAL in complex with α -GJ and α -HGJ reveals that these ligands adopt a position that is slightly rotated relative to the positioning of DGJ (however, the resolution of the solved structures limits the precision of the coordinates). The rotation of the ligand

is necessitated by steric constraints between the ligand functional group and the side-chain of D231. The degree of this rotation is dictated by the nature of the functional group. The structure of α -GAL in complex with α -HGJ reveals that the hydroxymethyl group of α -HGJ is positioned in close proximity to the carboxylate of D231. In the case of α -GJ a hydrogen bond exists between the carboxylate of D231 and the hydroxyl group of α -GJ. The enthalpy of the bond and the less restrictive steric restraints result in only a minor decrease in affinity.

In summary, we have shown that the addition of axially positioned groups to the C-1 position of a galactonojirimycin scaffold results in increased specificity towards α -selecting galactosidases. Compounds containing these groups utilize the same binding mode as the 1-deoxy forms of the compounds and have similar affinities towards their intended targets. These findings taken as a whole add to our understanding of rational design of molecules intended for PCT. Despite the slightly reduced binding affinity of molecules with adducts on the 1 position, the gained specificity decreases the degree to which off-target binding occurs and may decrease the likelihood of unintended effects during treatment.

Experimental Procedures

Crystallization and x-ray data collection

Crystals of α -GAL were grown using the hanging drop method with a 1:1 mixture of reservoir solution (8% PEG 8000, 0.1M sodium cacodylate pH 6.5, and 20 mM $\text{Mg}((\text{CH}_3\text{COO})_3)_2$ and 5 mg/mL α -GAL in 20mM Bis-Tris pH 6.0. Crystals were transferred into stabilization buffer (10% PEG 8000, .1M sodium cacodylate pH 6.5, and

20 mM Mg((CH₃COO)₃)₂ and then into a cryoprotectant solution (14% PEG 8000, .1M sodium cacodylate pH 6.5, 20 mM Mg((CH₃COO)₃)₂, and 20% glycerol) containing either 10mM α -HGJ or 20mM α -GJ. The crystals were then flash-frozen in liquid nitrogen. X-ray data were collected at the X25 beamline at Brookhaven National Laboratory. X-ray images were processed using iMosflm and phase determination was carried out using Fourier synthesis using α -GAL coordinates (PDB: 3HG3). Modeling and refinement were performed using Coot and Refmac5. Figures were prepared using Pymol.

Protein expression and purification

Expression of α -GAL and α -NAGAL was accomplished as previously described. In short, *Trichoplusia ni* (Tn5) cells expressing either α -GAL or α -NAGAL were grown to $\sim 3 \times 10^6$ cells/mL at which point the supernatant was separated through centrifugation. The supernatant was then buffer exchanged into nickel binding buffer (50mM phosphate buffer pH 7.0, 250mM NaCl, and 40mM Imidazole) using a tangential flow cartridge. The exchanged supernatant was then applied to a Nickel-FT column and eluted using a linear imidazole gradient. The fractions containing either α -GAL or α -NAGAL were pooled, concentrated, and buffer exchanged into ion exchange binding buffer (50mM Bis-tris pH 6.0) using an Amicon spin column. The fractions were then applied to a UNO-Q ion exchange column and eluted with a linear gradient of NaCl. The fractions containing either enzyme were then concentrated to 1mg/mL.

Enzyme inhibition assays

0.4 μ M aliquots of α -GAL in assay buffer (100mM citrate/phosphate buffer pH 4.5) were mixed with varying concentrations of α -GJ, α -HGJNAc, or α -HGJ, and incubated at 4°C for 30 minutes. 10 μ L of the enzyme/inhibitor solution was then added to 20 μ L of 15mM *para*-nitrophenyl- α -galactose (pNP- α -Gal) or 5mM *para*-nitrophenyl- α -N-acetylgalactosamine (pNP- α -GalNAc) and the reaction was incubated for 10 minutes at 37°C at which point 10 μ L of the reaction was added to 290 μ L of 200mM sodium borate pH 9.8 in 96-well plates. Abs₄₀₀ readings of the samples were measured and product concentrations were calculated using an extinction coefficient of 18.1 mM⁻¹ cm⁻¹. The averages of triplicate data sets were fitted to semilog sigmoid dose response curves using Kaleidagraph. IC₅₀ values were calculated and K_i was determined using the following:

$$\text{normalized activity} = \frac{A_{\max} - A_{\min}}{1 + 10^{b(\log[I] - \log IC_{50})}}$$

$$K_i = \frac{IC_{50}}{\left(1 + \frac{[S]}{K_M}\right)}$$

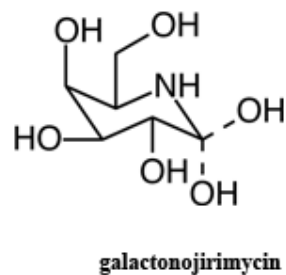
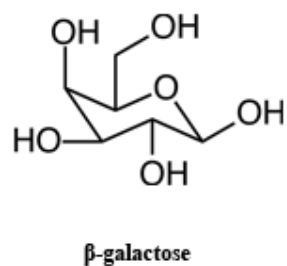
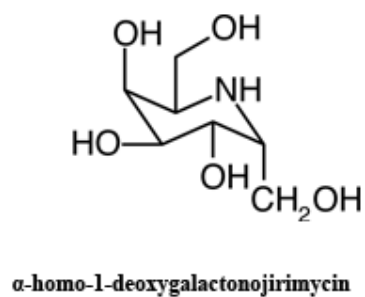
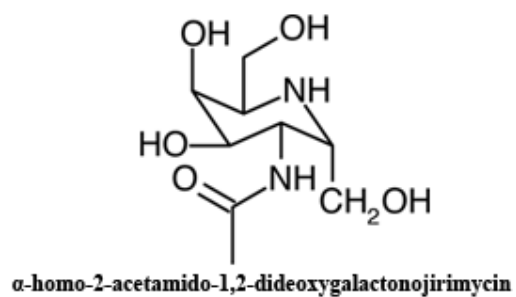
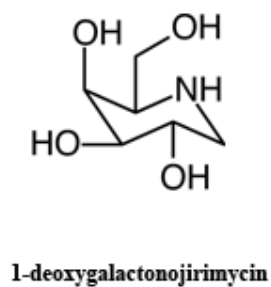
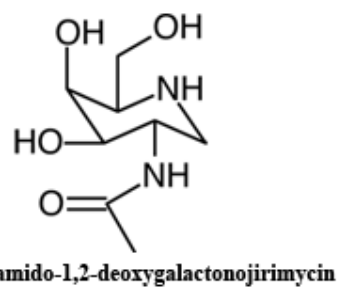
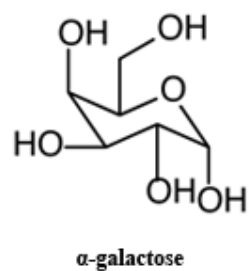
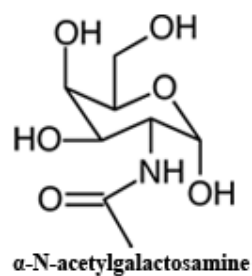


FIGURE 10. Structures of α -galactose, α -N-acetylgalactosamine, and imino sugar analogs

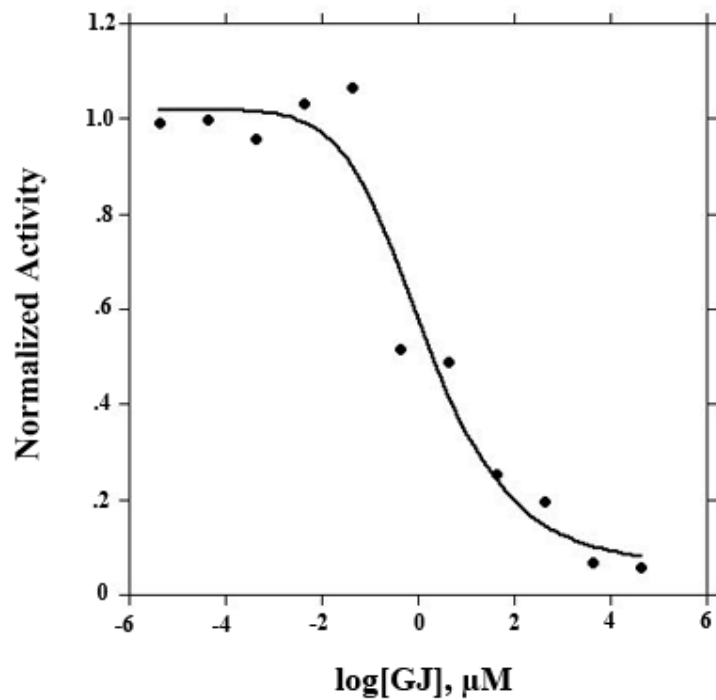


FIGURE 11. Enzymatic Inhibition of α -galactosidase by galactonojirimycin (GJ)
The ability of GJ to inhibit the cleavage of pNP- α -gal by human α -galactosidase at pH of 4.5 was measured and fit with a semi-log sigmoid dose response curve.

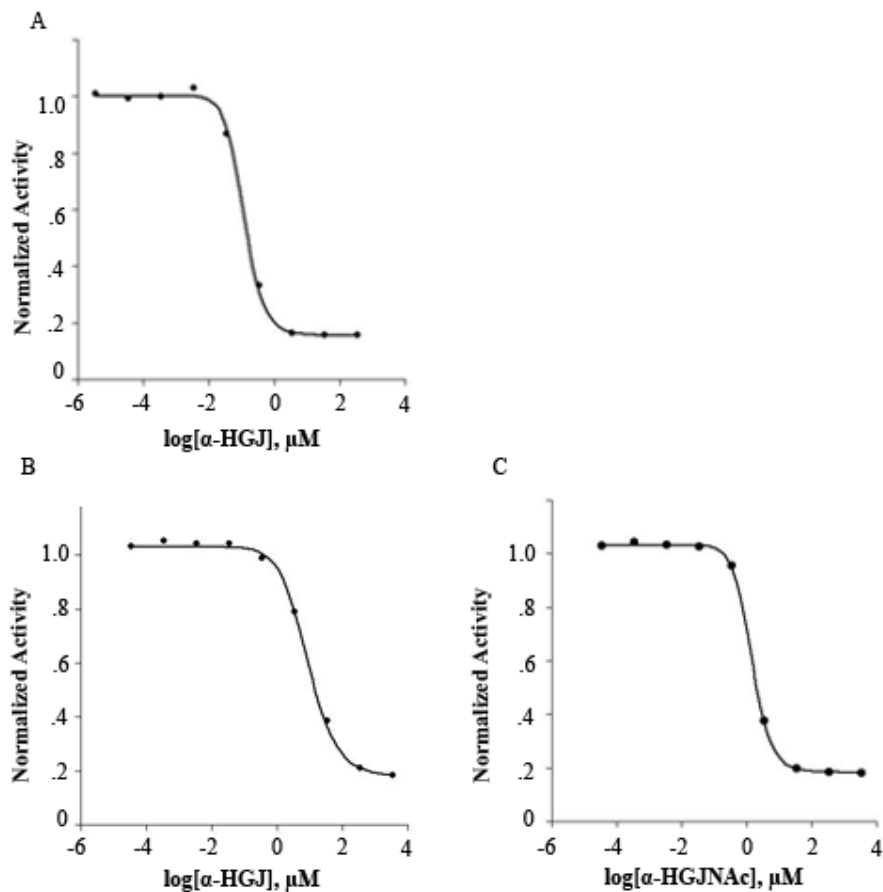


FIGURE 12. Enzymatic Inhibition of α -galactosidase and α -N-acetylgalactosaminidase by α -HGJ and α -HGJNAc

The ability of α -HGJ and α -HGJNAc to inhibit the cleavage of pNP substrate by either α -galactosidase or α -N-acetylgalactosaminidase. (A) Depicts the inhibitory effect of α -HGJ on the ability of α -galactosidase. pNP- α -gal. (B and C respectively) Depict the ability of α -HGJ and α -HGJNAc to α -N-acetylgalactosaminidase-catalyzed cleavage of pNP- α -galNAc

| | α -GAL | α -NAGAL |
|-----------------------------------|---------------|-----------------|
| Galactose | 16 mM | >16 mM |
| GalNAc | N/A | 13 mM |
| GJ | 84 nM | Not determined |
| DGJ | 39 nM ** | 1.5 μ M * |
| DGJNAc | N/A * | 51 nM * |
| α-HGJ | 686 nM | 2.4 μ M |
| α-HGJNAc | N/A | 440 nM |

TABLE 3. Summary of K_i Values

Summarization of known inhibitor constants of various compounds against α -GAL or α -NAGAL.

*data previously published by Clark et al [56]

**data previously published by Guce et al [14]

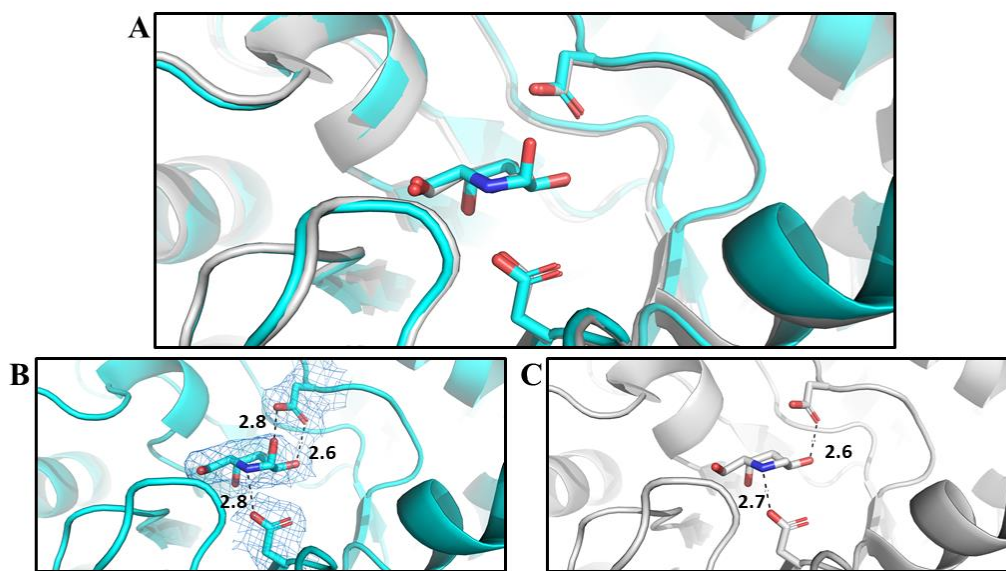


FIGURE 13. **Structure of α -GAL bound to α -GJ.** (A) Superposition of the active sites of α -GAL bound to DGJ (Grey) (pdb code 3S5Y) and α -GAL bound to α -GJ (Cyan). (B) Active site of α -GAL bound to α -GJ with density corresponding to α -GJ and active site residues D231 and D170 shown. Distance measurements are shown between critical interactions. (C) Active site of α -GAL bound to DGJ with distances shown for critical interactions

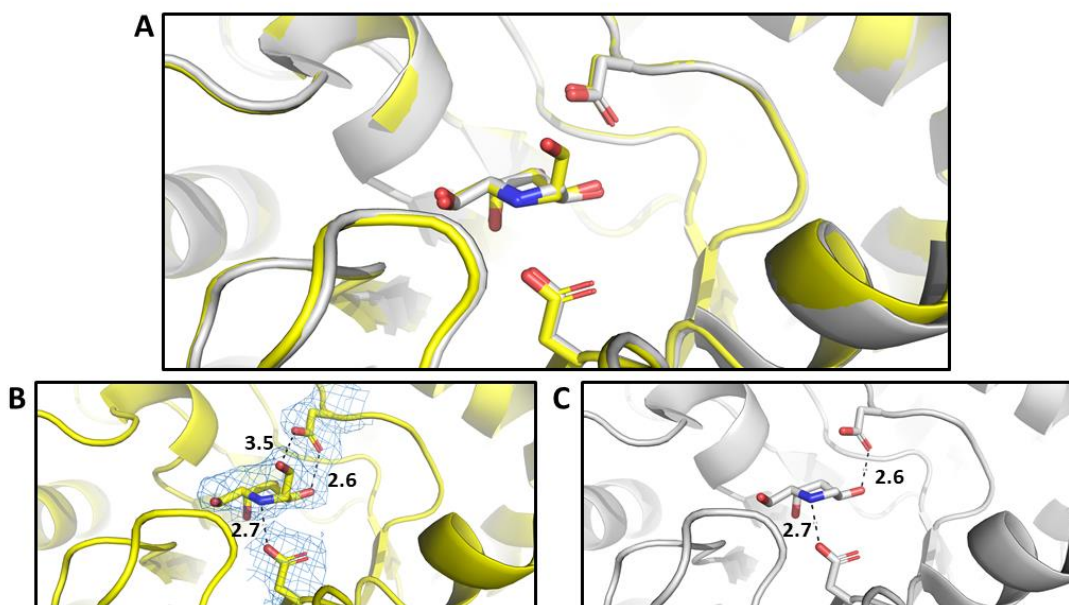


FIGURE 14. Structure of α -GAL bound to α -HGJ (A) Superposition of the active sites of α -GAL bound to DGJ (Grey) (PDB code 3S5Y) and α -GAL bound to α -HGJ (Yellow). (B) Active site of α -GAL α -HGJ with density corresponding to α -HGJ and active site residues D231 and D170 shown. Distance measurements are shown between critical interactions. (C) Active site of α -GAL bound to DGJ with distances shown for critical interactions

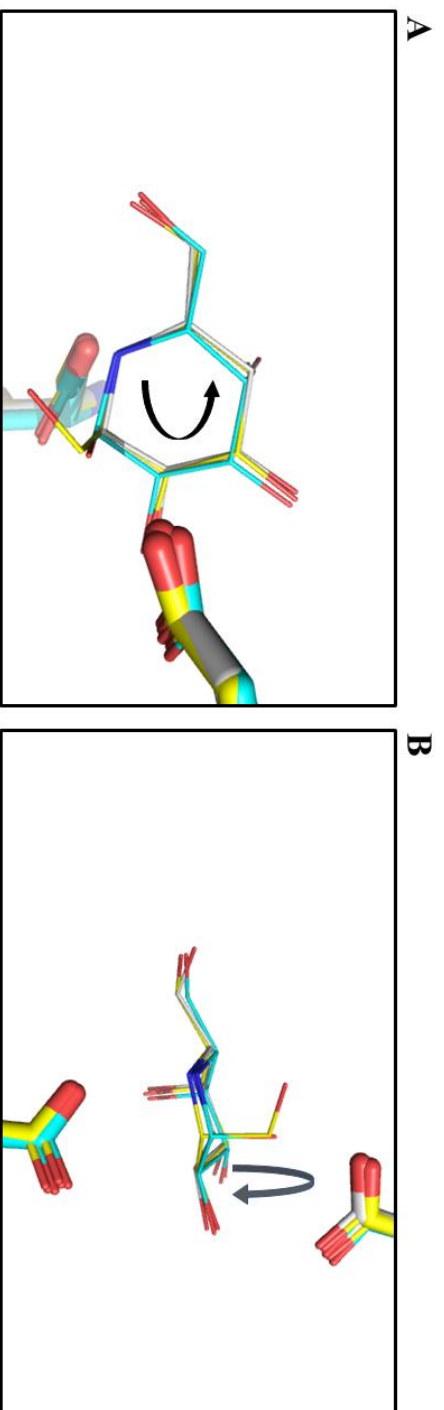


FIGURE 15. Comparison of iminosugar binding to α -GAL (A) Pair-fit alignment of α -HGJ and α -GJ structures to the active site of α -GAL + DGJ. Catalytic residues aspartic acids 170 and 231 are shown in all structures. α -GAL + DGJ is shown in grey, α -GAL + α -GJ shown in cyan, and α -GAL + α -HGJ (B) 90° rotated view.

| Data | | |
|-------------------------------------------|-------------------------|-------------------------|
| Ligand Soak | GJ | α -HGJ |
| X-ray source | BNL X25 | BNL X25 |
| Wavelength, Å | 1.100 | 1.100 |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ |
| Cell a, b, c in Å | 60.1, 107.9, 183.2 | 59.6, 106.3, 182.4 |
| Cell α , β , γ in ° | 90, 90, 90 | 90, 90, 90 |
| Resolution, Å (last shell) | 69.83 – 2.1 (2.4 – 2.1) | 53.71 – 2.3 (2.4 – 2.3) |
| No of observation (last shell) | 658,185 (91104) | 492,708 (70063) |
| Unique observation (last shell) | 70,812 (10039) | 52,439 (7543) |
| Completeness, % (last shell) | 99.7 (98.1) | 100 (100) |
| Multiplicity (last shell) | 9.3 (9.1) | 9.4 (9.3) |
| R_{sym} (last shell) | 0.197 (0.882) | 0.082 (0.455) |
| $\langle I/\sigma_1 \rangle$ (last shell) | 6.5 (2.0) | 15.2 (4.2) |
| Refinement | | |
| Rwork / Rfree | 0.206 (0.246) | 0.197 (0.240) |
| Ramachandran: Favored | 96.16 | 95.8 |
| Allowed | 3.32 | 3.82 |
| Generous | 0.26 | 0 |
| Forbidden | 0.26 | 0.38 |
| RMS deviations: Bonds, Å | .013 | 0.012 |
| Angles, ° | 1.63 | 1.58 |

TABLE 4. Structural Data Collection and Refinement Statistics

Chapter 5: Conclusions and Future Directions

Lysosomal storage diseases are a class of rare metabolic disease resulting from insufficient levels of enzymatic activity in the lysosome, resulting in the accumulation of substrate over time that ultimately leads to the development of disease symptoms. Normally, the degradative function of the lysosome is achieved through the sequential action of numerous lysosomal hydrolases that act to catabolize large macromolecules into smaller components that can then be recycled or expelled from the cell. An individual lysosomal storage results from the loss of one specific enzyme function, and results in the normal sequential process of lysosomal degradation halting, and the accumulation of the substrate specific to the missing enzyme. Due to the monogenic nature of lysosomal storage diseases, theoretical treatment options for these diseases are fairly simple when compared to diseases with more complex disease mechanisms. In fact enzyme replacement therapy, the most widely utilized treatment for lysosomal storage diseases, was suggested alongside the first study identifying the mechanism for a lysosomal storage diseases; *“In our pathogenic speculations and in our therapeutic attempts, it may be well to keep in mind that any substance which is taken up intracellularly in an endocytic process is likely to end up within lysosomes. This obviously opens up many possibilities for interaction, including replacement therapy”* [101, 102].

Despite the conceptual simplicity of enzyme replacement therapy, practical limitations have led to the need for developing alternate treatment strategies as well as modifications to ERT drugs. There are numerous universal challenges facing ERT for lysosomal storage diseases. These include poor cellular uptake of the ERT enzyme into afflicted cell types. This limitation led to the development of glycan-engineered enzymes

for the treatment of Gaucher's and Pompe disease, where these enzymes possess an increased ability to bind extracellular mannose 6-phosphate receptors and are more efficiently taken into cells. Low serum half-life is also a universal limitation of ERT as these enzymes have evolved to act in the acidic environment of the lysosome and are less stable in the near neutral pH of patient sera. Patient immune response is also a challenge shared across ERT. An inability to cross blood-brain barrier also limits ERT, under normal drug administration protocols, and is of particular concern for lysosomal storage disorders that have neurological implications. In addition to the numerous shared limitations of ERT, there exist specific limitations for each disease.

In Fabry disease, a major limitation of faced by ERT is incomplete clearance of Gb3 inclusions from podocyte cells. It has been shown that despite receiving ERT podocyte cells taken from Fabry patients still contain Gb3 inclusions [15]. Additionally, it has been observed that patients develop podocyte foot effacement during ERT [16]. Podocyte cells form the last layer of the glomerular filtration barrier and is naturally a difficult target for a circulating enzyme. Based on the observation that permeability across the glomerular filtration barrier is inversely related to molecular radius, we sought to design an engineered form of α -GAL with a decreased molecular radius. The identification of two positions differentially conserved between monomeric and dimeric forms of glycosyl hydrolase family 27 members allowed us to design a mutant form of human α -GAL that exists in a monomeric state. Our engineered α -GAL was shown to have reduced mobility on a size-exclusion chromatography column and increased permeability across size selective membranes. The engineered enzyme retained a large

portion of wild-type enzymatic activity. Our variant α -GAL showed signs of reduced stability when compared to the wild-type form.

The results of this study show that human α -GAL can be converted to a monomeric form and presents a novel strategy to increase efficacy of ERT of podocyte Gb3 inclusion clearance. Although; the current monomeric variant shows decreased stability and activity, if these shortcomings were to be addressed through further protein modification it would represent a promising enzyme for pre-clinical trials. Specifically, this enzyme would be of particular importance for patients suffering from renal complications (classical forms and the renal variants of the late-onset form). It could also potentially be used in combination with the current ERT drugs, which have higher stability and readily clear Gb3 inclusions from the bulk of tissue. This combination of a more generic enzyme and one specialized for podocyte cells would retain the general efficacy of the current drugs but would gain efficacy towards treating podocyte cells.

Another major issue associated with ERT for Fabry disease is immunogenicity. Eighty-eight percent of patients with the classical form of the disease develop IgG- or IgE- based reactions to the treatment. Working under the hypothesis that this immune response results from these patients being immune naïve to correctly folded α -GAL, we engineered a mutant form of α -NAGAL possessing the active site of α -GAL. This engineered form of α -NAGAL presents itself immunologically as α -NAGAL, which we confirmed testing antigenicity against polyclonal sera raised against α -GAL and α -NAGAL. We also showed that this active conversion did indeed result in a change in enzyme specificity as the engineered enzyme showed an increase in specificity towards α -galactoside substrate and was no longer capable of catalyzing the hydrolysis of

substrate with terminal N-acetylgalactosamine sugars. Additionally, we solved the structure of the inverse conversion, and determined that the binding mode was preserved when transplanting active sites.

Our engineered α -NAGAL strategy is not the only strategy being developed towards addressing the issue of immunogenicity. Pegunigalsidase alfa is a PEGylated form of α -GAL currently in clinical trials. Initial results indicate that this form of α -GAL is less immunogenic and has a higher serum half-life than current ERT drugs [57]. The increased serum-half life and lower immunogenicity are promising, but the same principle limiting immune response, in which the addition of PEG polymers limits epitope presentation, would also likely lower receptor binding. Additionally, this drug requires in vitro modifications further complicating drug production.

Our engineered α -NAGAL has the benefit of being immunologically neutral while being nearly as active as the wild-type α -GAL. It is an ideal design for gene therapy and has been used in proof-of-concept studies for gene therapy in Fabry models by UniQure. These initial studies showed proof of concept and will hopefully warrant progression towards clinical trials. This advancement is tremendously exciting as lysosomal storage diseases are prime candidates for gene therapy, due their monogenetic nature, and the use of this design is particularly appealing, because any immune response generated would be of dire consequence as gene therapy is non-reversible.

For Fabry disease, pharmacological chaperone therapy is the only currently approved alternate treatment to ERT and has only recently received approval. DGJ is the drug used in this therapy and possesses characteristics ideal for PCT. High affinity and pH dependence are critical to the success of DGJ as a PCT drug. One shortcoming that

was identified is a lack of specificity towards α - and β -selecting galactosidases/N-acetylgalactosaminidases [94]. We sought to address this through the use of iminosugars containing α -linked groups attached to the C1 position of the sugar ring. We found that these compounds selectively bound α -selecting enzymes with affinities slightly lower than that of DGJ or DGJNAc. Structural analysis revealed that these compounds utilize the same ionic interaction, which in DGJ provides the high pH-dependent affinity.

Specificity is of high importance for PCT, because off-target binding has a number of adverse effects. When off-target binding is present, the intended target competes with other proteins for binding to the chaperone. This competition may result in higher dosage of chaperone molecule. In addition to affecting dosages, off-target binding results in the inhibition of unintended proteins, which poses a risk for adverse clinical effects. The compounds we highlighted in our study provide an alternative to DGJ that utilize the same pH-dependent mechanism but are restricted to α -selecting enzymes. This is an important advancement, because DGJ has been shown to bind unintended targets, such as β -galactosidase, with relatively high affinity [94].

At the current time, the arrival of second-generation Fabry disease treatments are on the horizon. There are numerous exciting prospects, such as PEGylated ERT therapeutics, plant-expressed ERT therapeutics, substrate-reduction therapeutics, and gene-therapy treatments. Although there is likely to be an increase in treatment options available to Fabry patients, the opportunity for combination therapies has yet to be exploited and may provide the most efficacious option for Fabry disease. This combinatorial approach has already been shown to be efficacious through the combination of current generation of ERT drugs and DGJ in Fabry mouse models [93]. This

combination showed an increased in recombinant α -GAL serum half-life, increased cellular α -GAL levels, and greater Gb3 clearance when compared to ERT alone. The potential for additional combinations would allow for specialized treatments for different variations of the disease. For example, a combination of standard ERT therapeutic, PCT therapeutic, and ERT therapeutic specialized towards renal efficacy would be ideal for patients with more pronounced renal symptoms, but would not be ideal for a patient with the cardiac variant of the disease.

Summary

In the studies presented we have described strategies to develop second generation therapeutics for the treatment of Fabry disease. These strategies address shortcomings of the current treatment options for Fabry disease. These studies targeted enzyme replacement therapy and pharmacological chaperone therapy.

One of the shortcomings of enzyme replacement therapy addressed was immunogenicity. To address the issue of immunogenicity we interconverted the active sites of α -GAL and α -NAGAL. We demonstrated that these mutated enzymes have effectively had their specificities swapped. These interconversions have no effect on the antigenicity of the enzyme. Structural and enzymatic studies show that these converted active sites behave comparably to their counterpart wild-type enzymes. These enzymes, α -NAGAL^{EL} in particular, are attractive designs for future clinical studies. α -NAGAL^{EL} possesses the specificity of α -GAL but presents to immune system as α -NAGAL.

The second shortcoming addressed is poor podocyte substrate clearance. We engineered a monomeric form of α -GAL, α -GAL^{F273G+W277G}, and demonstrated that this

enzyme retains a significant portion of wild-type enzymatic activity. More crucially, this enzyme was confirmed to more readily permeate size selective membranes. Although, this enzyme showed signs of reduced stability, it shows promise as a scaffold for other engineered monomeric α -GAL proteins. The goal would be to produce an ERT therapeutic better able to transverse the glomerular filtration barrier and to clear substrate from podocyte cells.

Lastly, we identified a strategy to design small molecules that have increased specificity towards α -selecting galactosidases and N-acetylgalactosaminidases as compared to current pharmacological chaperones for Fabry disease. We demonstrated that these compounds were excluded from β -galactosidases but were still capable of binding α -GAL and/or α -NAGAL with high affinity. Structural studies showed that these compounds bound α -GAL with the same binding mode as DGJ, the current FDA-approved Fabry pharmacological chaperone.

Appendix

PHARMACOLOGICAL CHAPERONES FOR HUMAN α -N- ACETYLGALACTOSAMINIDASE

I made contributions to generating the stable cell lines used for protein expression. I also contributed to the purification of the proteins used in this study and performed part of the specificity assays used in this study.

This article is the first to explicitly identify potential pharmacological chaperones for Schindler/Kanzaki disease. It also lays out a template for designing compounds specific for α -N-acetylgalactosaminidases, and was the first to identify that DGJ has a chaperoning effect on α -N-acetylgalactosaminidase.

Pharmacological chaperones for human α -N-acetylgalactosaminidase

Nathaniel E. Clark^a, Matthew C. Metcalf^b, Daniel Best^b, George W. J. Fleet^{b,c}, and Scott C. Garman^{a,1}

^aDepartment of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, MA 01003; ^bChemistry Research Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3TA, United Kingdom; and ^cOxford Glycobiology Institute, University of Oxford, Oxford OX1 3QU, United Kingdom

Edited* by Gregory A. Petsko, Brandeis University, Waltham, MA, and approved September 11, 2012 (received for review March 6, 2012)

Schindler/Kanzaki disease is an inherited metabolic disease with no current treatment options. This neurologic disease results from a defect in the lysosomal α -N-acetylgalactosaminidase (α -NAGAL) enzyme. In this report, we show evidence that the iminosugar DGJNac can inhibit, stabilize, and chaperone human α -NAGAL both in vitro and in vivo. We demonstrate that a related iminosugar DGJ (currently in phase III clinical trials for another metabolic disorder, Fabry disease) can also chaperone human α -NAGAL in Schindler/Kanzaki disease. The 1.4- and 1.5-Å crystal structures of human α -NAGAL complexes reveal the different binding modes of iminosugars compared with glycosides. We show how differences in two functional groups result in >9 kcal/mol of additional binding energy and explain the molecular interactions responsible for the unexpectedly high affinity of the pharmacological chaperones. These results open two avenues for treatment of Schindler/Kanzaki disease and elucidate the atomic basis for pharmacological chaperoning in the entire family of lysosomal storage diseases.

glycosidase | protein folding | structural biology | X-ray crystallography | E.C. 3.2.1.49

The lysosome is an acidic organelle containing dozens of enzymes that catalyze the breakdown of cellular metabolites. One lysosomal enzyme is α -N-acetylgalactosaminidase (α -NAGAL) (E.C. 3.2.1.49), a hydrolase that catalyzes removal of terminal α -linked N-acetylgalactosamine (α -GalNAc) and (less efficiently) galactose monosaccharides (1). Substrates for human α -NAGAL include glycans found in O-linked glycosylation, blood group A antigens, and others with α -GalNAc glycosides. In humans, mutations in the *NAGA* gene that codes for α -NAGAL lead to a loss of enzymatic activity in the lysosome and the subsequent accumulation of substrates, which eventually results in the autosomal recessive Schindler/Kanzaki disease (2–4). Schindler/Kanzaki disease belongs to the lysosomal storage disease family, which includes such diverse diseases as Tay-Sachs, Sandhoff, Gaucher, and Fabry diseases. Lysosomal storage diseases are characterized by the accumulation of substrates in the absence of a functional enzyme, leading to progressive deterioration in patients. Schindler/Kanzaki disease presents a wide range of symptoms in patients but is primarily characterized by neuronal pathologies (5–11). In humans, the *NAGA* gene is most closely related to the *GLA* gene, which codes for α -galactosidase (α -GAL A), and mutations in *GLA* lead to Fabry disease (12–14).

There are three principal treatment strategies currently approved or in clinical trials for lysosomal storage diseases. The first is enzyme replacement therapy, where the missing enzymatic activity is provided by regular injections of enzyme purified from recombinant sources. This therapy has been approved for the treatment of Fabry, Gaucher, and Pompe diseases and mucopolysaccharidosis I, II, and VI (15). A second strategy is substrate reduction therapy, where an inhibitor of an enzyme upstream in a biosynthetic pathway leads to reduced substrate accumulation (16). A third approach is pharmacological chaperone therapy, where the mutant enzyme is stabilized by the addition of a small-molecule chaperone. This strategy has been proposed for Gaucher and

Pompe diseases and is currently in phase III clinical trials for Fabry disease (17, 18). Approximately 50% of Fabry disease mutations lead to defects in the folding or stability of the enzyme, and this subset responds to pharmacological chaperone in cellular studies (18). Additionally, the pharmacological chaperone strategy allows the possibility of treatment of lysosomal storage diseases with neurological manifestations, because small-molecule chaperones can potentially cross the blood–brain barrier, whereas macromolecular enzymes are typically unable to cross into the brain. To prevent the occurrence of lysosomal storage diseases, a modest 5–15% threshold of enzymatic activity may be sufficient (17, 19–21).

The most promising pharmacological chaperone in the clinic is the iminosugar analog of galactose, 1-deoxygalactonojirimycin (DGJ), used for the treatment of Fabry disease (22). DGJ can bind and stabilize α -GAL A at neutral pH in the endoplasmic reticulum (ER), allowing it to traffic to the lysosome, where it then dissociates at low pH (18, 22–28). Paradoxically, the addition of a small-molecule competitive inhibitor of an enzyme leads to an increase in the amount of activity of the enzyme, which slows the progression of the disease.

In this report, we sought specific pharmacological chaperones for the human α -NAGAL enzyme. We found that DGJ can bind, inhibit, and chaperone α -NAGAL nearly as well as it chaperones α -GAL A. Because DGJ has promiscuity for a number of lysosomal enzymes, we tested a related molecule with improved specificity for α -NAGAL. We synthesized an iminosugar, 2-acetamido-1,2-dideoxy-D-galactonojirimycin (DGJNac), specific for α -NAGAL (29, 30), and we examined its ability to bind, inhibit, and chaperone human α -NAGAL. We performed enzyme kinetics experiments to measure the binding and inhibitory properties of the two compounds on human α -NAGAL. We show that DGJNac and DGJ are able to protect human α -NAGAL from proteolytic degradation. We determined high-resolution crystal structures of the complexes of DGJNac and DGJ bound to human α -NAGAL at 1.4 and 1.5 Å, respectively, and a glucose-soaked structure with an empty active site at 1.6 Å. We then tested the ability of DGJNac and DGJ to chaperone α -NAGAL in cellular assays. Overall, our experiments reveal the atomic basis for the tight binding of iminosugars to lysosomal glycosidases. We show that both DGJNac and DGJ are suitable pharmacological chaperones for α -NAGAL, and the potency of DGJNac in vitro is ~10-fold greater. We propose DGJNac and DGJ as potential pharmacological chaperones for Schindler/Kanzaki disease patients.

Author contributions: N.E.C. and S.C.G. designed research; N.E.C., M.C.M., and S.C.G. performed research; N.E.C., D.B., G.W.J.F., and S.C.G. contributed new reagents/analytic tools; N.E.C., M.C.M., and S.C.G. analyzed data; and N.E.C., M.C.M., D.B., G.W.J.F., and S.C.G. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4DO4, 4DO5, and 4DO6).

To whom correspondence should be addressed. Email: garman@biochem.umass.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1209924109/-DCSupplemental.

Results

Inhibition of α -NAGAL by the Iminosugars DGJNac and DGJ. To test whether DGJNac and DGJ would make suitable pharmacological chaperones for human α -NAGAL, we first measured their binding in enzymatic inhibition assays. As predicted from their similarity to the catalytic products of the α -NAGAL reaction, both DGJNac and DGJ are inhibitors of α -NAGAL. DGJNac is a tight binding inhibitor of α -NAGAL, with a K_i of 51 nM, whereas DGJ binds 29-fold weaker with a K_i of 1.5 μ M (Fig. 1). We also tested the ability of catalytic products to inhibit human α -NAGAL: GalNac binds 250,000-fold weaker with a K_i of 12.7 mM (Fig. 1), and galactose is a very weak inhibitor, where concentrations up to the assay limit of 100 mM did not inhibit the enzyme. By comparison, in the interaction exploited in clinical trials of pharmacological chaperone therapy for Fabry disease, DGJ has a K_i of 40 nM for α -GAL A. Because the affinities of DGJNac for α -NAGAL and DGJ for α -GAL A are similar, DGJNac has potency in the appropriate range for pharmacological chaperone therapy in Schindler/Kanzaki disease.

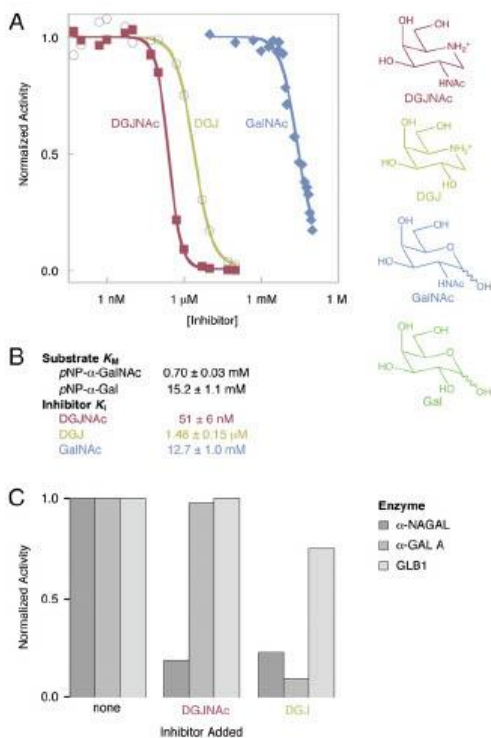


Fig. 1. Enzymatic inhibition of α -NAGAL by DGJ, DGJNac, and GalNac. (A) DGJNac (red), DGJ (yellow), and GalNac (blue) were tested for their ability to inhibit the α -NAGAL-catalyzed cleavage of pNP- α -GalNac. (B) The measured IC_{50} values were converted into K_i values as described in the text. K_M values for substrates are also shown. (C) Three enzymes with related activities (α -NAGAL, α -GAL A, and GLB1) were tested in enzymatic assays in the presence of no inhibitor (Left), 20 μ M DGJNac (Center), or 20 μ M DGJ (Right). DGJNac shows increased specificity over DGJ and inhibits only α -NAGAL.

Clark et al.

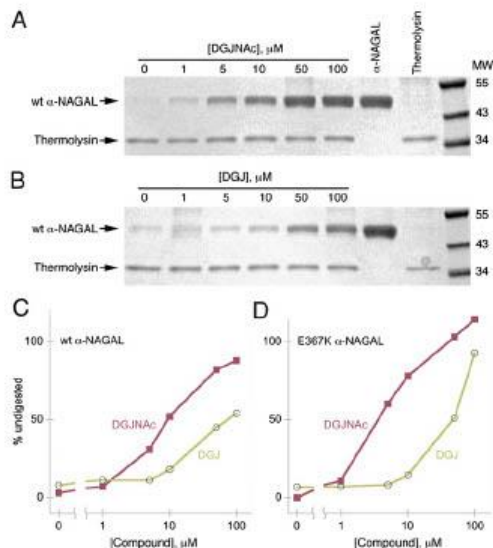


Fig. 2. DGJ and DGJNac confer protection against protease digestion of α -NAGAL. (A and B) Wild-type (wt) α -NAGAL was preincubated with 0–100 μ M DGJNac (A) or DGJ (B) before thermolysin digestion and SDS/PAGE. The intensity of the undigested α -NAGAL band indicates compound-induced protection from proteolysis. (C and D) Densitometry plots of the wild-type (C) and E367K mutant (D) α -NAGAL bands at different compound concentrations show the higher potency of DGJNac over DGJ.

To examine the specificity of DGJNac and DGJ, we performed the inhibition assays with two lysosomal enzymes with closely related activities, α -GAL A and GLB1 (also known as β -galactosidase). DGJNac is highly specific for α -NAGAL, whereas DGJ inhibits α -GAL A better than α -NAGAL and partially inhibits GLB1 (Fig. 1C), in agreement with previous enzymatic and crystallographic studies (18, 25, 27, 31–34).

Protease Protection Conferred by Iminosugar Binding. We tested whether DGJNac and DGJ could limit the proteolytic degradation of α -NAGAL when exposed to limited thermolysin digestion (35). Native α -NAGAL protein is resistant to many proteases, including trypsin, thermolysin, and pepsin, but addition of denaturant allowed for digestion. α -NAGAL was preincubated with different concentrations of DGJNac or DGJ, and then equilibrated in 4 M urea before limited digestion and SDS/PAGE, revealing the amount of undigested α -NAGAL. The digestions show that, at pH 7.2 (mimicking the pH in the ER), both DGJNac and DGJ are effective pharmacological chaperones for human α -NAGAL, where both compounds protect α -NAGAL from proteolysis, with DGJNac protecting at \sim 10-fold lower concentrations (Fig. 2). Both wild-type and the Schindler/Kanzaki mutant E367K α -NAGAL behave similarly in the digestion assay.

Atomic Basis of Iminosugar Interaction with Glycoside Hydrolase Family 27 Enzymes. To investigate the atomic basis of DGJNac and DGJ binding to α -NAGAL, we determined three crystal structures of α -NAGAL soaked with DGJNac, DGJ, or glucose (Fig. 3 and Table S1). Comparison of the new structures with the previously determined complexes of α -NAGAL with GalNac and Gal (36) reveals a highly favorable ionic interaction in the

PNAS | October 23, 2012 | vol. 109 | no. 43 | 17401

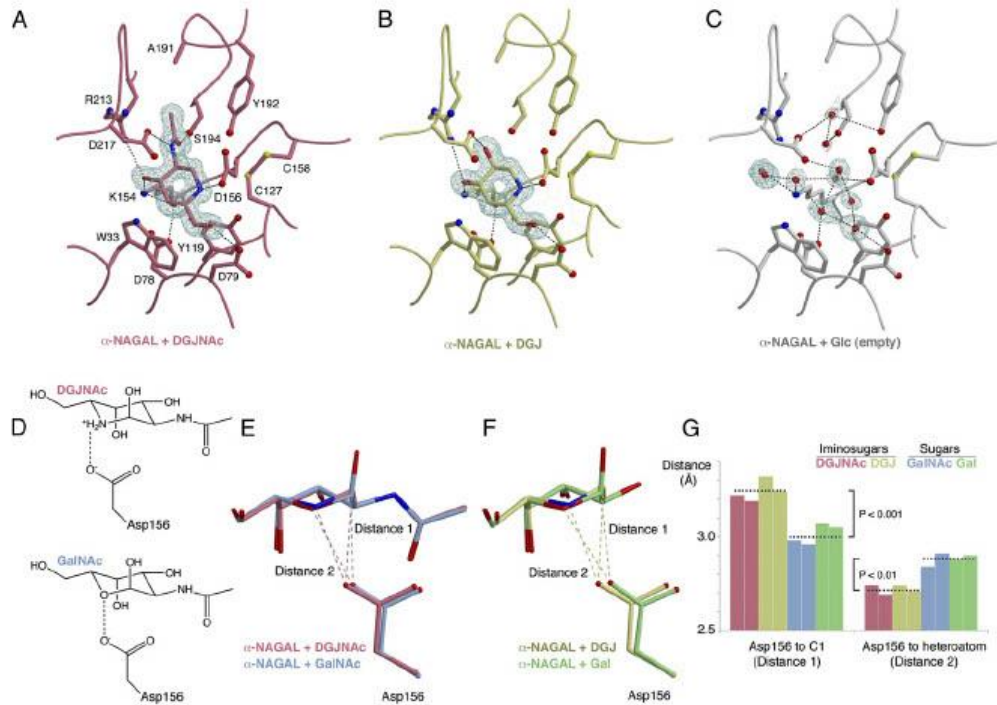


Fig. 3. Pharmacological chaperone binding in the active site of human α -NAGAL. (A–C) α -weighted 2Fo-Fc electron density maps of α -NAGAL soaked with DGJNac, DGJ, and glucose respectively. A and B are contoured at 1.8 σ around the ligand density and C at 1.5 σ . (D) Schematic indicating the key interaction between Asp156 and DGJNac or GalNac. (E and F) Superposition of structures of iminosugar complexes comparing DGJNac (red) and DGJ (yellow) with their sugar analogs GalNac (blue) and Gal (green), and the location of the catalytic nucleophile Asp156. (G) Plots of the distances of Asp156 to the ligand C1 and heteroatom, indicating the different mode of binding of the iminosugars and sugars. The colors are as in E and F, the dashed line shows the mean of four grouped measurements, and brackets show P values from Student *t* tests of the paired data.

iminosugar complexes (Fig. 3D). In sugar complexes from this family of enzymes, the catalytic nucleophile is poised to attack C1 of the ligand (31, 33, 36–40). The distances from the nucleophile D156 to C1 of GalNac (2.97 Å) and to C1 of Gal (3.06 Å) are shorter in the sugar complexes, where the *syn* lone pair of D156 points directly at C1 of the sugar, in position for nucleophilic attack. However, when an iminosugar is bound in the α -NAGAL active site, D156 shifts away from C1 (3.21 and 3.28 Å for DGJNac and DGJ, respectively) and toward N5 of the ligand (Fig. 3E–G). The DGJNac and DGJ iminosugars bind more like each other than like the GalNac and Gal sugars. When iminosugars are compared with sugars, the distances between D156 and the ligand shift by statistically significant amounts, with $P < 0.01$.

DGJNac and DGJ Chaperoning of Human α -NAGAL in Cells. To examine the ability of DGJNac and DGJ to chaperone wild-type α -NAGAL in vivo, we transfected human embryonic kidney (HEK 293T) cells with plasmids encoding FLAG-tagged human α -NAGAL in the presence or absence of the compounds. Immunoblots with anti-FLAG antibodies revealed higher levels of α -NAGAL protein expressed in cells cultured in the presence of DGJNac or DGJ (Fig. 4A and B). Despite the higher potency of DGJNac in vitro, cells cultured with DGJNac showed a lesser increase in α -NAGAL protein expression, possibly due

to differential uptake of DGJ and DGJNac in cells (30). Both DGJ and DGJNac showed statistically significant increases in protein levels compared with no compound (100 μ M DGJNac and 500 μ M DGJ have $P < 0.05$ in paired Student *t* tests). The ability of DGJ to chaperone wild-type α -GAL A in cellular assays led to its eventual development as a clinical candidate (22).

To explore the effect of the compounds on mutant forms of α -NAGAL found in Schindler/Kanzaki disease patients, we performed subcellular fractionation assays to separate lysosomal and ER-resident forms of α -NAGAL (Fig. 4C). Overexpressed α -NAGAL shows two distinct bands when blotted with a polyclonal antibody. Subcellular fractionation revealed that the higher molecular weight band (Fig. 4C, white arrowhead) fractionates with the ER marker Calnexin, retains the C-terminal FLAG tag, and has no Endo H-resistant glycans, indicative of an immature ER protein. The lower molecular weight band (Fig. 4C, black arrowhead) fractionates with the lysosomal marker LAMP2, loses the C-terminal FLAG tag, and has Endo H-resistant glycans, indicative of a mature lysosomal protein.

DGJNac and DGJ can restore wild-type cellular distribution to the S160C α -NAGAL, a Schindler/Kanzaki disease variant form. In the absence of pharmacological chaperone, cells expressing the S160C α -NAGAL mutant show little mature α -NAGAL (and almost no immature α -NAGAL, presumably due to ER-associated

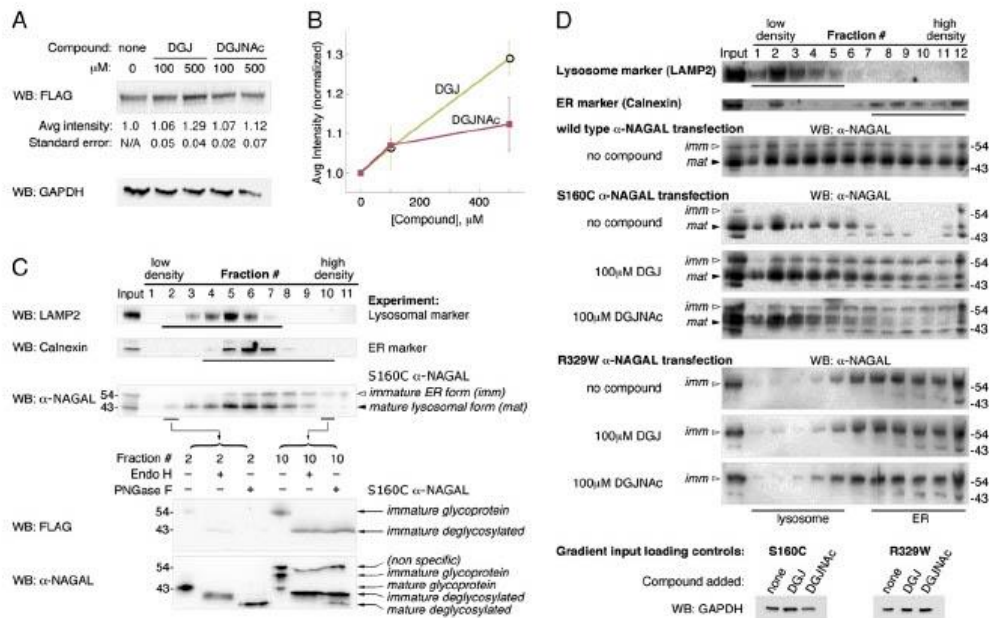


Fig. 4. Pharmacological chaperones can increase expression and lysosomal localization of human α -NAGAL in mammalian cells. (A and B) HEK 293T cells were transfected with FLAG-labeled human α -NAGAL and incubated in triplicate with 0, 100, or 500 μM DGJ or DGJNac for 24 h. Anti-FLAG Western blots (WB) of lysates show increases in the amount of α -NAGAL expressed in the presence of DGJ and DGJNac. α -NAGAL band intensities were normalized to expression of α -NAGAL in the absence of compound. Student *t* tests of paired data show $P < 0.05$ for 100 μM DGJNac and 500 μM DGJ compared with no compound. Lysates were blotted with a GAPDH antibody as a loading control. (C) Subcellular fractionation of α -NAGAL shows mature and immature forms. Gradient separation of postnuclear supernatant of FLAG-labeled α -NAGAL-expressing 293T cell homogenate was followed by Western blotting with antibodies against LAMP2 (lysosomal marker), Calnexin (ER marker), FLAG, or α -NAGAL. α -NAGAL segregates into a larger form that colocalizes with the ER marker (white arrowhead, labeled *imm*) and a smaller form that colocalizes with the lysosomal marker (black arrowhead, labeled *mat*). Fractions 2 and 10 were analyzed by blotting with FLAG and α -NAGAL antibodies, showing that the C-terminal FLAG epitope is largely removed in the lysosome. The immature form is fully deglycosylated by Endo H and PNGase F, whereas the smaller mature form is fully deglycosylated by PNGase F but shows resistance to Endo H, characteristic of glycan processing in the Golgi. (D) DGJ and DGJNac can increase the amount of mature α -NAGAL for select Schindler/Kanzaki mutations. Homogenates from HEK 293T cells stably expressing wild-type, S160C, or R329W α -NAGAL were subjected to subcellular fractionation and blotted with antibodies against α -NAGAL. Before fractionation, the gradient inputs were Western blotted with GAPDH (as a loading control). The S160C mutant shows increases in total protein and in lysosomal protein when DGJ or DGJNac is added, restoring the wild-type protein distribution. In contrast, the R329W mutant does not respond and shows the same distribution of total and lysosomal α -NAGAL with and without compound added. The white and black arrowheads mark the immature and mature forms of α -NAGAL as in C.

degradation). When DGJ or DGJNac is added to the S160C α -NAGAL cells, the amounts of overall α -NAGAL and of mature lysosomal α -NAGAL increase relative to controls (Fig. 4D). Additionally, the cellular distribution of S160C α -NAGAL in the presence of DGJ or DGJNac mimics that of wild-type α -NAGAL, suggesting that the S160C Schindler/Kanzaki mutant will respond to pharmacological chaperone therapy. In contrast, for the R329W Schindler/Kanzaki mutant, only immature α -NAGAL was detected, and there was response to neither DGJ nor DGJNac (Fig. 4D). This observation is consistent with α -GAL A defects in Fabry disease, where ~50% of the disease mutations respond to pharmacological chaperones (18).

Discussion

Since the initial structure of human α -NAGAL at 1.9-Å resolution (36), we have improved the resolution to 1.4 Å (among the highest resolution structures reported for human glycoproteins). The new structures provide unprecedented coordinate accuracy (with a maximum estimated coordinate error of 0.11 Å), revealing subtle

changes in the enzyme upon pharmacological chaperone binding, and providing a structural basis for the much higher potency of iminosugars. The structural results presented here are applicable to the entire glycoside hydrolase family 27 (41) including α -NAGAL and α -GAL A, as well as other proteins that bind iminosugars.

The high-resolution structures provide an atomic basis for the tighter binding of the iminosugars, by highlighting the critical importance of the ion pair between the amine of DGJNac/DGJ and the D156 carboxylate (Fig. 3 D–G). By revealing the interactions responsible for tight binding of ligands, the structures will aid in the design of high-affinity compounds for these and other lysosomal enzymes. We propose that future scaffolds for compounds that bind glycoside hydrolase active sites take advantage of a highly favorable ion pair with the nucleophilic carboxylate, leading to tight-binding inhibitors.

The structures and the enzyme inhibition data for the ligands allow us to construct a thermodynamic cycle for ligand binding to α -NAGAL (Fig. 5). Replacement of the ligand 2-OH by an *N*-acetyl group provides ~2 kcal/mol of additional binding energy.

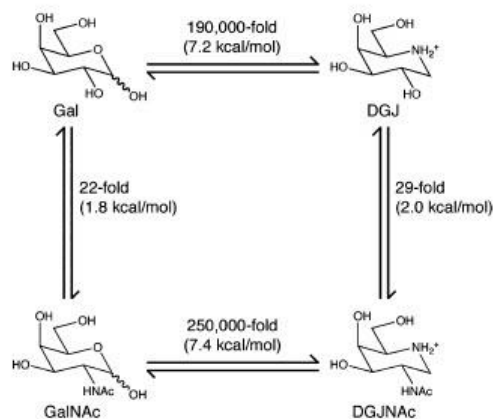


Fig. 5. Thermodynamic cycle for binding of ligands to α -NAGAL. A $\Delta\Delta G$ representation of ligand binding to human α -NAGAL shows that the presence of an *N*-acetyl group of GalNAc and DGJNAc provides ~ 2 kcal/mol of binding energy, whereas the ring nitrogen provides >7 kcal/mol. At the corners of the cycle, DGJNAc binds to α -NAGAL $>10^6$ -fold tighter than galactose does.

Replacement of the ring oxygen in hexose sugars with nitrogen in iminosugars provides >7 kcal/mol of additional binding energy. The two substitutions taken together lead to >9 kcal/mol improved binding, or better than 10^6 -fold improvement in binding affinity, a remarkable difference for highly similar compounds.

Because of the similarity of pharmacological chaperones to natural ligands, they can have off-target binding. For example, 20 μ M DGJ (a clinically relevant concentration for targeting α -GAL A in Fabry disease treatment) inhibits α -NAGAL by 70% and GLB1 by 25% in vitro (Fig. 1). In cellular assays, 10 μ M DGJ inhibits both α -NAGAL by 63% and GLB1 by 25% (18), potentially affecting the degradation of substrates with terminal α -GalNAc and β -Gal glycosides. Thus, the potency and efficacy of DGJ for α -NAGAL are both encouraging and problematic pharmacologically. In Schindler/Kanzaki disease, DGJ may be useful as a pharmacological chaperone. However, the DGJ- α -NAGAL interaction also represents unwanted off-target binding of the DGJ- α -GAL A interaction central to pharmacological chaperoning for Fabry disease. The 1.6 μ M K_i of DGJ for α -NAGAL is well below the therapeutic target concentration for treatment of Fabry disease, so off-target binding to α -NAGAL is expected. In contrast, DGJNAc contains an *N*-acetyl group that sterically precludes it from binding to α -GAL A and GLB1 (31–34), so it has improved selectivity for α -NAGAL.

In contrast to traditional enzyme inhibitors used in the clinic, effective pharmacological chaperones must meet additional criteria (17). Traditional inhibitors simply bind the active site of their target with high potency, rendering the enzyme ineffective due to competitive inhibition. On the other hand, pharmacological chaperones must not only bind to the target enzyme with high potency, but they must also dissociate from the active site of the enzyme for catalysis to occur. Additionally, an effective pharmacological chaperone must also distribute efficiently to the ER, where folding of the target enzyme occurs.

DGJNAc, the iminosugar analog of GalNAc, has been proposed as a possible pharmacological chaperone for human α -NAGAL. DGJNAc was originally hypothesized to be an inhibitor of α -NAGAL by extension of observations on the Fabry disease equivalents (DGJ and α -GAL A) (17). However, a report of the specific synthesis and testing of the compound appeared only

recently (29) [and not in a sometimes-cited earlier paper (42)]. Here, we describe biochemical, structural, and cellular studies on the interaction and chaperoning of human α -NAGAL by DGJNAc.

Our experiments point to DGJNAc and DGJ as reasonable candidates for pharmacological chaperones for Schindler/Kanzaki disease, as both compounds meet the affinity criterion suggested for developing new pharmacological chaperones. [A threshold of IC_{50} of <10 μ M is recommended for promising pharmacological chaperone lead compounds (17).] Our protease protection studies show that both DGJNAc and DGJ are capable of reducing the amount of α -NAGAL digested by protease, indicating that they shift the folding equilibrium toward the folded state. In cellular assays, we show that DGJNAc and DGJ are able to increase the amount of wild-type α -NAGAL protein produced by mammalian cells. Additionally, we show that DGJNAc and DGJ are able to increase the amount of S160C α -NAGAL (a Schindler/Kanzaki disease variant form) delivered to the lysosome compared with controls. DGJNAc and DGJ increase cellular expression of α -NAGAL by 20–30%. This increase is comparable to the chaperoning seen with the compounds DGJ and isofagomine, which have been tested in clinical trials for Fabry and Gaucher diseases (18, 43).

Because the population of patients with Schindler/Kanzaki disease is extremely small, it is unlikely to be an active target for clinical development by the pharmaceutical industry, and the neurological manifestations of the disease make enzyme replacement therapy unsuitable. We suggest that DGJ (which has already succeeded in phase I clinical trials, demonstrating its safety in humans) might be repurposed for use as a pharmacological chaperone in the Schindler/Kanzaki disease patient population. Alternatively, DGJNAc (which has higher potency than DGJ for α -NAGAL in vitro but unexpectedly shows less efficacy in cellular experiments) might be adapted into a more effective pharmacological chaperone used in compassionate care for Schindler/Kanzaki disease. For example, iminosugar compounds for Gaucher disease were developed into clinical candidates by derivatization of the heterocyclic nitrogen with alkyl substituents (44, 45).

The role of α -NAGAL in cancer is controversial (46). It is well established that α -NAGAL in the serum is a biomarker for cancers including melanoma and that α -NAGAL activity is raised in patients with large tumor burdens (47). However, others have proposed a causal link between α -NAGAL and cancer progression: α -NAGAL in the serum is hypothesized to deglycosylate vitamin D binding protein, removing the latter's ability to activate macrophages, leading to immunosuppression (48). The DGJNAc molecule we describe here, a specific inhibitor of α -NAGAL, allows for direct testing of this hypothesis. If the proposed mechanism of α -NAGAL exacerbating cancer progression proves to be correct, DGJNAc may then be useful in the treatment of cancer by reducing immunosuppression in patients.

In summary, the structures, results, and design principles reported here will improve therapeutic approaches toward protein folding diseases. We have synthesized and tested DGJNAc, a compound for the treatment of Schindler/Kanzaki disease, showing that it binds and chaperones human α -NAGAL both in vitro and in vivo. We have established that DGJ, a compound currently in clinical trials for Fabry disease, could be repurposed to treat Schindler/Kanzaki patients, who currently have no treatment options available to them. By characterizing the inhibitory properties of DGJNAc toward α -NAGAL, we describe a compound that can be used to directly test the role of α -NAGAL in immunosuppression of several cancers. If the role of α -NAGAL in immunosuppression is proven, DGJNAc, as a high-affinity inhibitor of α -NAGAL, might be useful as an anticancer compound.

Materials and Methods

For biochemistry and crystallography experiments, α -NAGAL was produced in Tn5 insect cells. Crystal structures of N201Q α -NAGAL, an engineered variant

that produces high-quality crystals (36), were solved with DGJ bound, DGJNAC bound, or with an empty active site. DGJNAC was synthesized as described (29). For limited proteolysis studies, purified wild-type or E367K α -NAGAL was incubated overnight in 4 M urea, pH 7.2, in the presence and absence of DGJ or DGJNAC. Samples were digested briefly with thermolysin and analyzed by SDS/PAGE. Cellular studies used HEK 293T cells transfected with FLAG-tagged wild-type, S160C, or R329W α -NAGAL. Density gradient cellular fractionation was used to analyze the cellular distribution of wild-type and mutant forms of α -NAGAL, and the ability of DGJ and DGJNAC to rescue mutant forms of α -NAGAL. For complete methods, see *SI Materials and Methods*.

- Dean KJ, Sung SS, Sweeley CC (1977) The identification of α -galactosidase B from human liver as an α -N-acetylgalactosaminidase. *Biochem Biophys Res Commun* 77(4): 1411–1417.
- Desnick RJ, Schindler D (2001) *N-Acetylgalactosaminidase deficiency: Schindler disease. The Metabolic and Molecular Bases of Inherited Disease*, eds Scriver CR, Beaudet AL, Sly WS, Valle D (McGraw-Hill, New York), 8th Ed, pp 3483–3505.
- van Diggelen OP, et al. (1987) Lysosomal α -N-acetylgalactosaminidase deficiency: A new inherited metabolic disease. *Lancet* 2(8562):804.
- Kanzaki T, Yokota M, Mizuno N, Matsumoto Y, Hirabayashi Y (1989) Novel lysosomal glycoaminocid storage disease with angiokeratoma corporis diffusum. *Lancet* 1 (8643):875–877.
- van Diggelen OP, et al. (1988) α -N-acetylgalactosaminidase deficiency, a new lysosomal storage disorder. *J Inher Metab Dis* 11(4):349–357.
- Wang AM, Kanzaki T, Desnick RJ (1994) The molecular lesion in the α -N-acetylgalactosaminidase gene that causes angiokeratoma corporis diffusum with glycopeptiduria. *J Clin Invest* 94(2):839–845.
- Wang AM, Schindler D, Desnick R (1990) Schindler disease: The molecular lesion in the α -N-acetylgalactosaminidase gene that causes an infantile neuroaxonal dystrophy. *J Clin Invest* 86(5):1752–1756.
- Keulemans JL, et al. (1996) Human α -N-acetylgalactosaminidase (α -NAGA) deficiency: New mutations and the paradox between genotype and phenotype. *J Med Genet* 33 (6):458–464.
- Kodama K, et al. (2001) A new case of α -N-acetylgalactosaminidase deficiency with angiokeratoma corporis diffusum, with Ménétre's syndrome and without mental retardation. *Br J Dermatol* 144(2):363–368.
- Bakker HD, et al. (2001) Human α -N-acetylgalactosaminidase (α -NAGA) deficiency: No association with neuroaxonal dystrophy? *Eur J Hum Genet* 9(2):91–96.
- Chabás A, Duque J, Gort L (2007) A new infantile case of α -N-acetylgalactosaminidase deficiency. Cardiomyopathy as a presenting symptom. *J Inher Metab Dis* 30(1):108.
- Desnick RJ, Ioannou YA, Eng CM (2001) Galactosidase A deficiency: Fabry disease. *The Metabolic and Molecular Bases of Inherited Disease*, eds Scriver CR, Beaudet AL, Sly WS, Valle D (McGraw-Hill, New York), 8th Ed, pp 3733–3774.
- Garman SC (2007) Structure-function relationships in α -galactosidase A. *Acta Paediatr Suppl* 96(455):6–16.
- Germain DP (2010) Fabry disease. *Orphanet J Rare Dis* 5:30.
- Parenti G (2009) Treating lysosomal storage diseases with pharmacological chaperones: From concept to clinic. *EMBO Mol Med* 1(5):268–279.
- Platt FM, Jayakumar M (2008) Substrate reduction therapy. *Acta Paediatr Suppl* 97 (457):88–93.
- Fan JQ (2008) A counterintuitive approach to treat enzyme deficiencies: Use of enzyme inhibitors for restoring mutant enzyme activity. *Biol Chem* 389(1):1–11.
- Benjamin ER, et al. (2009) The pharmacological chaperone 1-deoxygalactonojirimycin increases α -galactosidase A levels in Fabry patient cell lines. *J Inher Metab Dis* 32(3): 424–440.
- Leinekugel P, Michel S, Conzelmann E, Sandhoff K (1992) Quantitative correlation between the residual activity of β -hexosaminidase A and arylsulphatase A and the severity of the resulting lysosomal storage disease. *Hum Genet* 88(5):513–523.
- Schueler UH, et al. (2004) Correlation between enzyme activity and substrate storage in a cell culture model system for Gaucher disease. *J Inher Metab Dis* 27(5):649–658.
- Ries M, et al. (2005) Pediatric Fabry disease. *Pediatrics* 115(3):e344–e355.
- Fan JQ, Ishii S, Asano N, Suzuki Y (1999) Accelerated transport and maturation of lysosomal α -galactosidase A in Fabry lymphoblasts by an enzyme inhibitor. *Nat Med* 5 (1):112–115.
- Yam GH, Zuber C, Roth J (2005) A synthetic chaperone corrects the trafficking defect and disease phenotype in a protein misfolding disorder. *FASEB J* 19(1):12–18.
- Ishii S, et al. (2007) Mutant α -galactosidase A enzymes identified in Fabry disease patients with residual enzyme activity: Biochemical characterization and restoration of normal intracellular processing by 1-deoxygalactonojirimycin. *Biochem J* 406(2): 285–295.
- Lieberman RL, D'acquino JA, Ringe D, Petsko GA (2009) Effects of pH and iminosugar pharmacological chaperones on lysosomal glycosidase structure and stability. *Biochemistry* 48(22):4816–4827.
- Sugawara K, et al. (2009) Molecular interaction of imino sugars with human α -galactosidase: Insight into the mechanism of complex formation and pharmacological chaperone action in Fabry disease. *Mol Genet Metab* 96(4):233–238.
- Guce AJ, Clark NE, Rogich JJ, Garman SC (2011) The molecular basis of pharmacological chaperoning in human α -galactosidase. *Chem Biol* 18(12):1521–1526.
- Benjamin ER, et al. (2012) Co-administration with the pharmacological chaperone AT1001 increases recombinant human α -galactosidase A tissue uptake and improves substrate reduction in Fabry mice. *Mol Ther* 20(4):717–726.
- Best D, et al. (2010) Synthesis of 2-acetamido-1,2-dideoxy-D-galactonojirimycin [DGJNAC] from D-glucuronolactone: The first sub-micromolar inhibitor of α -N-acetylgalactosaminidases. *Tetrahedron Lett* 51(17):2222–2224.
- Glawar AF, et al. (2012) Scalable syntheses of both enantiomers of DNJNAC and DGJNAC from glucuronolactone: The effect of N-alkylation on hexosaminidase inhibition. *Chem Eur J* 18(30):9341–9359.
- Garman SC, Harnick L, Zhu A, Garboczi DN (2002) The 1.9 Å structure of α -N-acetylgalactosaminidase: Molecular basis of glycosidase deficiency diseases. *Structure* 10 (3):425–434.
- Garman SC, Garboczi DN (2004) The molecular defect leading to Fabry disease: Structure of human α -galactosidase. *J Mol Biol* 337(2):319–335.
- Tomasik JB, Metcalfe MC, Guce AJ, Clark NE, Garman SC (2010) Interconversion of the specificities of human lysosomal enzymes associated with Fabry and Schindler diseases. *J Biol Chem* 285(28):21560–21566.
- Ohno J, et al. (2012) Crystal structure of human β -galactosidase: Structural basis of GM1 gangliosidosis and morquio B diseases. *J Biol Chem* 287(3):1801–1812.
- Park CW, Marqusee S (2005) Pulse proteolysis: A simple method for quantitative determination of protein stability and ligand binding. *Nat Methods* 2(3):207–212.
- Clark NE, Garman SC (2009) The 1.9 Å structure of human α -N-acetylgalactosaminidase: The molecular basis of Schindler and Kanzaki diseases. *J Mol Biol* 393(2): 435–447.
- Fujimoto Z, Kaneko S, Momma M, Kobayashi H, Mizuno H (2003) Crystal structure of rice α -galactosidase complexed with α -galactose. *J Biol Chem* 278(22):20313–20318.
- Golubev AM, et al. (2004) Crystal structure of α -galactosidase from *Trichoderma reesei* and its complex with galactose: Implications for catalytic mechanism. *J Mol Biol* 339(2):413–422.
- Garman SC (2006) Structural studies on α -GAL and α -NAGAL: The atomic basis of Guce and Schindler diseases. *Bioact Biotrans* 24(1/2):129–136.
- Guce AJ, et al. (2010) Catalytic mechanism of human α -galactosidase. *J Biol Chem* 285 (6):3625–3632.
- Cartarel BL, et al. (2009) The Carbohydrate-Active Enzymes database (CAZY): An expert resource for glycomics. *Nucleic Acids Res* 37(Database issue):D233–D238.
- Ichikawa Y, Igarashi Y, Ichikawa M, Suhara Y (1998) 1-N-Iminosugars: Potent and selective inhibitors of β -glycosidases. *J Am Chem Soc* 120(13):3007–3018.
- Benito JM, García Fernández JM, Ortiz Mellet C (2011) Pharmacological chaperone therapy for Gaucher disease: A patent review. *Expert Opin Ther Pat* 21(6):885–903.
- Brumshtein B, et al. (2007) Crystal structures of complexes of N-butyl- and N-nonyl-deoxyojirimycin bound to acid β -glucosidase: Insights into the mechanism of chemical chaperone action in Gaucher disease. *J Biol Chem* 282(39):29052–29058.
- Yu L, et al. (2006) α -1-Glcyl-1-deoxyojirimycin as a pharmacological chaperone for Gaucher disease. *Bioorg Med Chem* 14(23):7736–7744.
- Rehder DS, Nelson RW, Borges CR (2009) Glycosylation status of vitamin D binding protein in cancer patients. *Protein Sci* 18(10):2036–2042.
- Greco M, et al. (2009) Serum proteomic profile of cutaneous malignant melanoma and relation to cancer progression: Association to tumor derived α -N-acetylgalactosaminidase activity. *Cancer Lett* 283(2):222–229.
- Gregory KJ, et al. (2010) Vitamin D binding protein-macrophage activating factor directly inhibits proliferation, migration, and uPAR expression of prostate cancer cells. *PLoS One* 5(10):e13428.

HUMAN ACID SPHINGOMYELINASE STRUCTURE PROVIDE INSIGHT TO
MOLECULAR BASIS OF NIEMANN-PICK DISEASE

I made contributions to the identifying crystallographic conditions for the enzyme. This article has had a large impact in the Niemann-Pick field of research as it is the first structure of human acid sphingomyelinase. It clarified the catalytic mechanism of the enzyme and identified the importance of zinc ions in the catalytic mechanism. In addition to clarifying the catalytic mechanism it also allowed for docking simulations, which modelled the collaborative process of the separate protein domains required for substrate hydrolysis.

ARTICLE

Received 23 Feb 2016 | Accepted 1 Sep 2016 | Published 11 Oct 2016

DOI: 10.1038/ncomms13082

OPEN

Human acid sphingomyelinase structures provide insight to molecular basis of Niemann–Pick disease

Yan-Feng Zhou^{1,†}, Matthew C. Metcalf², Scott C. Garman², Tim Edmunds¹, Huawei Qiu¹ & Ronnie R. Wei¹

Acid sphingomyelinase (ASM) hydrolyzes sphingomyelin to ceramide and phosphocholine, essential components of myelin in neurons. Genetic alterations in ASM lead to ASM deficiency (ASMD) and have been linked to Niemann–Pick disease types A and B. Olipudase alfa, a recombinant form of human ASM, is being developed as enzyme replacement therapy to treat the non-neurological manifestations of ASMD. Here we present the human ASM holoenzyme and product bound structures encompassing all of the functional domains. The catalytic domain has a metallophosphatase fold, and two zinc ions and one reaction product phosphocholine are identified in a histidine-rich active site. The structures reveal the underlying catalytic mechanism, in which two zinc ions activate a water molecule for nucleophilic attack of the phosphodiester bond. Docking of sphingomyelin provides a model that allows insight into the selectivity of the enzyme and how the ASM domains collaborate to complete hydrolysis. Mapping of known mutations provides a basic understanding on correlations between enzyme dysfunction and phenotypes observed in ASMD patients.

¹Protein Engineering Department, Biologics Research, Sanofi, Framingham, Massachusetts 01701, USA. ²Department of Biochemistry & Molecular Biology, University of Massachusetts Amherst, Amherst, Massachusetts 01003, USA. [†]Present address: Analytical Discovery Therapeutics, Shire, Lexington, Massachusetts 02421, USA. Correspondence and requests for materials should be addressed to Y.-F.Z. (email: Yanfeng.Zhou@sanofi.com) or to R.R.W. (email: Ronnie.Wei@sanofi.com).

Acid sphingomyelinase deficiency (ASMD) causes abnormal accumulation of sphingomyelin in lysosomes, resulting in the progressive, life-threatening disorder historically referred to as Niemann–Pick disease types A (NPD A), presenting with an early-onset neuronopathic phenotype, and type B (NPD B), the non-neuronopathic form¹. ASMD is a result of mutations in acid sphingomyelinase (ASM, E.C. 3.1.4.12), which hydrolyzes sphingomyelin to ceramide and phosphocholine (Fig. 1a)². Recent studies on this rare disease illustrate that ASM plays an important role in the ceramide-mediated signalling pathway. Recombinant human ASM expressed from Chinese hamster ovary (CHO) cells is being developed as enzyme replacement therapy (ERT) for the non-neurological manifestations of ASMD, and is currently in the phase 1/2 pediatric trial and phase 2/3 trial for adults^{3,4}.

ASM is highly conserved in animals, from *C. elegans* to human, with over 35% protein sequence identities (Supplementary Fig. 1). Sequence analysis suggests ASM is a multi-domain protein, including a saposin domain, a proline-rich linker, a metallo-dependent phosphatase catalytic domain, and an ill-defined C-terminal domain (Fig. 1b)⁵. Saposin proteins are sphingolipid activator proteins that present lipids from membranes to the active sites of various enzymes in an acidic cellular compartment⁶. The saposin domain in ASM is sufficient to support the hydrolysis of sphingomyelin without external saposin proteins^{7,8}. In addition, ASM differs from neutral and alkaline sphingomyelinases in both domain architecture and protein sequence, with identities <10%.

Previous biochemical characterizations reveal that proper function of human ASM (UniProt database ID: P17405) requires zinc ions, multiple post-translational modifications, and acidic pH (ref. 5). On the basis of its localization, two forms of ASM originating from one gene have been reported, an intracellular lysosomal form and an extracellular secreted form⁹. Zinc ions are prerequisites for activity of both forms¹⁰. The lysosomal ASM is preloaded with zinc, while the secreted form requires exogenous zinc⁹. Also, human ASM has six potential *N*-linked glycosylation sites, and its *in vitro* activity pH optimum is around 5 (ref. 11). However, it is not clear how zinc and other physiological factors are involved in the enzyme function.

Despite its clinical significance, limited structural studies on ASM have been reported that elucidate its substrate selectivity and enzymatic mechanism. A model of human ASM catalytic domain has been generated based on a distant homolog, purple acid phosphatase¹². Two recent crystal structures of catalytic domain of sphingomyelinase phosphodiesterase like 3a (SMPDL3A), sharing 31% identities to human ASM, are reported^{13,14}. However, the understanding on ASMD is limited to the metal-binding site due to the low sequence identities and absence of the *N*-terminal saposin domain.

Here we report two crystal structures of recombinant human ASM and observe how co-factor zinc ions and a product phosphocholine are coordinated by the enzyme. The multi-domain structures, representing two states in the catalytic cycle, shed light on the mechanism of how natural substrate sphingomyelin is recognized and degraded. Moreover, currently reported ASMD mutations in the UniProt database were mapped onto the structure to build genotype–phenotype correlations with atomic details.

Results

Overall structure. Full-length human ASM was expressed from HEK293S *N*-acetylglucosaminyl transferase 1-deficient (Gnt1[−]) cells with Man₅ type glycans. ASM in cell culture media was purified with a monoclonal antibody coupled affinity column,

followed by size exclusion column. Purified protein was crystallized at pH 5.5. The recombinant human ASM being developed for ERT, olipudase alfa, was also crystallized without any protease in the crystallization drops. Olipudase alfa was expressed in CHO cells with complex glycans¹⁵. The olipudase alfa crystals diffracted to around 4 Å at synchrotron (Table 1), and clear glycosylation and zinc cofactors densities were observed (Supplementary Fig. 2). Its refined structure was essentially identical to ASM from HEK cells. The descriptions in this report are focused on ASM from HEK cells unless specified.

ASM shares <10% sequence identity to structurally known homologues. Therefore, the holoenzyme structure was determined using single-wavelength anomalous diffraction (SAD) from a Pt derivative at 2.43 Å, and the phases was extended to 2.25 Å using a Zn-SAD data set. The final model was refined to R_{free} of 0.203 (Table 1). Clear electron density allowed us to build all residues from W84 to M611 as well as two zinc ions, six *N*-linked glycans, and eight disulfides (Fig. 1c). The same residue range is visible in the olipudase alfa structure, suggesting ASM *N* and *C* termini are inherently disordered and flexible. The structure includes an *N*-terminal saposin domain (residues 84–167), a proline-rich linker (168–195), a catalytic metallophosphatase domain (196–538), and a helical C-terminal domain (539–611) (Fig. 1).

Saposin domain and proline-rich linker. The presence of saposin domain and proline-rich linker distinguishes ASM from all the other structurally known proteins. A number of crystal structures have demonstrated that saposin proteins undergo open and close conformational changes in response to lipid binding (Supplementary Fig. 3)^{16–18}. The four helices in ASM saposin domain were observed in an extended open conformation in which helices H1 and H2 run nearly antiparallel to H3 and H4 (Fig. 1c). The helices are stabilized by two disulfide bonds connecting the saposin domain *N* and *C* termini and a third disulfide between H2 and H3. As a result, helices H3 and H4 form tight pairwise helical interactions with H1 and H2. The kinked H3 helix mediates saposin and catalytic domain interactions. In contrast, the H1 and H4 helices extend away from the catalytic domain, and there is no direct contact between H4 and catalytic domain (Fig. 1c). The four ASM saposin helices are amphipathic. The concave inner surface is hydrophobic and faces towards the zinc ions in active site, while the convex outer surface is lined with hydrophilic residues and exposed to solvent.

A 30-residue linker connects the saposin domain to the catalytic domain (Fig. 1c,f). The structure reveals that this linker spans 57 Å and covers 1128 Å² surface of the catalytic domain. The linker is bent in the middle at position T180, resulting in a 97° kink. The linker can be divided to prolineless *N*-terminal half and proline-rich *C*-terminal half, based on amino-acid features. In the prolineless half, there is one glycosylation and several serine residues, but no proline. In contrast, the *C*-terminal half is dominated by proline, with one triple and two double proline motifs (P¹⁸²KPPPKPPSPAP). The prolineless half runs along the α5 and α6 in catalytic domain and is stabilized through extensive hydrophobic interactions with α5 and α6. The proline-rich half is extended and wraps around loops following α6, β8 and β10 in catalytic domain.

Catalytic domain. The catalytic domain structure shows two layers of β-sheets, both six-stranded, form the central core, with dimensions 47 Å × 50 Å × 53 Å. Six α-helices and the helical C-terminal domain flank the β-sheets resulting in a four-layer α/β/β/α sandwich architecture. The catalytic domain starts at β1 in the center of the first sheet and ends at β12 at the edge of this

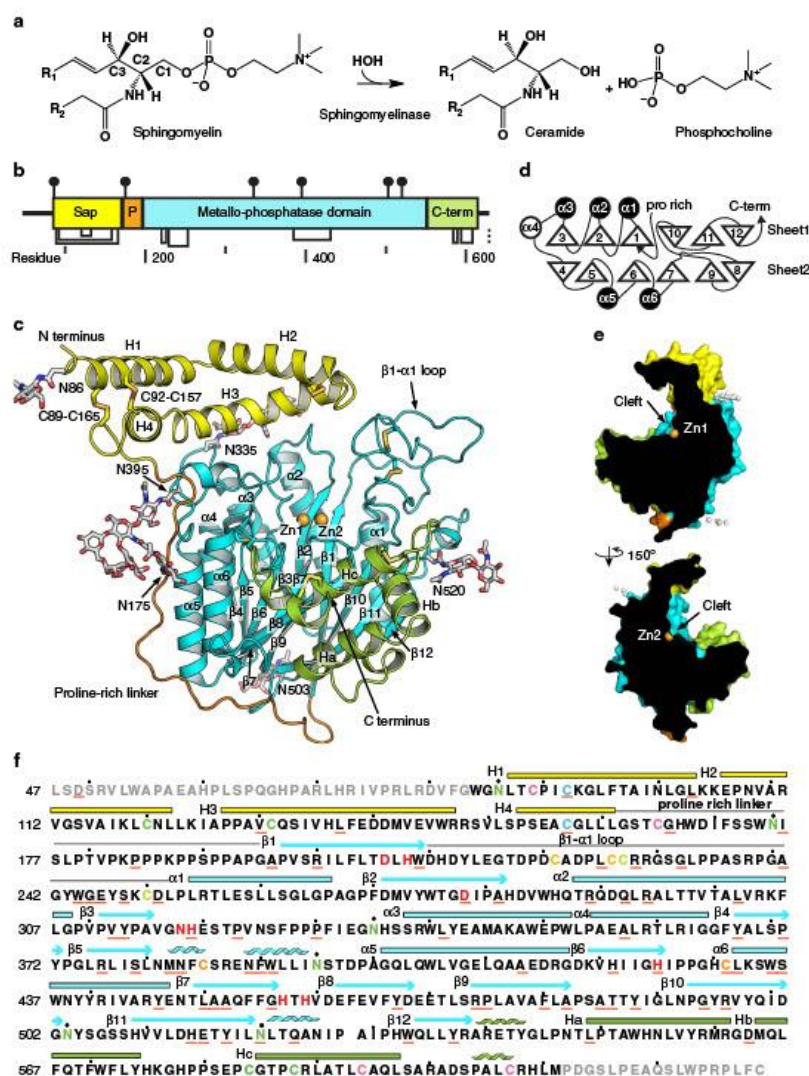


Figure 1 | Structure of human ASM. (a) Sphingomyelin degradation catalyzed by ASM. (b) Diagram of ASM domains drawn to scale, with yellow saposin, orange proline-rich linker, and green C-terminal domain. (c) Overall structure of ASM. Glycosylation and disulfide bonds are in sticks. Two zinc ions are in spheres. (d) Topology of ASM. Upward and downward pointing triangles indicate the N to C terminus orientation of β strands. Open and closed circles indicate upward and downward going helices, respectively. (e) Cross-section illustration of cleft on ASM surface, with zinc and glycans as spheres. Arrow points to the center of substrate binding cleft. (f) ASM protein sequence with structural features. Disulfides are shown as matched cysteine colors. Residues that are not visible in the structure are grey. Underlined residues are reported ASMD mutation sites. Catalytic residues are bold and red. Glycosylated asparagine residues are green.

sheet (Fig. 1d). The topological arrangement suggests catalytic domain adopts a metallophosphatase fold (SCOP d.159) and belongs to a calcineurin-like phosphoesterase superfamily in Pfam database (PF00149). ASM catalytic domain and C-terminal

domain shares high structural similarity to SMPDL3A (refs 13,14), with identical topology and root mean square deviation (r.m.s.d.) of 1.4 Å of all C α atoms. Superimposition of ASM catalytic domain to pig purple acid phosphatase (PAP, PDB

Table 1 | Data collection and refinement statistics.

| Data collection | Holo ASM | Pt derivative | Phosphocholine | Olipudase alfa |
|-------------------------------------------------|------------------------|------------------------|------------------------|------------------------|
| X-ray wavelength (Å) | 1.2830 | 1.0721 | 1.5418 | 1.2825 |
| Resolution (Å) | 43.81-2.25 (2.33-2.25) | 42.33-2.43 (2.52-2.43) | 31.18-2.50 (2.64-2.50) | 45.00-3.65 (3.74-3.65) |
| Space group | P6 ₄ 22 | P6 ₄ 22 | P6 ₄ 22 | I222 |
| No. of reflections | 88030 | 68691 | 33985 | 118574 |
| <i>Cell dimensions</i> | | | | |
| a, b, c (Å) | 132.5, 132.5, 189.8 | 132.2, 132.2, 189.3 | 131.6, 131.6, 188.6 | 191.0, 230.9, 252.3 |
| α, β, γ (°) | 90.0, 90.0, 120.0 | 90.0, 90.0, 120.0 | 90.0, 90.0, 120.0 | 90.0, 90.0, 90.0 |
| <i>I/σ(I)</i> | 22.9 (1.9) | 22.1 (2.0) | 12.6 (3.2) | 7.44 (0.34) |
| Rmerge | 0.092 (0.732) | 0.081 (0.735) | 0.200 (0.822) | 0.229 (5.35) |
| Completeness (%) | 99.5 (94.9) | 99.0 (90.9) | 98.9 (100.0) | 98.9 (86.3) |
| Redundancy | 10.3 (5.0) | 7.3 (4.9) | 17.3 (13.9) | 6.9 (5.4) |
| CC1/2 in highest shell | 0.669 | 0.743 | 0.852 | 0.108 |
| <i>Phasing and refinement</i> | | | | |
| Phasing figure of merit | — | 0.251 | — | — |
| $R_{\text{work}}/R_{\text{free}}$ | 0.184/0.203 | — | 0.190/0.222 | 0.248/0.254 |
| Monomers per asymmetric unit | 1 | 1 | 1 | 3 |
| Composition per asymmetric unit | | | | |
| Amino acid/sugar | 527/14 | | 527/13 | 1581/29 |
| Water/SO ₄ | 177/11 | | 261/11 | 0/0 |
| Zn/Phosphocholine | 2/0 | | 2/1 | 6/0 |
| rmsd | | | | |
| Bond (Å) | 0.009 | | 0.018 | 0.013 |
| Angle (°) | 0.902 | | 0.885 | 0.931 |
| Ramachandran plot (favoured/allowed/outlier) | 96.4/3.6/0.0 | | 95.4/4.6/0.0 | 95.6/4.2/0.2 |

ASM, acid sphingomyelinase; rmsd, root mean square deviation.
Holo ASM and Pt derivative are anomalous data collected at Zn peak and Pt L-III edge, respectively. Phosphocholine data set is non-anomalous. Olipudase Alfa data set is collected at Zn peak. Values in parentheses are for the highest-resolution shell.

IUTE) and bacteria neutral sphingomyelinase (PDB 2DDR) shows r.m.s.d. values of 3.1 and 3.9 Å, respectively. The unique topological arrangement of β9–10 and β10–12 strands makes the catalytic domains in ASM and SMPDL3A differ from currently identified members in metallophosphatase fold and neutral sphingomyelinases (DNase I-like fold, SCOP d.151)¹⁹.

One feature of the ASM metallophosphatase domain is the long loops following certain β strands and extending toward the zinc ions (Fig. 2). These loops not only mediate zinc and substrate binding but also form the interfaces to the proline-rich linker, saposin and C-terminal domains. H3 helix in saposin domain forms extensive contacts to β1-α1, β2-α2, β3-α3 and β5-α5 loops (Fig. 2a,b). The interface between H3 and these loops is 700 Å² and over 30 Å long. It is worth noting that many ASMD mutations occur on the interface residues, such as L137P in H3, P323A on the β3-α3 loop, ΔF390 and W391G on the β5-α5 loop. I170 and F171 on proline-rich linker interact with V150 on saposin H3 and L393 on β5-α5 loop, resulting in a turn towards the catalytic domain. The proline-rich linker is further stabilized by hydrophobic interactions between W174, I176, L178 and the α5 α6 helices (Fig. 2c). Moreover, the 44-residue β1-α1 loop extends out from the catalytic domain and folds into a small globular unit next to the H2-H3 turn in saposin domain (Fig. 2b). The conformation of this long loop is enforced by internal two pairs of charged interactions, two pairs of disulfides, and hydrophobic interactions. This loop is also interacting with β2-α2, β9-β10 and β11-β12 loops.

The catalytic domain shares a 1,632 Å² interface with the C-terminal domain (Fig. 2d). The C-terminal domain starts after β12 in sheet1 and contains four helices, Ha-c, and a 3₁₀ at the C terminus. The β8-9 strands at the edge of sheet2 deeply inserts into the space between these helices. Moreover, C-terminal domain contributes to the stabilization of the active site conformation, especially the β7-β8 loop. It has been reported

that ΔT592 mutation causes severe ASMD in patients²⁰. T592 is in the middle of Hc helix, which forms extensive interactions with the catalytic domain (Fig. 2d). Thus, we predict ΔT592 mutation has severe influence on ASM folding.

Zinc binding in the active site. The long loops in the catalytic domain, together with saposin and C-terminal domains, create a deep and wide open cleft (Fig. 1c,e). The saposin and C-terminal domains are on the opposite sides of the catalytic domain. This is the only large and deep cleft (15 Å × 30 Å × 10 Å) on ASM. Six N-linked glycans scatter on the surface and all distribute far away from this cleft.

Two neighbouring zinc ions are identified in the center of the cleft. The presence of zinc was confirmed by an X-ray fluorescence scan, and strong zinc anomalous signal has been observed in the data (Supplementary Fig. 4b). The location of zinc is at the C-terminal end of strands β1, β2, β6 and β7 (Figs 1c and 3a). The two zinc ions are separated by 3.5 Å and referred to as Zn1 and Zn2. Both zincs have trigonal bipyramidal geometry. D278 bridges between Zn1 and Zn2 at the axial positions, while H457 and H459 are in the other axial positions. Two of the equatorial positions are occupied by N318 and H425 for Zn1 and D206 and H208 for Zn2. Mutation of H425A or D206A abolishes ASM activity¹², which is consistent with our structure. The conformations of the axial ligands H457 and H459 are constrained by side chains from helices in the C-terminal domain (Fig. 2d).

The precise side-chain arrangement around the metals facilitates recruitment of a key catalytic water molecule to the third equatorial position bridging between Zn1 and Zn2. The presence of this water is confirmed by observation on omit Fo-Fc map (Supplementary Fig. 4b). The catalytic water is stabilized and positioned by hydrogen bonds to the main chain carbonyl of

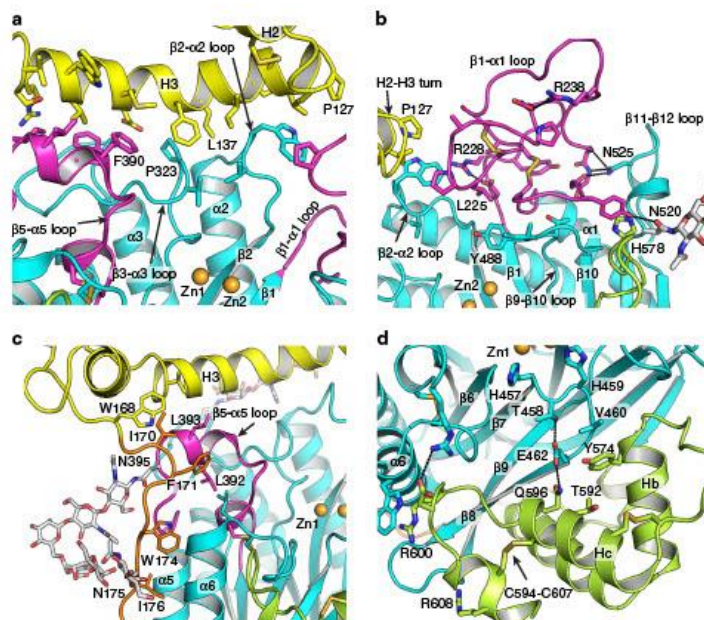


Figure 2 | Structural details of the substrate binding cleft. (a) Interface between H3 in saposin and the $\beta 5$ - $\alpha 5$, $\beta 3$ - $\alpha 3$, $\beta 2$ - $\alpha 2$ loops in the catalytic domain. (b) $\beta 1$ - $\alpha 1$ loop in catalytic domain near by the tip of H2-H3 turn in saposin domain. (c) Hydrophobic interactions between N-terminal part of the proline-rich linker and catalytic domain. (d) Interface between the C-terminal domain and catalytic domain.

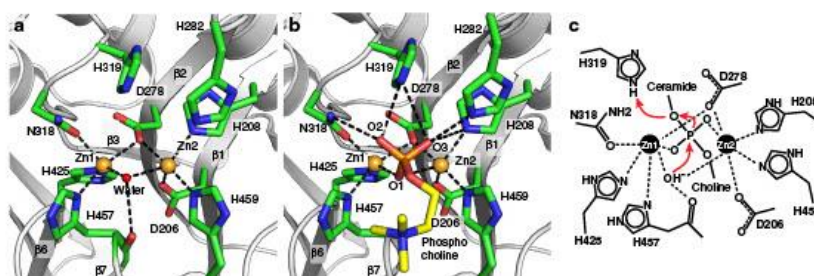


Figure 3 | ASM active site. (a,b) Active sites of the holo (a) and phosphocholine bound (b) structures. α atoms are shown in grey ribbon. Two zinc atoms are shown as gold spheres, and one water molecule as red sphere. Phosphocholine is shown as sticks with yellow carbon, red oxygen and orange phosphorus. (c) Proposed catalytic mechanism in 2D diagram. Red arrows indicate electron relay during the cleavage. Only phosphate group in sphingomyelin are shown in details. Ceramide and choline indicate orientation of substrate in ASM.

H457 (Fig. 3a). There is an asymmetric arrangement in which the carboxylate oxygen of D278 is slightly closer to Zn2 (D-Zn1 2.5 Å, D-Zn2 1.9 Å), while the catalytic water is closer to Zn1 (Wat-Zn1 1.9 Å, Wat-Zn2 2.6 Å). In contrast, the coordination of this water in PAP and SMPDL3A is more symmetric (Supplementary Fig. 5)¹⁴.

Phosphocholine binding. To better understand the ASM catalytic mechanism and how the Zn ions mediate hydrolysis, we soaked a product phosphocholine into ASM crystal and solved

the structure to 2.5 Å (Table 1). Clear density was observed near the zinc atoms in the Fo-Fc difference map (Supplementary Fig. 4c). One phosphocholine was built into the electron density, with the choline pointing toward the C-terminal domain (Fig. 3b). The phosphoryl group is tightly coordinated by both zinc ions and the neighbouring catalytic residues. Oxygen (O1) on the phosphoryl group replaces the position of the catalytic water molecule in the holo ASM structure. N318 and H319 on $\beta 3$ - $\alpha 3$ loop, H208 on $\beta 1$, and H282 on $\beta 2$ - $\alpha 2$ loop interact with two oxygen atoms (O2 and O3) on the phosphoryl group.

Discussion

Sphingomyelin is a key membrane component from bacteria to human. Understanding of how sphingomyelin is recognized and hydrolysed by ASM greatly enhances our knowledge of sphingolipid metabolism and associated human diseases. Due to low sequence identity, reliable prediction of human ASM function cannot be made based on structurally known proteins. The human ASM structures we report here illustrate how the co-factor zinc ions activate the enzyme and how domain arrangement contributes to substrate specificity.

A putative catalytic mechanism can be deduced for ASM, based on the structures (Fig. 3c). We crystallized two states in the ASM catalytic cycle, the holoenzyme state, and product phosphocholine bound state. We propose that ASM catalysis is an associative general acid–base type reaction.

The active site in ASM defines a stringent specificity for zinc and the phosphate group of sphingomyelin. The two zinc ions are coordinated by four histidines, two aspartic acids, one asparagine and the catalytic water molecule. During the substrate recognition, we speculate that two zinc co-factors serve as anchors for the two phosphate oxygens from sphingomyelin. The water can be deprotonated to hydroxide, because the two close zinc ions extract proton away from this water. Following that, the hydroxide initiates nucleophilic attack on the phosphorus of sphingomyelin. The arrangement of the active site allows the phosphate oxygen on the ceramide side to extra a proton from either H319 or H282 (Fig. 4a). The H282 imidazole ring is stabilized by π - π stacking interaction with Y488 and polar interaction to D251 carboxylate group, while the H319 is involved in water mediated hydrogen bonding. As a result, H319 is a more energy favoured hydrogen donor in the reaction. We propose the imidazole ring on H319 donates a proton to the oxyanion of the ceramide leaving group. It is worth noting that the cleft is open and solvent accessible, which allows quick exchange of protons on both H319 and the nucleophile water in the regeneration step. Consistent with this hypothesis, mutation of H319 to tyrosine causes severe ASMD²¹.

The proposed mechanism is further supported by observations in other di-metal-binding metallophosphatases. Despite <10% protein sequence identities, ASM shares a conserved catalytic site and di-metal binding pattern with other metallophosphatases, such as mammalian purple acid phosphatase (Supplementary Fig. 5)²². The di-metal bridging water molecule is found at the active site in this class of catalytic domains/enzymes. In addition, the phosphate location is also conserved, with the same trigonal bipyramidal coordination geometry.

Most of the sphingolipid degradation enzymes require activator proteins, like saposins, to present substrates to the catalytic domain. ASM has its own saposin domain for sphingomyelin recognition. Our structure provides direct observation on how saposin and catalytic domain are arranged for sphingomyelin recognition.

Major sphingomyelin species in biological membranes have long acyl chains, with three dominant lengths, C-16, C-18, and C-24 (ref. 23). The substrate binding cleft in ASM is created by the catalytic domain in the middle, and saposin domain and C-terminal domain at two opposite sides. Zinc ions and phosphocholine occupy the center and the C-terminal domain side of the cleft. The saposin concave surface facing towards the active site is rich in hydrophobic residues (Fig. 4a) and not highly charged (Fig. 4b), which can accommodate the hydrophobic ceramide chains in sphingomyelin. The distance between Zn1 and the phenyl ring of F138 in H3 is 14 Å, which is close to the 18 Å length of an extended 16 carbon chain Ahn *et al.*¹⁶ have shown that hydrophobic chains of a phosphatidylethanolamine are accommodated within a hydrophobic cavity at the middle of the kinked H3 helices in saposin B protein (Supplementary Fig. 3c). Recent crystal structure of saposin A lipoprotein discs and cryoelectron microscopic study show saposin proteins can wrap around lipids and membrand protein and form lipoprotein-like nanoparticles^{18,24}. These observations indicate saposin might be functional as long as it is under open conformation, regardless its oligomeric states. With limited knowledge of binding between saposin proteins and lipids, we did docking of C-16 sphingomyelin onto ASM to visualize how the ceramide chains can be placed between the active site and the saposin domain (Fig. 4). The saposin is under open conformation in our structures. Even with the constraints of phosphocholine and the kinked region in H3 as docking site, the flexible ceramide chains were placed various positions (Fig. 4b), instead of form a specific interaction network with saposin domain. The modelling result suggests that sphingomyelins can fit well into the cleft on ASM surface. We propose it undergoes close–open conformational changes in response to substrate binding like other saposin proteins. The proline-rich linker could act as a hinge that switches the saposin domain conformational changes, facilitating presentation of substrate to the catalytic domain.

The chemical composition of the hydrocarbon chains allows sphingomyelin to fit better into ASM cleft than other phospholipids. Oninla *et al.*²⁵ has found that ASM catalyses degradation of C18-sphingomyelin 10-fold faster than dipalmitoyl-phosphatidylcholine, although they share an

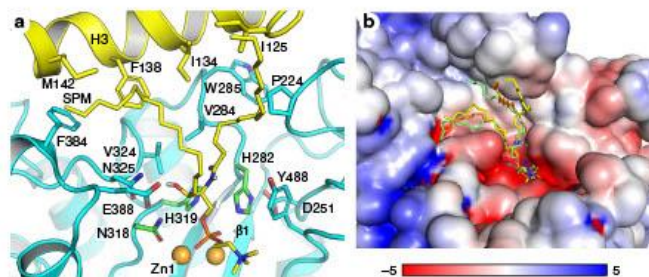


Figure 4 | Docked model of sphingomyelin on ASM. (a) Surface residues around C-16 sphingomyelin. Zinc atoms are shown gold spheres. Sphingomyelin is coloured as yellow carbon, blue nitrogen, red oxygen and orange phosphorus. (b) Electrostatic potential surface calculated at pH 5.0 with PropKa (ref. 46) and PDB2PQR (ref. 47). Red and blue surfaces correspond to negative and positive electrostatic potential scaled from $-5 k_B T$ to $5 k_B T$. Three docked poses were shown on the surface, and the yellow pose was selected and shown in a.

identical phosphocholine head group. Sphingomyelin in nature is synthesized with two stereo chemical requirements in the hydrocarbon chain: the double bond in sphingosine is in *trans* configuration, and the C2 amine and C3 hydroxyl must have three relationship²⁶. We speculate that these are the two key factors for proper substrate alignment on ASM. First, the amine at C2 and hydroxyl at C3 positions are good hydrogen-bond donors (Fig. 1a), while the glycerol backbone in phosphatidylcholine has fewer hydrogen-bond donors. N325, E388 and Y488 in ASM are within 7 Å from the zinc ions and in position for hydrogen bonding with substrate (Fig. 4a), and N325 is highly conserved among different species (Supplementary Fig. 1). Hydrogen bonding plus charged interactions with zinc ions together make substrate recognition highly specific. Second, naturally existing sphingomyelin is much more saturated than phosphatidylcholine, which normally contains one or more double bonds in *cis* configuration in the middle of the acyl chains²³. The rigidity of the *cis* double bond causes the acyl chain to bend and restricts the conformational freedom. As a result, we predict the recognition and presentation of phosphatidylcholine and other phospholipids to ASM active site is slower than that of sphingomyelin.

ASMD is caused by loss of function mutations in ASM. There are ninety known mutations listed in UniProt database and summarized in supplementary Table 1. All mutations are covered in our structure, except D49V located in the signal peptide.

Mapping mutations onto one dimensional (1D) protein sequence and 3D structure reveals interesting occurrence pattern in human ASM (Fig. 5). Overall, 82% of the mutations are located in the catalytic domain, resulting in 21.6% mutation rate in protein sequence. In contrast, the mutation rates in the saposin, proline-rich linker, and C-terminal domains are no > 11% (Fig. 5a).

Those mutations are predicted into two categories according to their disruptive effects on catalytic activity and protein folding (Fig. 5b; Supplementary Table 1). First, the catalytic group includes mutation of residues that are directly or indirectly involved in zinc binding or phosphocholine binding in the specificity pocket. The indirect residues disrupt the folding near the active site and thus affect catalysis. For direct coordination, D278A mutation eliminates the coordination to both zinc atoms. H319Y not only introduces a larger group to the compact active site but also loses the imidazole group which plays an essential role in the phosphodiester bond cleavage. H425 points the imidazole ring towards zinc atoms, and mutation to the larger arginine completely disrupts zinc coordination. There are many other mutations that affect the active site indirectly. For example, the A281T mutation affects H319, and W209R mutation disrupts the $\beta 1$ stability and affects D206 and H208, which coordinate Zn2. Q292 and H319 are involved in a buried hydrogen-bond network including A281 and Q287. The Q292 amide forms a 3.2 Å hydrogen bond with the H319 carbonyl. The Q292K mutation brings a longer side chain into this well-defined

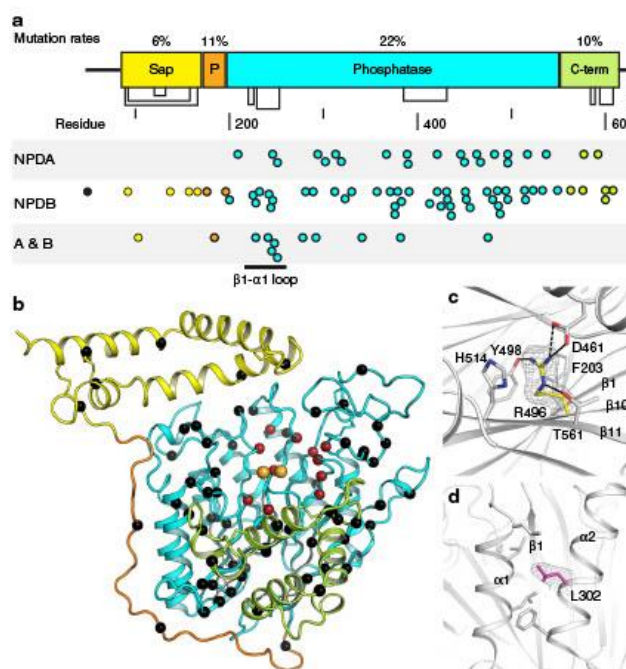


Figure 5 | Mutations mapped on structure. (a) Distribution of mutations on human ASM 1-D sequence. Each point represents one reported mutation in the UniProt database. Points are coloured according to domains. (b) Distribution of mutations on 3D structure. α atoms of mutated residues are shown as spheres. Activity-related mutations are red, and folding related ones are black. Zinc atoms are gold spheres. (c) Network of R496 with neighbouring residues in the hydrophobic core between β -sheets. (d) Hydrophobic interactions mediated by L302 between $\alpha 1$ and $\alpha 2$ helices. In c-d, 2Fo-Fc density for side chains of mutation sites are shown and contoured at 1 σ .

hydrogen bonding network and perturbs the catalytic role of H319.

Among the mutations in the catalytic domain, there are two clusters in the 3D structure. One is on the tightly folded $\beta 1$ - $\alpha 1$ loop (Fig. 2b). Mutations like A241V, G242R, G245S, introduce larger side chains and disrupt local folding. The other cluster spreads onto $\beta 4$, $\beta 5$, $\beta 5$ - $\alpha 5$ loop and $\beta 6$ - $\alpha 6$ loop in the protein sequence. According to the structure, these mutations are in close proximity to H3 in the saposin domain, a critical location for binding of sphingomyelin.

The second group of mutations is on residues that are widely distributed in the structure and distant from the active site. These mutations destabilize local interaction networks and affect enzyme stability and folding (Fig. 5b). Most of the residues are sequestered from solvent (Supplementary Table 1), and some surface exposed ones are heavily involved in interactions with neighboring residues in the folded structure. R496L and L302P mutations are two of the three prevalent mutations in the Jewish population and cause NPD A. Both R496 and L302 are highly conserved among species (Supplementary Fig. 1). The guanidinium group of R496 forms extensive interactions with neighboring residues, and stabilize the buried core between the β -sheets (Fig. 5c). L302 is located in the middle of helix $\alpha 2$, and the local hydrophobic interactions maintain the stability of $\alpha 1$ and $\alpha 2$, which are next to $\beta 1$ strand in the active site (Fig. 5d). In addition, mutation to a proline, which often interrupts helical structures, further decreases the local stability.

Deletion of R608 (Δ R608) is the most common mutation in NPD B, and patients with homozygous Δ R608 mutations have mild phenotype. R608 is located after the C594-C607 disulfide bond on a 3_{10} helix. Residues after M611 are missing electron density and presumably disordered. The residues after R608 show very low sequence identities among ASM from different species (Supplementary Fig. 1), but a free cysteine at the C-terminal end is conserved and the length of this region in most species is around 20 amino acids. Mutation or deletion of the last residue C629 in human ASM is known to increase ASM activity and the cysteine is proposed to form direct contacts to active site²⁷. It has been reported in other metallophosphatase, like PhoD in *Bacillus subtilis*, that the C-terminal end folds back to active site to inhibit enzyme activity²⁸. We predict that deletion of R608 will disturb the C terminus and affect its activity. In contrast to Δ R608, Δ T592 has been reported as an NPD A mutation. As described earlier, T592 forms direct contacts with active site. Δ T592 will disrupt the interactions between the C-terminal domain and the catalytic domain, resulting in a perturbed active site and severe enzyme deficiency (Fig. 2d).

Taken together, we predict how individual mutation affect enzyme folding and function based on the atomic resolution structure of ASM. This information allows us to better understand gene mutations and ASMD phenotypes in the context of enzyme folding and catalytic activity. As abovementioned, we observe consistent correlations between disease severity and a disruptive effect on the ASM active site or folding. The structural information could serve as a guide to analyse and predict the phenotypic outcome in ASMD individuals. The structures provide us the ability to understand the molecular mechanism behind diseases. Following tests of thermostability and enzymatic assay with ASM mutants will help us better reveal the predicted correlations.

The abovementioned structures open a door for us to decipher ASM catalytic mechanism and mutations in ASMD patients. However, many questions remain unanswered. First one will be how ASM interacts with membrane and reaches sphingomyelin. ASM in solution is monomeric, and it has not been reported that ASM needs to dimerize to function. New evidence with different

techniques may better reveal the details. Another question is how ASM is inhibited by the anti-depressants drugs. Model has been proposed that those tricyclic drugs would insert into membrane and adjust charges on the surface, which will interfere the electrostatic adherence of ASM to membrane²⁹. The structure shown here does not conflict with this model. However, these hydrophobic tricyclic drugs could also potentially bind to the hydrophobic surface in saposin domain thus prevent saposin-sphingomyelin interaction. A following co-crystal structure of ASM or saposin domain with these small molecules would provide more insights. Moreover, ASM activity is regulated by different modifications and physiological partners, such as phosphorylation on S508 and association with acid ceramidase^{30,31}. We observe S508 on $\beta 10$ - $\beta 11$ loop is partially buried by the proline-rich linker and hydroxyl group is hydrogen bonded to D501 in our structures. This serine residue is not phosphorylated and located in a position distant from substrate binding site. A crystal structure with phosphorylated S508 may address the question. ASM could also be activated by redox conditions and possible through the last free cysteine C629 (refs 27,32). We see that most of the disulfides in ASM are solvent accessible and the C-terminal C629 are exposed and too flexible to be seen in our structures. The exposure of these residues may make ASM sensitive to redox conditions.

We present the atomic resolution structures of human ASM here, with detailed information about zinc and phosphocholine binding. The structures allow us to interpret zinc activation and substrate hydrolysis in ASM. In addition, the results provide insights about how catalytic domain and its neighboring saposin, proline-rich linker, and C-terminal domains together create a cleft specific for sphingomyelin binding. These data also give us the opportunity to start correlating mutations in patients and enzyme deficiency.

Methods

Protein expression. Codon optimized DNA encoding human ASM (M1 to C629) was cloned into the NheI and BamHI sites of plasmid pIRES2-EGFP and confirmed by sequencing. HEK293S Gnt1⁻ cells (ATCC CRL-3022) were transfected using polyethylenimine³³. Single colonies were selected in Dulbecco's modified Eagle's medium and 5% fetal bovine serum with 1 mg ml⁻¹ G418. Stably transfected cells were cultured in suspension in Freestyle 293 serum-free medium, and supernatants were harvested 7 days after expansion.

Olipudase alfa was produced at Sanofi Genzyme²⁷. Human full-length ASM cDNA was cloned into a dihydrofolate reductase selection vector and stably expressed in CHO-DXB11 cells (ATCC CRL-9096).

Protein purification. HEK293S Gnt1⁻ cell expressed ASM was purified from the supernatant using antibody affinity and gel-filtration chromatography. Anti-ASM monoclonal antibody was coupled to AminoLink Plus resin (Thermo Scientific, USA) following the manufacturer's protocol. ASM supernatant was adjusted to 20 mM Bis-Tris pH 6.0, 0.15 M sodium chloride and 0.1 mM zinc acetate, and incubated with resin overnight at 4 °C. ASM was eluted with 20 mM sodium citrate pH 3.0, and neutralized with 0.1 M Bis-Tris pH 6.5. The eluate was further purified using Superdex 200 10/300 GL (GE healthcare, USA) in 20 mM Bis-Tris pH 6.0, 0.15 M sodium chloride (Supplementary Fig. 6). Peak fractions were concentrated to 20 mg ml⁻¹.

Olipudase alfa was purified from culture media using hydrophobic interaction chromatography followed by ion exchange chromatography. Eluate was concentrated and subjected to gel-filtration as described in purification of ASM from HEK293S Gnt1⁻ cells. The monomer peak was eluted 0.9 ml earlier than ASM from HEK293 S Gnt1⁻ cells (Supplementary Fig. 6). Peak fractions were collected and concentrated to 20 mg ml⁻¹ for crystallization.

Crystallization. ASM was crystallized with *in situ* proteolysis of endoprotease Glu-C at 1:200 protease:ASM mass ratio. Crystals were obtained in sitting drops at 21 °C using a well solution of 1.5 M ammonium sulfate, 0.1 M sodium acetate pH 5.0-5.5, 12 % glycerol. Olipudase alfa crystals were obtained at 21 °C in the same condition, but without any protease in the drops.

Platinum derivatives were prepared by soaking native crystals in well solution supplemented with 10 mM K₂PtCl₄ for 10 min at 21 °C. Derivatized crystals were

back soaked for 5 min in a cryo protectant solution supplemented with 20% glycerol and flash frozen in liquid nitrogen.

The phosphocholine bound crystals were generated by soaking native crystals in well solution supplemented with 30 mM phosphocholine for 15 min at 21 °C. Crystals were collected as with Pt-derivatives.

Data collection. X-ray fluorescence scans identified zinc in crystals. SAD data were collected at beamline SER-CAT 22ID at the Advanced Photon Source (APS, USA). Diffraction from native crystals was collected at the Zn peak wavelength (1.2830 Å) and Pt-derivatized crystals at the Pt L-III edge (1.0721 Å). Phosphocholine-soaked data were collected on a Rigaku FR-E+ X-ray generator with a Cu anode and Saturn 944+ CCD detector. Zn and Pt SAD data were processed with HKL2000 (ref. 34), and phosphocholine data were processed with iMosflm and Scala^{35,36}. Diffraction of olipudase alfa was performed at the Zn peak wavelength (1.2825 Å) at beamline CMCF-08ID at the Canadian light source, and the SAD data were processed with XDS^{37,38}. Data collection statistics are listed in Table 1.

Structure determination. Pt-SAD data were used in *phenix.autosol* for Pt location search, phasing and density modification³⁹. Six Pt sites were found (Supplementary Fig. 4a), and the figure of merit for phasing was 0.251. The density-modified map was loaded into *buccaneer* in CCP4 for protein-chain building⁴⁰. Over 90% of the ASM residues were traced.

The holo ASM structure was completed using the higher-resolution Zn-SAD data. The output model from Pt-SAD was put into *phaser* MR-SAD⁴¹ against the Zn-SAD data for Zn searching and phasing. Two zinc atoms were located in each protein monomer (Supplementary Fig. 4b). The model was refined against anomalous amplitudes with *phenix.refine*⁴², and Hendrickson-Lattman coefficients from *phaser* MR-SAD were applied during refinement. Model building was completed with *Coot*⁴³. The sequence-to-structure register was guided by the glycosylation sites and disulfide bonds. Continuous electron density was observed from W84 to M611, and all main chain atoms are visible in the electron density.

The holo structure was used as a search model for MR phasing of phosphocholine data with *phaser*. One phosphocholine was built into the Fo-Fc difference map with *Coot* (Supplementary Fig. 4c), and refinement was completed with *phenix.refine*.

Structure determination and refinement of olipudase alfa were the same as described for holo ASM. Three protein monomers were found in each asymmetric unit. NCS restraints were applied, and high resolution ASM holoenzyme structure was used as reference structure during the refinement with *phenix.refine*. Two zinc atoms were located in each monomer according to the anomalous difference map (Supplementary Fig. 2b). Glycosylations on N175, N335, N503 and N520 were clearly visible in omit Fo-Fc maps (Supplementary Fig. 2a). Residues W84 to M611 were built into the electron density.

The final statistics for model building and refinement are listed in Table 1. Figures of structures are generated with *PyMOL*⁴⁴.

Docking sphingomyelin to ASM. To better visualize how the hydrocarbon chains might be placed on ASM upon substrate binding, a carbon-16 (C-16) sphingomyelin was docked to holo ASM with MOE 2014 (ref. 45). Amber10: EHT force field and R-field solvation were used. The ASM structure was prepared for 3D protonation with default parameters except that histidine, asparagine and glutamine side-chain conformations were retained. Sphingomyelin and ASM were energy minimized with default parameters. Phosphocholine interacting residues and zinc ions were selected as dock sites. In addition, I134, F138, and M142 in the saposin H3 helix were also included as dock sites based on the published saposin B and lipid crystal structure¹⁶. Phosphocholine was selected as ligand template, and template similarity placement protocol and induced fit protocol were applied in the docking procedure. Duplicated poses were removed and 30 energy favoured poses were retained after two rounds of rescoring. Other parameters were set as default. This docking is limited by the lack of computational restraints on two nearby zinc ions and numerous rotatable bonds in sphingomyelin. The wide open substrate binding cleft also contributed to the various conformations of hydrocarbon chains in docked results. The output 30 poses of sphingomyelin were visually inspected and validated based on whether the pose made reasonable interactions with ASM, instead of only looking for the best energy favored position. Three poses were selected and shown in Fig. 4.

Data availability. Coordinates and structure factors have been deposited in the Protein Data Bank under accession code 5I81 (holo-ASM), 5I85 (ASM with phosphocholine) and 5I8R (olipudase alfa).

References

- Schuchman, E. H. & Desnick, R. J. In *The Online Metabolic and Molecular Bases of Inherited Disease* (eds Beaudet, A. L. et al.) (The McGraw-Hill Companies, Inc., 2013).

- Besley, G. T., Hoogheboom, A. J., Hoogveen, A., Kleijer, W. J. & Galjaard, H. Somatic cell hybridisation studies showing different gene mutations in Niemann-Pick variants. *Hum. Genet.* 54, 409–412 (1980).
- McGovern, M. M. et al. Novel first-dose adverse drug reactions during a phase I trial of olipudase alfa (recombinant human acid sphingomyelinase) in adults with Niemann-Pick disease type B (acid sphingomyelinase deficiency). *Genet. Med.* 18, 34–40 (2016).
- Wasserstein, M. P. et al. Successful within-patient dose escalation of olipudase alfa in acid sphingomyelinase deficiency. *Mol. Genet. Metab.* 116, 88–97 (2015).
- Jenkins, R. W., Canals, D. & Hannun, Y. A. Roles and regulation of secretory and lysosomal acid sphingomyelinase. *Cell Signal.* 21, 836–846 (2009).
- 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).
- Paton, B. C., Schmid, B., Kustermann-Kuhn, B., Poulos, A. & Harzer, K. Additional biochemical findings in a patient and fetal sibling with a genetic defect in the sphingolipid activator protein (SAP) precursor, prosaposin. Evidence for a deficiency in SAP-1 and for a normal lysosomal neuraminidase. *Biochem J* 285(Pt 2): 481–488 (1992).
- Bradova, V. et al. Prosaposin deficiency: further characterization of the sphingolipid activator protein-deficient sibs. Multiple glycolipid elevations (including lactosylceramidosis), partial enzyme deficiencies and ultrastructure of the skin in this generalized sphingolipid storage disease. *Hum. Genet.* 92, 143–152 (1993).
- Schissel, S. L., Schuchman, E. H., Williams, K. J. & Tabas, I. Zn²⁺-stimulated sphingomyelinase is secreted by many cell types and is a product of the acid sphingomyelinase gene. *J. Biol. Chem.* 271, 18431–18436 (1996).
- Schissel, S. L., Kesler, G. A., Schuchman, E. H., Williams, K. J. & Tabas, I. The cellular trafficking and zinc dependence of secretory and lysosomal sphingomyelinase, two products of the acid sphingomyelinase gene. *J. Biol. Chem.* 273, 18250–18259 (1998).
- Callahan, J. W., Jones, C. S., Davidson, D. J. & Shankaran, P. The active site of lysosomal sphingomyelinase: evidence for the involvement of hydrophobic and ionic groups. *J. Neurosci. Res.* 10, 151–163 (1983).
- Seto, M. et al. A model of the acid sphingomyelinase phosphoesterase domain based on its remote structural homolog purple acid phosphatase. *Protein Sci.* 13, 3172–3186 (2004).
- Lim, S. M., Yeung, K., Tresaugues, L., Ling, T. H. & Nordlund, P. The structure and catalytic mechanism of Human Sphingomyelin Phosphodiesterase like 3a - an acid sphingomyelinase homolog with a novel nucleotide hydrolase activity. *FEBS J.* 1107–1123 (2016).
- Gorelik, A., Iles, K., Superti-Furga, G. & Nagar, B. Structural basis for nucleotide hydrolysis by the acid sphingomyelinase-like phosphodiesterase SMPD13A. *J. Biol. Chem.* 291, 6376–6385 (2016).
- Stanley, P., Schachter, H. & Taniguchi, N. In *Essentials of Glycobiology* (eds Varki, A., Cummings, R. D. & Esko, J. D.) (Cold Spring Harbor Laboratory Press, 2009).
- Ahn, V. E., Faull, K. F., Whiteleggs, J. P., Fluharty, A. L. & Prive, G. G. Crystal structure of saposin B reveals a dimeric shell for lipid binding. *Proc. Natl Acad. Sci. USA* 100, 38–43 (2003).
- Ahn, V. E., Leyko, P., Alattia, J. R., Chen, L. & Prive, G. G. Crystal structures of saposins A and C. *Protein Sci.* 15, 1849–1857 (2006).
- Popovic, K., Holyoake, J., Pomes, R. & Prive, G. G. Structure of saposin A lipoprotein discs. *Proc. Natl Acad. Sci. USA* 109, 2908–2912 (2012).
- Openshaw, A. E., Race, P. R., Monzo, H. J., Vazquez-Boland, J. A. & Banfield, M. J. Crystal structure of SmcL, a bacterial neutral sphingomyelinase C from *Listeria*. *J. Biol. Chem.* 280, 35011–35017 (2005).
- Rodriguez-Pascual, L. et al. Identification and characterization of SMPD1 mutations causing Niemann-Pick types A and B in Spanish patients. *Hum. Mutat.* 30, 1117–1122 (2009).
- Sikora, J., Pavlu-Pereira, H., Elleder, M., Roelofs, H. & Wevers, R. A. Seven novel acid sphingomyelinase gene mutations in Niemann-Pick type A and B patients. *Ann. Hum. Genet.* 67, 63–70 (2003).
- Guddat, L. W. et al. Crystal structure of mammalian purple acid phosphatase. *Structure* 7, 757–767 (1999).
- Barenholz, Y. & Thompson, T. E. Sphingomyelin: biophysical aspects. *Chem. Phys. Lipids* 102, 29–34 (1999).
- Frauenfeld, J. et al. A saposin-lipoprotein nanoparticle system for membrane proteins. *Nat. Methods* 13, 345–351 (2016).
- Onilka, V. O., Breiden, B., Babalola, J. O. & Sandhoff, K. Acid sphingomyelinase activity is regulated by membrane lipids and facilitates cholesterol transfer by NPC2. *J. Lipid Res.* 55, 2606–2619 (2014).
- Sribney, M. & Kennedy, E. P. The enzymatic synthesis of sphingomyelin. *J. Biol. Chem.* 233, 1315–1322 (1958).
- Qiu, H. et al. Activation of human acid sphingomyelinase through modification or deletion of C-terminal cysteine. *J. Biol. Chem.* 278, 32744–32752 (2003).

28. Rodriguez, F. *et al.* Crystal structure of the *Bacillus subtilis* phosphodiesterase PhoD reveals an iron and calcium-containing active site. *J. Biol. Chem.* 289, 30889–30899 (2014).
29. Beckmann, N., Sharma, D., Gulbins, E., Becker, K. A. & Edelmann, B. Inhibition of acid sphingomyelinase by tricyclic antidepressants and analogs. *Front. Physiol.* 5, 331 (2014).
30. Zeidan, Y. H. & Hannun, Y. A. Activation of acid sphingomyelinase by protein kinase C δ -mediated phosphorylation. *J. Biol. Chem.* 282, 11549–11561 (2007).
31. He, X. *et al.* Purification and characterization of recombinant, human acid ceramidase. Catalytic reactions and interactions with acid sphingomyelinase. *J. Biol. Chem.* 278, 32978–32986 (2003).
32. Dumitru, C. A. & Gulbins, E. TRAIL activates acid sphingomyelinase via a redox mechanism and releases ceramide to trigger apoptosis. *Oncogene* 25, 5612–5625 (2006).
33. Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl Acad. Sci. USA* 99, 13419–13424 (2002).
34. Otwinowski, Z. & Minor, W. in *Methods in Enzymology* Vol. 276, 307–326 (Academic Press, 1997).
35. Batty, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. IMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* 67, 271–281 (2011).
36. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* 67, 235–242 (2011).
37. Kabach, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66, 125–132 (2010).
38. Karpplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* 336, 1030–1033 (2012).
39. Terwilliger, T. C. *et al.* Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D Biol. Crystallogr.* 65, 582–601 (2009).
40. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* 62, 1002–1011 (2006).
41. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–674 (2007).
42. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenixrefine. *Acta Crystallogr. D Biol. Crystallogr.* 68, 352–367 (2012).
43. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501 (2010).
44. Schrödinger, L. L. C. The PyMOL Molecular Graphics System, Version 1.3r1 (2010).
45. Chemical Computing Group Inc. Molecular Operating Environment, Version 2014.09, www.chemcomp.com (2014).
46. Li, H., Robertson, A. D. & Jensen, J. H. Very fast empirical prediction and rationalization of protein pKa values. *Proteins* 61, 704–721 (2005).
47. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 32, W665–W667 (2004).

Acknowledgements

We thank Venkat Ryakala, Kevin Brower and Verna Warikoo for help with ASM purification, Bharat Datta and Scott Van Patten for supply of oligoputase alfa, and Ana Cristina Puga and Robert Mataliano for critical reading of the manuscript.

Author contributions

Y.-E.Z. designed research, performed experiments and wrote the manuscript. R.R.W. designed and supervised research and wrote the manuscript. H.Q. and T.E. supervised research and edited the manuscript. M.C.M. and S.C.G. performed crystal screening of oligoputase alfa. S.G. edited the manuscript.


Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: Y.-E.Z., T.E., H.Q. and R.R.W. are employees of Sanofi. The remaining authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Zhou, Y.-F. *et al.* Human acid sphingomyelinase structures provide insight to molecular basis of Niemann-Pick disease. *Nat. Commun.* 7, 13082 doi: 10.1038/ncomms13082 (2016).

 This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016

LIGAND-PROMOTED PROTEIN FOLDING BY BIASED KINETIC
PARTITIONING

I helped design the eukaryotic experiments in this study and performed these experiments. This study provides an experimentally validated model for the effect ligand binding has on protein folding. In addition to testing this model in a prokaryotic model, this model was validated for the application to a eukaryotic system. This is the first model created for an FDA-approved pharmacological chaperone in a disease model.



HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2017 August 20.

Published in final edited form as:

Nat Chem Biol. 2017 April ; 13(4): 369–371. doi:10.1038/nchembio.2303.

Ligand-Promoted Protein Folding by Biased Kinetic Partitioning

Karan S. Hingorani^{1,2}, Matthew C. Metcalf^{1,2}, Derrick T. Deming^{1,2}, Scott C. Garman^{1,2}, Evan T. Powers^{4,*}, and Lila M. Gierasch^{1,2,3,*}

¹Program in Molecular and Cellular Biology, University of Massachusetts Amherst, Amherst, MA 01003

²Department of Biochemistry & Molecular Biology, University of Massachusetts Amherst, Amherst, MA 01003

³Department of Chemistry, University of Massachusetts Amherst, Amherst, MA 01003

⁴Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037

Abstract

Protein folding in cells occurs in the presence of high concentrations of endogenous binding partners, and exogenous binding partners have been exploited as pharmacological chaperones. A combined mathematical modeling and experimental approach shows that a ligand improves the folding of a destabilized protein by biasing the kinetic partitioning between folding and alternative fates (aggregation or degradation). Computationally predicted inhibition of test protein aggregation and degradation as a function of ligand concentration are validated by experiments in two disparate cellular systems.

Proteins in a physiological milieu can have many binding partners. Surprisingly, little attention has been paid to how such ligands affect protein folding, despite the current interest in *in vivo* protein folding^{1–5}. In the intracellular environment, proteins must navigate an intersecting folding–aggregation landscape to reach their functional states⁶. In principle, native ligands as well as pharmacological chaperones—ligands added to cells as potential therapeutics⁷—can influence this landscape. In the present study, we explore how ligand binding affects folding in the cell by combining mathematical modeling with experiments.

Our model for *in vivo* protein folding in the presence of ligands is pictured in Fig. 1. In this model, protein is synthesized in an unfolded state (U), which can fold to the native state (N) with forward and reverse rate constants k_f and k_u . Alternatively, the unfolded state can be

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>

*To whom to address correspondence. epowers@scripps.edu and gierasch@biochem.umass.edu.

Author contributions: K. S. H., L. M. G., E. T. P. and S. C. G. conceived and designed the experiments. K. S. H. performed the experiments with dDHF. K. S. H., M. C. M., and D. T. D. performed the experiments with R301Q α -GAL. E. T. P. developed the mathematical model. All authors contributed to the experimental design, data analysis, and manuscript preparation.

Competing financial interests: The authors declare no competing financial interests.

Additional information: Any supplementary information, chemical compound information and source data are available in the online version of the paper.

degraded (Deg) or aggregate (A). Degradation is irreversible with a rate constant k_{deg} . We also treat aggregation as being irreversible with a rate constant k_{agg} , although intracellular aggregates can be disaggregated by chaperones or degraded by autophagy⁸. These and other effects of the cellular protein folding machinery are subsumed into the rate constants in our model. The native state can bind to a ligand (L) to form a complex (N:L) with association and dissociation rate constants k_a and k_d . Finally, the N or N:L states can be secreted with a rate constant k_{sec} . We assume that there is a large extracellular reservoir of ligand with a constant concentration and that the intracellular ($[L]_{in}$) and extracellular ($[L]_{out}$) ligand concentrations are proportional: $[L]_{in} = K_{io}[L]_{out}$. The model's rate equations and more detailed discussion of the model are given in Supplementary Results, Supplementary Note.

The fraction of protein that remains soluble (*i.e.*, is not degraded and does not aggregate) at a given time, or F_r , is experimentally measurable. F_r will evolve in two stages during a protein expression time course when a ligand is present. It rapidly reaches a first "pseudo-steady state" during which ligand binding is effectively irreversible ($k_a[N][L] \gg k_d[N:L]$). The system then gradually approaches a second, more robust pseudo-steady state as ligand dissociation becomes appreciable. In both of these stages, F_r is given by:

$$F_r = 1 - \frac{[A]_t + [Deg]_t}{[P_{tot}]_t} = 1 - \frac{1}{1 + B_1 \left(\frac{1}{1 + B_2 + B_3[L]_{out}} \right)} \quad (1)$$

where $[A]_t$, $[Deg]_t$, and $[P_{tot}]_t$ are the concentrations of aggregated, degraded, and total protein synthesized at time t , and $B_1 = k_f/(k_{agg} + k_{deg})$ and $B_2 = k_{sec}/k_u$ during both pseudo-steady states (see Supplementary Note). However, during the first pseudo-steady state $B_3 = k_a K_{io}/k_u$, whereas during the second $B_3 = k_a K_{io}/(k_u(1 + k_d/k_{sec}))$. When necessary, we will write the former as $B_{3,1}$ and the latter as $B_{3,2}$. The F_r vs. $[L]_{out}$ curve shifts to higher ligand concentrations during the second pseudo-steady state because $B_{3,2} < B_{3,1}$, indicating that maintaining a given protein level requires more ligand when ligand dissociation is non-negligible. The time needed to transition from the first to the second pseudo-steady state is discussed in the Supplementary Note, but we note that it is greater in systems that do not secrete protein than in systems that do.

Equation (1) predicts that F_r increases with $[L]_{out}$ until it reaches the following limit:

$$F_{r([L] \rightarrow \infty)} = F_{r,max} = 1 - \frac{1}{1 + B_1} = 1 - \frac{1}{1 + \frac{k_f}{k_{agg} + k_{deg}}} \quad (2)$$

Thus, at very high ligand concentrations, a protein's fate depends on folding kinetics (k_f) but not folding thermodynamics. In addition, the fact that $F_{r,max}$ is determined by the ratio $k_f/(k_{agg} + k_{deg})$ illustrates that analyzing F_r vs. $[L]$ curves using our model generally yields information on rate constant ratios. This circumstance arises because F_r measures the

partitioning of protein molecules among their possible fates at pseudo-steady states, which is a function of relative, not absolute, rates.

To test the predictions of equation (1), we first studied a mutant of the *E. coli* protein dihydrofolate reductase (dDHFR) expressed in *E. coli* in the presence of trimethoprim (TMP), a DHFR ligand⁹. The mutation in question was a -Gly-Gly- insertion between K106 and A107 in a surface loop (Fig. 2a). This mutation destabilizes DHFR and makes it aggregation-prone but leaves TMP binding largely unperturbed ($K_d = 37 \pm 6$ nM compared to 9-15 nM for wild type DHFR⁹; Supplementary Fig. 1). *In vitro*, dDHFR aggregation is efficiently, though not completely, inhibited by an equimolar concentration of TMP (Supplementary Fig. 2).

In the absence of TMP, dDHFR is mostly aggregated ($F_r = 0.15 \pm 0.02$) after expression in *E. coli* BL21 (DE3) cells for 1.5 hours at 37 °C (Fig. 2b; Supplementary Fig. 3a), unlike wild type DHFR, which does not aggregate when overexpressed¹⁰. Titration with TMP increases F_r , but dDHFR still partially aggregates even at [TMP] = 100 μ M (Fig. 2b; Supplementary Fig. 3a). This behavior is consistent with equations (1) and (2). In addition, equation (1) fits well to the data in Fig. 2b, arguing that our model captures the essential features of the system. Arresting dDHFR synthesis by adding chloramphenicol in the presence of excess TMP did not change F_r after 1 h, consistent with our assumption that disaggregation is negligible (Supplementary Fig. 4a).

Interpreting the fit of equation (1) to the data in Fig. 2b requires knowing whether the dDHFR expression system is in the first or second pseudo-steady state, or somewhere in between. Two experiments suggest that the system is in the first pseudo-steady state. First, arresting synthesis with chloramphenicol and shifting to medium lacking TMP led to only a modest change in F_r after 1 h, arguing that dDHFR did not rapidly re-equilibrate from the ligand-bound to the free state (Supplementary Fig. 4b). Second, measurements of F_r vs. ligand concentration at the 0.5 h and 1 h time points were similar to those at the 1.5 h time point (Supplementary Fig. 4c), consistent with ligand-bound dDHFR accumulating as a fixed proportion of the total amount of protein synthesized (which increased roughly linearly with time: [dDHFR] = 1.5 mM, 2.1 mM and 3.3 mM at 0.5, 1 and 1.5 h). Both findings are consistent with irreversible ligand binding being a reasonable assumption at the 1.5 h time point. See the Supplementary Note for further discussion.

Of the parameters from the fit of equation (1), $B_1 = k_f/(k_{agg} + k_{deg}) = k_f/k_{agg} = 2.9 \pm 0.8$ is of particular interest because it reveals that folding is about three times faster than aggregation (*E. coli* BL21 lack Lon protease, so $k_{deg} \sim 0$). According to equation (2), this ratio controls the maximum rescue of dDHFR by TMP. No matter how high the concentration of TMP is, some dDHFR will be diverted to aggregation before TMP binding can rescue it; for dDHFR, that portion is about 25% (Fig. 2b). See the Supplementary Note for further discussion of these points.

As a second test of our model, we chose human α -galactosidase A (α -GAL) (Fig. 2c) expressed in HEK-293T cells in the presence of 1-deoxygalactonojirimycin (DGJ; migalastat), a galactose mimic that binds in α -GAL's active site. Destabilized α -GAL

mutants can cause Fabry disease because they are often degraded by ER associated degradation (ERAD) before they can be exported from the ER¹¹, leading to a loss-of-function. DGJ increases the activity of α -GAL mutants despite being an inhibitor because it stabilizes α -GAL in the ER but is out-competed by natural α -GAL substrates in the lysosome^{7,11,12}. Note that α -GAL self-associates to form a dimer; in our model the effects of this process are subsumed into k_u (see Supplementary Note).

We tested the effect of DGJ on the intracellular levels of FLAG-tagged R301Q α -GAL (which causes a mild form of Fabry disease¹³) that had been transfected into HEK-293T cells and expressed for 24 h. To determine the total amount of α -GAL synthesized, we used the proteasome inhibitor lactacystin to inhibit ERAD in one sample (note that since total protein levels are of interest in this experiment, any α -GAL aggregation caused by proteasome inhibition would be immaterial). Since no α -GAL was detected in the media, F_T is the ratio of the protein level at a given DGJ concentration to that of the lactacystin sample.

As before, F_T for α -GAL increases with increasing [DGJ] until a plateau is reached. Also, equation (1) fits the F_T vs. [DGJ] data well (Fig. 2d; Supplementary Figure 3b). Since no R301Q α -GAL aggregation was observed ($k_{agg} = 0$), $B_1 = k_f/(k_{agg} + k_{deg}) = k_f/k_{deg}$. Thus, the best-fit value of $B_1 = 1.2 \pm 0.2$ indicates that folding and degradation are almost evenly poised. Furthermore, according to equation (2), the maximal rescue for α -GAL R301Q is ~55% of the protein. See the Supplementary Note for further discussion.

To further probe the role of ligand dissociation in the pharmacological chaperoning of α -GAL, we measured F_T as a function of the concentration of another, weaker binding ligand: α -homogalactonojirimycin (HGJ)¹⁴. Since the α -GAL expression system is likely in the second pseudo-steady state (see Supplementary Note), equation (1) predicts that decreases in k_a or increases in k_d —either of which could cause the lower affinity of HGJ for α -GAL—would shift the F_T vs. [HGJ] response curve to higher ligand concentrations. Consistent with this expectation, HGJ did not affect F_T over the experimentally accessible concentration range (Supplementary Fig. 5). We note, however, that this result could be partly due to HGJ being taken up into cells less efficiently than DGJ.

This study has implications for identifying good candidates for pharmacological chaperoning. Ligands will always increase productive folding because $dF_T/d[L]$ must be positive. However, equation (2) shows that the upper limit of F_T is dictated by the relative kinetics of folding vs. degradation and/or aggregation. Pharmacological chaperoning is therefore most effective for thermodynamically unstable but fast folding proteins. For example, proteins fused to destabilized constructs of the fast-folding protein FKBP12¹⁵ are efficiently rescued by the FKBP ligand Shield-1¹⁶. Slow folding proteins, however, benefit less from pharmacological chaperoning, as illustrated by the limited success of pharmacological chaperones as monotherapies for cystic fibrosis¹⁷. Cystic fibrosis is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR), a large slow-folding membrane protein. Many CFTR mutants partition almost exclusively to ERAD leading to loss of CFTR function¹⁸. Treating cells expressing a disease-associated CFTR mutant ($\Delta F508$ CFTR) with lumacaftor modestly increased the trafficking of CFTR to the cell surface (to 14% of wild type levels)¹⁹. Since lumacaftor likely operates as a

pharmacological chaperone for CFTR¹⁹, this “low ceiling” for its efficacy is an unavoidable consequence of CFTR’s slow folding, and it suggests that other strategies could yield better results. For example, stimulating an alternative secretion pathway—and thereby increasing k_{sec} —has shown some promise in a mouse model of $\Delta F508$ -CFTR²⁰.

Finally, this study suggests that, like chaperones²¹, ligands could be “evolutionary buffers” in that their ability to stabilize the native state may allow proteins to evolve via otherwise risky mutations²². An example of naturally occurring binding partners preventing aggregation *in vivo* is the antibody light and heavy chains, each offsetting the aggregation threat of the other²³. However, this protective effect should not extend to mutations that slow down folding. Such mutations could be devastating if the protein’s folding rate was already marginal because it would invariably diminish F_r , perhaps leading to a loss-of-function phenotype. Based on this reasoning, one would expect residues in a slow-folding protein’s folding nucleus to be strongly conserved.

Online Methods

Cloning, expression, and purification of dihydrofolate reductase

The wild type cysteine-free dihydrofolate reductase (DHFR) gene was a gift from the Matthews lab (University of Massachusetts Medical School, Worcester MA)²⁶. It was subcloned into the pET28 expression vector using NdeI and BamHI cut sites. The -Gly-Gly- insertion between residues K106 and A107 was introduced by site-directed mutagenesis using a QuikChange protocol (Stratagene) yielding the “dDHFR” mutant. The vector was transformed into *E. coli* BL21 (DE3) cells. Cultures were grown at 37 °C to optical density at 600 nm (OD_{600}) of 0.6, induced with 1 mM final IPTG concentration and expressed at 37 °C for 4 h. The cells were harvested by centrifugation and re-suspended in buffer containing 20 mM HEPES, pH 7.4. The cells were lysed using a microfluidizer, and the protein was purified by refolding from inclusion bodies. Inclusion bodies were dissolved in 8 M urea and subsequently dialyzed to refold the protein, which was further purified using anion exchange chromatography on a DEAE column using a 0 M to 1 M sodium chloride gradient. After elution, the salt was removed by dialysis and the protein was flash frozen and stored at -80 °C until further use.

In vitro trimethoprim binding assay for dDHFR

The *in vitro* trimethoprim (TMP) binding assay was performed with 50 nM purified dDHFR following the protocol from Watson and co-workers²⁷. Briefly, dDHFR was incubated at 25 °C in phosphate buffer with increasing concentrations of TMP and the degree of quenching of dDHFR tryptophan fluorescence was monitored at 345 nm. Spectra were corrected for any contribution to the fluorescence by TMP by running a parallel titration without protein. Fraction bound (F_{bound}) was calculated by using the formula below.

$$F_{\text{bound}} = \frac{\text{fluorescence}_{\text{max}} - \text{fluorescence}_{\text{measured}}}{\text{fluorescence}_{\text{max}} - \text{fluorescence}_{\text{min}}}$$

The F_{bound} vs. TMP concentration data were fit to a single site binding model by using non-linear regression as implemented by the “NonlinearModelFit” command in Mathematica 10.4.

Cell growth, protein induction and partitioning for dDHFR expression experiments

For all experiments involving partitioning of dDHFR, *E. coli* BL21(DE3) cells were grown in LB medium until mid-log phase ($OD_{600} = 0.6$). Cells were induced with 1 mM final IPTG concentration for 0.5, 1, or 1.5 h. All growth and induction was performed at 37 °C. The cells were harvested after equalization for OD_{600} by centrifugation at $4,000 \times g$ for 2 minutes, the media was discarded, and the cells were lysed using BPER-II reagent at room temperature for 10 minutes with $1\mu\text{g/ml}$ DNase I in the lysis mixture. The sample was then centrifuged at $18,000 \times g$ for 30 minutes to separate insoluble (pellet) and soluble (supernatant) components. The samples were subsequently prepared by boiling in SDS running buffer for 12% SDS-PAGE.

TMP titration in *E. coli* cultures

TMP was dissolved in 100% methanol and then diluted to the desired concentration in media immediately after the addition of IPTG. Final methanol concentration did not exceed 7% v/v at the end of the experiment. Cells were harvested and partitioned as mentioned above.

Chloramphenicol shut off experiments

Protein was induced for 90 minutes in the presence of $80\mu\text{M}$ TMP. A fraction of the cells was harvested before any chloramphenicol was added and a second sample was collected after an hour of incubation with chloramphenicol at a final concentration of $50\mu\text{g/ml}$. The partitioning measurements were performed as described above. For the change of media experiment, a fraction of cells was harvested after the 90 minute induction of DHFR in the presence of $80\mu\text{M}$ TMP. The remaining cells were centrifuged for 10 minutes at $4000 \times g$, washed with equal volume LB centrifuged again and resuspended in LB supplemented with chloramphenicol at $50\mu\text{g/ml}$ without any TMP. The second sample was collected after an hour of incubation with chloramphenicol and processed as described above.

In vitro aggregation of dDHFR in the presence and absence of TMP

A solution of dDHFR ($20\mu\text{M}$) was prepared in pH 7.0 buffer and divided into two aliquots. To one aliquot was added TMP ($20\mu\text{M}$) from a stock solution in methanol (the final methanol concentration was 5% v/v). After incubating the aliquots for 2 days with agitation the turbidities of the samples were measured as optical density at 400 nm.

Mammalian cell growth, and experiments with α -galactosidase

We seeded 10 cm culture dishes with HEK 293T cells and allowed them to reach ~80% confluency at 37 °C, 5% CO_2 . We then transfected the cells with the R301Q α -galactosidase pCMV 3xFLAG-14 (R301Q α -GAL) vector using Lipofectamine 2000. At 24 hours post-transfection we trypsinized the cells, and used the resulting suspension to seed 6-well plates for the DGJ titration experiment. The cells were allowed to adhere over a 4-hour incubation period, after which desired amounts of DGJ, HGJ, or lactacystin ($20\mu\text{M}$ final concentration)

were added. After 24 hours the cells were washed thoroughly with PBS, then lysed with ice-cold lysis buffer (1% Triton X-100 in PBS). The lysate was centrifuged at $18,000 \times g$ for 30 minutes and the soluble fraction was used for analysis. The samples were normalized prior to loading using the BCA protein quantification method to ensure that equal amounts of protein were loaded in each case. The western blot was performed using as primary a polyclonal rabbit antibody against human α -galactosidase (GeneTex catalog number GTX 101178).

Gel band quantification

SDS-PAGE gels for dDHFR were stained with Coomassie G-250, scanned and analyzed on the LI-COR ODYSSEY CLx and quantified using the associated image studio software (version 4.0)²⁸. Western blots for the α -GAL experiments were similarly analyzed.

Fits of equation (1) to F_r vs. ligand concentration data

Equation (1) was fit to the F_r vs. ligand concentration data in Figures 2b and 2d by using non-linear regression as implemented by the “NonlinearModelFit” command in Mathematica 10.4. The uncorrected R^2 values are calculated from fits of equation (1) to all of the individual F_r measurements. The R^2 values that are corrected for measurement error are calculated from fits of equation (1) to the F_r measurements after averaging the triplicates. The latter procedure filters out measurement error, but yields exactly the same parameter estimates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank C. Robert Matthews (University of Massachusetts Medical School, Worcester) for providing us with a plasmid containing the wild type, cysteine-free *E. coli* DHFR gene. This work was supported by NIH grants GM101644 to L. M. G. and E. T. P. and DK76877 to S. C. G.

References

1. Hingorani KS, Gierasch LM. Comparing protein folding in vitro and in vivo: foldability meets the fitness challenge. *Curr Opin Struct Biol.* 2014; 24:81–90. [PubMed: 24434632]
2. Monteith WB, Cohen RD, Smith AE, Guzman-Cisneros E, Pielak GJ. Quinary structure modulates protein stability in cells. *Proc Natl Acad Sci U S A.* 2015; 112:1739–42. [PubMed: 25624496]
3. Guzman I, Gruebele M. Protein folding dynamics in the cell. *J Phys Chem B.* 2014; 118:8459–70. [PubMed: 24878167]
4. Powers ET, Powers DL, Gierasch LM. FoldEco: a model for proteostasis in *E. coli*. *Cell Rep.* 2012; 1:265–76. [PubMed: 22509487]
5. Bershtein S, Choi JM, Bhattacharyya S, Budnik B, Shakhnovich E. Systems-level response to point mutations in a core metabolic enzyme modulates genotype-phenotype relationship. *Cell Rep.* 2015; 11:645–56. [PubMed: 25892240]
6. Jahn TR, Radford SE. Folding versus aggregation: polypeptide conformations on competing pathways. *Arch Biochem Biophys.* 2008; 469:100–17. [PubMed: 17588526]

7. Parenti G, Andria G, Valenzano KJ. Pharmacological chaperone therapy: preclinical development, clinical translation, and prospects for the treatment of lysosomal storage disorders. *Mol Ther.* 2015; 23:1138–48. [PubMed: 25881001]
8. Sontag EM, Vonk WI, Frydman J. Sorting out the trash: the spatial nature of eukaryotic protein quality control. *Curr Opin Cell Biol.* 2014; 26:139–46. [PubMed: 24463332]
9. Baccanari DP, Daluge S, King RW. Inhibition of dihydrofolate reductase: effect of reduced nicotinamide adenine dinucleotide phosphate on the selectivity and affinity of diaminobenzylpyrimidines. *Biochemistry.* 1982; 21:5068–75. [PubMed: 6814484]
10. Cho Y, et al. Individual and collective contributions of chaperoning and degradation to protein homeostasis in *E. coli*. *Cell Rep.* 2015; 11:321–33. [PubMed: 25843722]
11. Ishii S, et al. Mutant alpha-galactosidase A enzymes identified in Fabry disease patients with residual enzyme activity: biochemical characterization and restoration of normal intracellular processing by 1-deoxygalactonojirimycin. *Biochem J.* 2007; 406:285–95. [PubMed: 17555407]
12. Fan JQ, Ishii S, Asano N, Suzuki Y. Accelerated transport and maturation of lysosomal alpha-galactosidase A in Fabry lymphoblasts by an enzyme inhibitor. *Nat Med.* 1999; 5:112–5. [PubMed: 9883849]
13. Sakuraba H, et al. Identification of point mutations in the alpha-galactosidase A gene in classical and atypical hemizygotes with Fabry disease. *Am J Hum Genet.* 1990; 47:784–9. [PubMed: 2171331]
14. Asano N, et al. In vitro inhibition and intracellular enhancement of lysosomal alpha-galactosidase A activity in Fabry lymphoblasts by 1-deoxygalactonojirimycin and its derivatives. *Eur J Biochem.* 2000; 267:4179–86. [PubMed: 10866822]
15. Main ER, Fulton KF, Jackson SE. Folding pathway of FKBP12 and characterisation of the transition state. *J Mol Biol.* 1999; 291:429–44. [PubMed: 10438630]
16. Banaszynski LA, Chen LC, Maynard-Smith LA, Ooi AG, Wandless TJ. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell.* 2006; 126:995–1004. [PubMed: 16959577]
17. Clancy JP, et al. Results of a phase IIa study of VX-809, an investigational CFTR corrector compound, in subjects with cystic fibrosis homozygous for the F508del-CFTR mutation. *Thorax.* 2012; 67:12–8. [PubMed: 21825083]
18. Lukacs GL, et al. Conformational maturation of CFTR but not its mutant counterpart (delta F508) occurs in the endoplasmic reticulum and requires ATP. *EMBO J.* 1994; 13:6076–86. [PubMed: 7529176]
19. Van Goor F, et al. Correction of the F508del-CFTR protein processing defect in vitro by the investigational drug VX-809. *Proc Natl Acad Sci U S A.* 2011; 108:18843–8. [PubMed: 21976485]
20. Gee HY, Noh SH, Tang BL, Kim KH, Lee MG. Rescue of DeltaF508-CFTR trafficking via a GRASP-dependent unconventional secretion pathway. *Cell.* 2011; 146:746–60. [PubMed: 21884936]
21. Tokuriki N, Tawfik DS. Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature.* 2009; 459:668–73. [PubMed: 19494908]
22. Gershenson A, Gierasch LM, Pastore A, Radford SE. Energy landscapes of functional proteins are inherently risky. *Nat Chem Biol.* 2014; 10:884–91. [PubMed: 25325699]
23. Corcos D, et al. Immunoglobulin aggregation leading to Russell body formation is prevented by the antibody light chain. *Blood.* 2010; 115:282–8. [PubMed: 19822901]
24. Bystroff C, Oatley SJ, Kraut J. Crystal structures of *Escherichia coli* dihydrofolate reductase: the NADP⁺ holoenzyme and the folate.NADP⁺ ternary complex. Substrate binding and a model for the transition state. *Biochemistry.* 1990; 29:3263–77. [PubMed: 2185835]
25. Guce AI, Clark NE, Rogich JJ, Garman SC. The molecular basis of pharmacological chaperoning in human alpha-galactosidase. *Chem Biol.* 2011; 18:1521–6. [PubMed: 22195554]

References for Online Methods

26. Ionescu RM, Smith VF, O'Neill JC Jr, Matthews CR. Multistate equilibrium unfolding of *Escherichia coli* dihydrofolate reductase: thermodynamic and spectroscopic description of the native, intermediate, and unfolded ensembles. *Biochemistry*. 2000; 39:9540–50. [PubMed: 10924151]
27. Watson M, Liu JW, Ollis D. Directed evolution of trimethoprim resistance in *Escherichia coli*. *FEBS J*. 2007; 274:2661–71. [PubMed: 17451440]
28. Luo S, Wehr NB, Levine RL. Quantitation of protein on gels and blots by infrared fluorescence of Coomassie blue and Fast Green. *Anal Biochem*. 2006; 350:233–8. [PubMed: 16336940]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

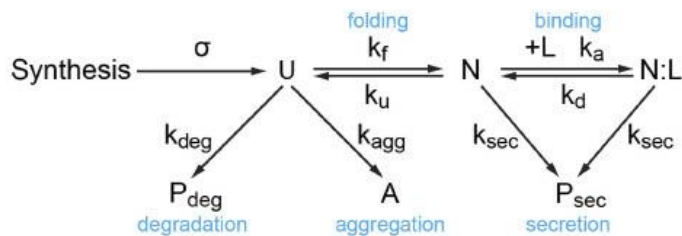


Figure 1. A model for the partitioning of protein among folding, aggregation, and degradation pathways

The species in the model are: U, unfolded protein; N, natively folded protein; L, unbound ligand; N:L, ligand bound natively folded protein; Deg, degraded protein; A, aggregated protein; Sec, secreted protein. The total protein synthesized, P_{tot} , includes all of these states. The rate constants are: σ , protein synthesis rate ($\mu\text{M s}^{-1}$); k_f , folding rate constant (s^{-1}); k_u , unfolding rate constant (s^{-1}); k_a , protein–ligand association rate constant ($\mu\text{M}^{-1} \text{s}^{-1}$); k_d , protein–ligand dissociation rate constant (s^{-1}); k_{deg} , degradation rate constant (s^{-1}); k_{agg} , aggregation rate constant (s^{-1}); k_{sec} , secretion rate constant (s^{-1}).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

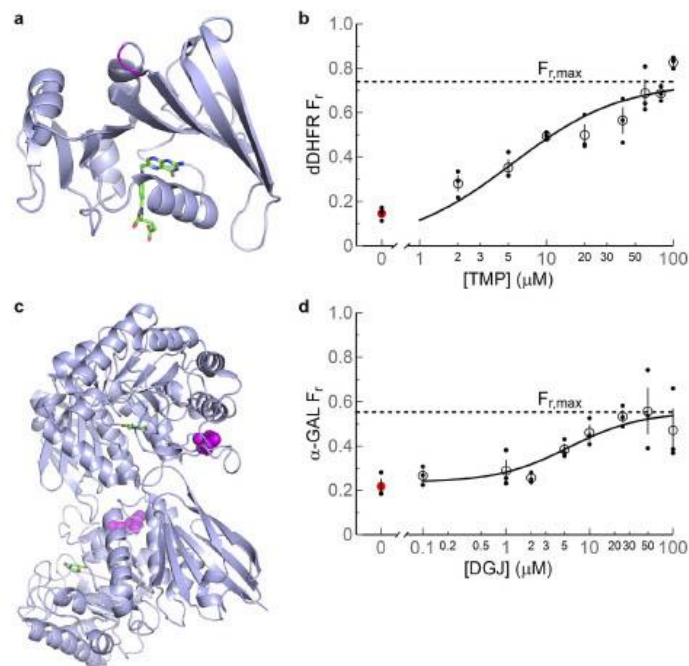


Figure 2. The proteins studied and plots of fraction soluble protein remaining (F_r) vs. ligand concentration
(a) Wild type *E. coli* DHFR with bound folate (colored by atom) (PDB ID: 7DFR²⁴). The site of the -Gly-Gly- insertion in the dDHFR mutant is in magenta. **(b)** F_r vs. [TMP] (expressed for 1.5 h at 37 °C). The fit of equation (1) to the data is shown (solid curve) with $B_1(k_f/(k_{agg} + k_{deg})) = 2.9 \pm 0.8$, $B_{3,1} = k_a K_{io}/k_u = 0.048 \pm 0.025 \mu\text{M}^{-1}$, and $R^2 = 0.76$ (0.84 after correction for measurement error). **(c)** Dimeric wild type human α -GAL with bound DGJ (colored by atom) (PDB ID: 3S5Y²⁵). The R301Q mutation site is in magenta. **(d)** F_r vs. [DGJ] (expressed for 24 h at 37 °C). The fit of equation (1) to the data is shown (solid curve) with $B_1 = 1.2 \pm 0.2$, $B_2 (k_{sec}/k_u) = 0.33 \pm 0.11$, $B_{3,2} = k_a K_{io}/(k_u(1+k_d/k_{sec})) = 0.14 \pm 0.10 \mu\text{M}^{-1}$, and $R^2 = 0.59$ (0.88 after correction for measurement error). In **(b)** and **(d)**, data from triplicate measurements are shown as smaller filled circles; the dashed line represents the maximum value of F_r according to equation (2); the value of F_r at [ligand] = 0 is shown as a filled red circle; means are shown as open circles; and error bars represent the standard error of the mean. The consistency of observed standard errors argues for similar underlying variability. Sample sizes were chosen based on errors observed in previous studies¹⁰.

BIBLIOGRAPHY

1. Spada, M., et al., *High incidence of later-onset fabry disease revealed by newborn screening*. Am J Hum Genet, 2006. **79**(1): p. 31-40.
2. Bishop, D.F., et al., *Human alpha-galactosidase A: nucleotide sequence of a cDNA clone encoding the mature enzyme*. Proc Natl Acad Sci U S A, 1986. **83**(13): p. 4859-63.
3. Kint, J.A., *Fabry's disease: alpha-galactosidase deficiency*. Science, 1970. **167**(3922): p. 1268-9.
4. Desnick, R.I., YA.; Eng, CM., *a-Galactosidase A deficiency: Fabry disease*, in *The metabolic & molecular bases of inherited disease*, C.R. Scriver, Editor. 2001, McGraw-Hill: New York. p. 3733-3774.
5. Mayes, J.S., et al., *Differential assay for lysosomal alpha-galactosidases in human tissues and its application to Fabry's disease*. Clin Chim Acta, 1981. **112**(2): p. 247-51.
6. Ries, M., et al., *Pediatric Fabry disease*. Pediatrics, 2005. **115**(3): p. e344-55.
7. Lee, K., et al., *A biochemical and pharmacological comparison of enzyme replacement therapies for the glycolipid storage disorder Fabry disease*. Glycobiology, 2003. **13**(4): p. 305-13.
8. Schiffmann, R., et al., *Enzyme replacement therapy in Fabry disease: a randomized controlled trial*. JAMA, 2001. **285**(21): p. 2743-9.
9. Fan, J.Q., *A contradictory treatment for lysosomal storage disorders: inhibitors enhance mutant enzyme activity*. Trends Pharmacol Sci, 2003. **24**(7): p. 355-60.

10. Guce, A.I., et al., *The molecular basis of pharmacological chaperoning in human alpha-galactosidase*. Chem Biol, 2011. **18**(12): p. 1521-6.
11. Moran, N., *FDA approves Galafold, a triumph for Amicus*. Nat Biotechnol, 2018. **36**(10): p. 913.
12. Bodensteiner, D., et al., *Successful reinstatement of agalsidase beta therapy in Fabry disease patients with previous IgE-antibody or skin-test reactivity to the recombinant enzyme*. Genet Med, 2008. **10**(5): p. 353-8.
13. Eng, C.M., et al., *Safety and efficacy of recombinant human alpha-galactosidase A--replacement therapy in Fabry's disease*. N Engl J Med, 2001. **345**(1): p. 9-16.
14. Schiffmann, R., et al., *Long-term therapy with agalsidase alfa for Fabry disease: safety and effects on renal function in a home infusion setting*. Nephrol Dial Transplant, 2006. **21**(2): p. 345-54.
15. Thurberg, B.L., et al., *Globotriaosylceramide accumulation in the Fabry kidney is cleared from multiple cell types after enzyme replacement therapy*. Kidney Int, 2002. **62**(6): p. 1933-46.
16. Tondel, C., et al., *Foot process effacement is an early marker of nephropathy in young classic Fabry patients without albuminuria*. Nephron, 2015. **129**(1): p. 16-21.
17. Bainton, D.F., *The discovery of lysosomes*. J Cell Biol, 1981. **91**(3 Pt 2): p. 66s-76s.
18. Doray, B., et al., *Interaction of the cation-dependent mannose 6-phosphate receptor with GGA proteins*. J Biol Chem, 2002. **277**(21): p. 18477-82.

19. Klumperman, J., et al., *Mannose 6-phosphate receptors are sorted from immature secretory granules via adaptor protein AP-1, clathrin, and syntaxin 6-positive vesicles*. J Cell Biol, 1998. **141**(2): p. 359-71.
20. Attar, N. and P.J. Cullen, *The retromer complex*. Adv Enzyme Regul, 2010. **50**(1): p. 216-36.
21. Blanz, J., et al., *Mannose 6-phosphate-independent Lysosomal Sorting of LIMP-2*. Traffic, 2015. **16**(10): p. 1127-36.
22. Bellettato, C.M. and M. Scarpa, *Pathophysiology of neuropathic lysosomal storage disorders*. J Inherit Metab Dis, 2010. **33**(4): p. 347-62.
23. Kudo, M., M.S. Brem, and W.M. Canfield, *Mucopolidosis II (I-cell disease) and mucopolidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha / beta -subunits precursor gene*. Am J Hum Genet, 2006. **78**(3): p. 451-63.
24. Mazrier, H., et al., *Inheritance, biochemical abnormalities, and clinical features of feline mucopolidosis II: the first animal model of human I-cell disease*. J Hered, 2003. **94**(5): p. 363-73.
25. Wang, P., et al., *A GNPTAB nonsense variant is associated with feline mucopolidosis II (I-cell disease)*. BMC Vet Res, 2018. **14**(1): p. 416.
26. Berkovic, S.F., et al., *Array-based gene discovery with three unrelated subjects shows SCARB2/LIMP-2 deficiency causes myoclonus epilepsy and glomerulosclerosis*. Am J Hum Genet, 2008. **82**(3): p. 673-84.

27. Reczek, D., et al., *LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase*. *Cell*, 2007. **131**(4): p. 770-83.
28. Blanz, J., et al., *Disease-causing mutations within the lysosomal integral membrane protein type 2 (LIMP-2) reveal the nature of binding to its ligand beta-glucocerebrosidase*. *Hum Mol Genet*, 2010. **19**(4): p. 563-72.
29. Ferrans, V.J., R.G. Hibbs, and C.D. Burda, *The heart in Fabry's disease. A histochemical and electron microscopic study*. *Am J Cardiol*, 1969. **24**(1): p. 95-110.
30. von Scheidt, W., et al., *An atypical variant of Fabry's disease with manifestations confined to the myocardium*. *N Engl J Med*, 1991. **324**(6): p. 395-9.
31. Beck, M., et al., *Anderson-Fabry disease in children and adolescents*. *Contrib Nephrol*, 2001(136): p. 251-5.
32. Waldek, S., et al., *Life expectancy and cause of death in males and females with Fabry disease: findings from the Fabry Registry*. *Genet Med*, 2009. **11**(11): p. 790-6.
33. Nakao, S., et al., *An atypical variant of Fabry's disease in men with left ventricular hypertrophy*. *N Engl J Med*, 1995. **333**(5): p. 288-93.
34. Sachdev, B., et al., *Prevalence of Anderson-Fabry disease in male patients with late onset hypertrophic cardiomyopathy*. *Circulation*, 2002. **105**(12): p. 1407-11.

35. Nakao, S., et al., *Fabry disease: detection of undiagnosed hemodialysis patients and identification of a "renal variant" phenotype*. *Kidney Int*, 2003. **64**(3): p. 801-7.
36. Kotanko, P., et al., *Results of a nationwide screening for Anderson-Fabry disease among dialysis patients*. *J Am Soc Nephrol*, 2004. **15**(5): p. 1323-9.
37. Echevarria, L., et al., *X-chromosome inactivation in female patients with Fabry disease*. *Clin Genet*, 2016. **89**(1): p. 44-54.
38. Lyon, M.F., *Gene action in the X-chromosome of the mouse (Mus musculus L.)*. *Nature*, 1961. **190**: p. 372-3.
39. Sagie, S. and E. Monovich, *[Mary Lyon (1925-2014) and the Random Inactivation of Chromosome X]*. *Harefuah*, 2016. **155**(3): p. 140-4, 197.
40. Garman, S.C. and D.N. Garboczi, *The molecular defect leading to Fabry disease: structure of human alpha-galactosidase*. *J Mol Biol*, 2004. **337**(2): p. 319-35.
41. Guce, A.I., et al., *Catalytic mechanism of human alpha-galactosidase*. *J Biol Chem*, 2010. **285**(6): p. 3625-32.
42. Radin, N.S., *Treating glucosphingolipid disorders by chemotherapy: use of approved drugs and over-the-counter remedies*. *J Inher Metab Dis*, 2000. **23**(8): p. 767-77.
43. Asano, N., et al., *In vitro inhibition and intracellular enhancement of lysosomal alpha-galactosidase A activity in Fabry lymphoblasts by 1-deoxygalactonojirimycin and its derivatives*. *Eur J Biochem*, 2000. **267**(13): p. 4179-86.
44. Ozawa, K., *[Gene therapy using AAV]*. *Uirusu*, 2007. **57**(1): p. 47-55.

45. Mynarek, M., et al., *Allogeneic hematopoietic SCT for alpha-mannosidosis: an analysis of 17 patients*. Bone Marrow Transplant, 2012. **47**(3): p. 352-9.
46. Platt, F.M., et al., *N-butyldeoxynojirimycin is a novel inhibitor of glycolipid biosynthesis*. J Biol Chem, 1994. **269**(11): p. 8362-5.
47. Abe, A., et al., *Glycosphingolipid depletion in fabry disease lymphoblasts with potent inhibitors of glucosylceramide synthase*. Kidney Int, 2000. **57**(2): p. 446-54.
48. Abe, A., et al., *Reduction of globotriaosylceramide in Fabry disease mice by substrate deprivation*. J Clin Invest, 2000. **105**(11): p. 1563-71.
49. Banikazemi, M., et al., *Agalsidase-beta therapy for advanced Fabry disease: a randomized trial*. Ann Intern Med, 2007. **146**(2): p. 77-86.
50. Sakuraba, H., et al., *Comparison of the effects of agalsidase alfa and agalsidase beta on cultured human Fabry fibroblasts and Fabry mice*. J Hum Genet, 2006. **51**(3): p. 180-188.
51. Choi, S.P., T.H., *Enhancement of sialyltransferase-catalyzed transfer of sialic acid onto glycoprotein oligosaccharides using silkworm hemolymph and its 30K protein*. J Mol. Catal B: Enz, 2006. **43**(1-4): p. 128-132.
52. Zhou, Q., et al., *Strategies for Neoglycan conjugation to human acid alpha-glucosidase*. Bioconjug Chem, 2011. **22**(4): p. 741-51.
53. Zhu, Y., et al., *Glycoengineered acid alpha-glucosidase with improved efficacy at correcting the metabolic aberrations and motor function deficits in a mouse model of Pompe disease*. Mol Ther, 2009. **17**(6): p. 954-63.

54. Zhu, Y., et al., *Conjugation of mannose 6-phosphate-containing oligosaccharides to acid alpha-glucosidase improves the clearance of glycogen in pompe mice*. J Biol Chem, 2004. **279**(48): p. 50336-41.
55. Zhu, Y., et al., *Carbohydrate-remodelled acid alpha-glucosidase with higher affinity for the cation-independent mannose 6-phosphate receptor demonstrates improved delivery to muscles of Pompe mice*. Biochem J, 2005. **389**(Pt 3): p. 619-28.
56. Grabowski, G.A., M. Golembo, and Y. Shaaltiel, *Taliglucerase alfa: an enzyme replacement therapy using plant cell expression technology*. Mol Genet Metab, 2014. **112**(1): p. 1-8.
57. Schiffmann, R., et al., *Pegunigalsidase alfa, a novel PEGylated enzyme replacement therapy for Fabry disease, provides sustained plasma concentrations and favorable pharmacodynamics: A 1-year Phase 1/2 clinical trial*. J Inherit Metab Dis, 2019. **42**(3): p. 534-544.
58. Frustaci, A., et al., *Improvement in cardiac function in the cardiac variant of Fabry's disease with galactose-infusion therapy*. N Engl J Med, 2001. **345**(1): p. 25-32.
59. Fan, J.Q., et al., *Accelerated transport and maturation of lysosomal alpha-galactosidase A in Fabry lymphoblasts by an enzyme inhibitor*. Nat Med, 1999. **5**(1): p. 112-5.
60. Asano, N., et al., *Nitrogen-in-the-ring pyranoses and furanoses: structural basis of inhibition of mammalian glycosidases*. J Med Chem, 1994. **37**(22): p. 3701-6.

61. Wang, A.M., D.F. Bishop, and R.J. Desnick, *Human alpha-N-acetylgalactosaminidase-molecular cloning, nucleotide sequence, and expression of a full-length cDNA. Homology with human alpha-galactosidase A suggests evolution from a common ancestral gene.* J Biol Chem, 1990. **265**(35): p. 21859-66.
62. Cantarel, B.L., et al., *The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics.* Nucleic Acids Res, 2009. **37**(Database issue): p. D233-8.
63. Clark, N.E. and S.C. Garman, *The 1.9 Å structure of human alpha-N-acetylgalactosaminidase: The molecular basis of Schindler and Kanzaki diseases.* J Mol Biol, 2009. **393**(2): p. 435-47.
64. Andreotti, G., et al., *Prediction of the responsiveness to pharmacological chaperones: lysosomal human alpha-galactosidase, a case of study.* Orphanet J Rare Dis. **5**: p. 36.
65. Fujimoto, Z., et al., *Crystal structure of rice alpha-galactosidase complexed with D-galactose.* J Biol Chem, 2003. **278**(22): p. 20313-8.
66. Garman, S.C., et al., *The 1.9 Å structure of alpha-N-acetylgalactosaminidase: molecular basis of glycosidase deficiency diseases.* Structure, 2002. **10**(3): p. 425-34.
67. C., G.S., *Structural studies on α-GAL and α-NAGAL: The atomic basis of Fabry and Schindler diseases.* Biocatalysis and Biotransformation, 2006. **24**:1-2: p. 129-136.

68. Tajima, Y., et al., *Use of a modified alpha-N-acetylgalactosaminidase in the development of enzyme replacement therapy for Fabry disease*. *Am J Hum Genet*, 2009. **85**(5): p. 569-80.
69. Benichou, B., et al., *A retrospective analysis of the potential impact of IgG antibodies to agalsidase beta on efficacy during enzyme replacement therapy for Fabry disease*. *Mol Genet Metab*, 2009. **96**(1): p. 4-12.
70. Vedder, A.C., et al., *Treatment of Fabry disease with different dosing regimens of agalsidase: effects on antibody formation and GL-3*. *Mol Genet Metab*, 2008. **94**(3): p. 319-25.
71. Starzyk, K., et al., *The long-term international safety experience of imiglucerase therapy for Gaucher disease*. *Mol Genet Metab*, 2007. **90**(2): p. 157-63.
72. Golubev, A.M., et al., *Crystal structure of alpha-galactosidase from *Trichoderma reesei* and its complex with galactose: implications for catalytic mechanism*. *J Mol Biol*, 2004. **339**(2): p. 413-22.
73. Perona, J.J., et al., *Structural origins of substrate discrimination in trypsin and chymotrypsin*. *Biochemistry*, 1995. **34**(5): p. 1489-99.
74. Toscano, M.D., K.J. Woycechowsky, and D. Hilvert, *Minimalist active-site redesign: teaching old enzymes new tricks*. *Angew Chem Int Ed Engl*, 2007. **46**(18): p. 3212-36.
75. Minor W, O.Z., in *Methods in Enzymology, Macromolecular Crystallography*, S.R.M. Carter CW, Editor. 1997, Academic Press: New York. p. 307-326.
76. Collaborative Computational Project, N., *The CCP4 suite: programs for protein crystallography*. *Acta Crystallogr D Biol Crystallogr*, 1994. **50**(Pt 5): p. 760-3.

77. Jones, T.A., et al., *Improved methods for building protein models in electron density maps and the location of errors in these models*. Acta Crystallogr A, 1991. **47 (Pt 2)**: p. 110-9.
78. Laskowski, R.A., et al., *AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR*. J Biomol NMR, 1996. **8(4)**: p. 477-86.
79. Kleywegt, G.J. and R.J. Read, *Not your average density*. Structure, 1997. **5(12)**: p. 1557-69.
80. P.J., K., *MOLSCRIPT : a program to produce both detailed and schematic plots of protein structures*. J. Appl. Cryst., 1991. **24**: p. 946-950.
81. Fenn, T.D., Ringe, D. Petsko, G.A., *POVScript+ : a program for model and data visualization using persistence of vision ray-tracing*. J. Appl. Cryst., 2003. **36**: p. 944-947.
82. Meikle, P.J., et al., *Prevalence of lysosomal storage disorders*. JAMA, 1999. **281(3)**: p. 249-54.
83. Eng, C.M., et al., *A phase 1/2 clinical trial of enzyme replacement in fabry disease: pharmacokinetic, substrate clearance, and safety studies*. Am J Hum Genet, 2001. **68(3)**: p. 711-22.
84. Rennke, H.G., R.S. Cotran, and M.A. Venkatachalam, *Role of molecular charge in glomerular permeability. Tracer studies with cationized ferritins*. J Cell Biol, 1975. **67(3)**: p. 638-46.

85. Scandling, J.D. and B.D. Myers, *Glomerular size-selectivity and microalbuminuria in early diabetic glomerular disease*. *Kidney Int*, 1992. **41**(4): p. 840-6.
86. Brady, R.O., *Enzymatic abnormalities in diseases of sphingolipid metabolism*. *Clin Chem*, 1967. **13**(7): p. 565-77.
87. Askari, H., et al., *Cellular and tissue localization of globotriaosylceramide in Fabry disease*. *Virchows Arch*, 2007. **451**(4): p. 823-34.
88. Beutler, E., *Lysosomal storage diseases: natural history and ethical and economic aspects*. *Mol Genet Metab*, 2006. **88**(3): p. 208-15.
89. Benjamin, E.R., et al., *The validation of pharmacogenetics for the identification of Fabry patients to be treated with migalastat*. *Genet Med*, 2017. **19**(4): p. 430-438.
90. Giugliani, R., et al., *A Phase 2 study of migalastat hydrochloride in females with Fabry disease: selection of population, safety and pharmacodynamic effects*. *Mol Genet Metab*, 2013. **109**(1): p. 86-92.
91. Young-Gqamana, B., et al., *Migalastat HCl reduces globotriaosylsphingosine (lyso-Gb3) in Fabry transgenic mice and in the plasma of Fabry patients*. *PLoS One*, 2013. **8**(3): p. e57631.
92. Benjamin, E.R., et al., *The pharmacological chaperone 1-deoxygalactonojirimycin increases alpha-galactosidase A levels in Fabry patient cell lines*. *J Inherit Metab Dis*, 2009. **32**(3): p. 424-40.

93. Benjamin, E.R., et al., *Co-administration With the Pharmacological Chaperone AT1001 Increases Recombinant Human alpha-Galactosidase A Tissue Uptake and Improves Substrate Reduction in Fabry Mice*. Mol Ther.
94. Clark, N.E., et al., *Pharmacological chaperones for human alpha-N-acetylgalactosaminidase*. Proc Natl Acad Sci U S A, 2012. **109**(43): p. 17400-5.
95. van Diggelen, O.P., et al., *alpha-N-acetylgalactosaminidase deficiency, a new lysosomal storage disorder*. J Inherit Metab Dis, 1988. **11**(4): p. 349-57.
96. Scriver, C.R., *The metabolic and molecular bases of inherited disease*. 7th ed. ed. 1995, New York: McGraw-Hill, Health Professions Division.
97. Bleriot, Y., et al., *Synthesis of 1,2-cis-Homoiminosugars Derived from GlcNAc and GalNAc Exploiting a beta-Amino Alcohol Skeletal Rearrangement*. Org Lett, 2014. **16**(21): p. 5512-5.
98. Yu, Y., et al., *Molecular basis of 1-deoxygalactonojirimycin arylthiourea binding to human alpha-galactosidase a: pharmacological chaperoning efficacy on Fabry disease mutants*. ACS Chem Biol, 2014. **9**(7): p. 1460-9.
99. Mena-Barragan, T., et al., *pH-Responsive Pharmacological Chaperones for Rescuing Mutant Glycosidases*. Angew Chem Int Ed Engl, 2015. **54**(40): p. 11696-700.
100. Fan, J.Q., *A counterintuitive approach to treat enzyme deficiencies: use of enzyme inhibitors for restoring mutant enzyme activity*. Biol Chem, 2008. **389**(1): p. 1-11.
101. Deduve, C., *From Cytases to Lysosomes*. Fed Proc, 1964. **23**: p. 1045-9.
102. Hers, H.G., *Inborn Lysosomal Diseases*. Gastroenterology, 1965. **48**: p. 625-33.

