



## Molecular Mechanisms of Disease Pathogenesis Differ in Krabbe Disease Variants

Journal:	<i>Traffic</i>
Manuscript ID	TRA-16-0491.R2
Manuscript Type:	Original Research Article
Date Submitted by the Author:	06-Apr-2016
Complete List of Authors:	Spratley, Samantha; University of Cambridge, Cambridge Institute for Medical Research Hill, Chris; University of Cambridge, Cambridge Institute for Medical Research Viuff, Agnete; Aarhus Universitet, Department of Chemistry Edgar, James; University of Cambridge, Cambridge Institute for Medical Research Skjodt, Karsten; University of Southern Denmark, Department of Cancer and Inflammation Deane, Janet; University of Cambridge, Cambridge Institute for Medical Research
Key Words:	Krabbe Disease, Globoid cell leukodystrophy, galactocerebrosidase, glycosphingolipid, lysosomal storage disease

Original Research Article

TRA.12404

doi: 10.1111/TRA.12404

Society number TRA-16-0491

Wellcome Trust Funded- YES

Manuscript received 3 March 2016

Revised and accepted 11 April 2016

Sent to press 11 April 2016

Color figures: Figures 3, 4, 5, 7

Halftone figures: Figures 1, 2, 6

Supplemental figures (legends only to be included in typeset manuscript): 4 figures 3 tables

Synopsis included

Abstract figure included

Editorial process file included

## **Molecular Mechanisms of Disease Pathogenesis Differ in Krabbe Disease Variants**

Samantha J. Spratley<sup>1</sup>, Chris H. Hill<sup>1,2</sup>, Agnete H. Viuff<sup>3</sup>, James R. Edgar<sup>4</sup>, Karsten Skjødtt<sup>5</sup> and Janet E. Deane<sup>1,\*</sup>

<sup>1</sup> Cambridge Institute for Medical Research, Department of Pathology, University of Cambridge, Cambridge Biomedical Campus, Hills Rd, Cambridge, CB2 0XY, UK

<sup>2</sup> Current address: MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge CB2 0QH, UK

<sup>3</sup> Department of Chemistry, Aarhus University, Langelandsgade 140, 8000 Aarhus C., Denmark

<sup>4</sup> Cambridge Institute for Medical Research, Department of Clinical Biochemistry, University of Cambridge, Cambridge Biomedical Campus, Hills Rd, Cambridge, CB2 0XY, UK

<sup>5</sup> Department of Cancer and Inflammation, University of Southern Denmark, 5000 Odense, Denmark

\* Corresponding author: Janet E. Deane, Cambridge Institute for Medical Research, Department of Pathology, University of Cambridge, Cambridge Biomedical Campus, Hills Rd, Cambridge, CB2 0XY, UK, Phone: +44 1223 762815, Email: jed55@cam.ac.uk.

Running title: Differing Molecular Pathogenesis in Krabbe Disease

Keywords: Krabbe Disease, Globoid cell leukodystrophy, galactocerebrosidase, glycosphingolipid, lysosomal storage disease

**Synopsis**

Krabbe disease is a devastating neurodegenerative disorder caused by defects in the lysosomal hydrolase galactocerebrosidase (GALC). Over 100 different mutations have been identified in the *GALC* gene and are located throughout the protein. Here, we identify a number of distinct molecular mechanisms underlying the disease pathogenesis, including misfolding and incorrect post-translational modification resulting in the lack of delivery of GALC (green) to lysosomal compartments (red, cathepsin D). Understanding these molecular defects will aid the targeting of specific new therapeutics for this fatal disease.

For Peer Review

## Abstract

Krabbe disease is a severe, fatal neurodegenerative disorder caused by defects in the lysosomal enzyme galactocerebrosidase (GALC). The correct targeting of GALC to the lysosome is essential for the degradation of glycosphingolipids including the primary lipid component of myelin. Over 100 different mutations have been identified in GALC that cause Krabbe disease but the mechanisms by which they cause disease remain unclear. We have generated monoclonal antibodies against full-length human GALC and used these to monitor the trafficking and processing of GALC variants in cell-based assays and by immunofluorescence microscopy. Striking differences in the secretion, processing and endosomal targeting of GALC variants allows the classification of these into distinct categories. A subset of GALC variants are not secreted by cells, not proteolytically processed, and remain trapped in the ER; these are likely to cause disease due to protein misfolding and should be targeted for pharmacological chaperone therapies. Other GALC variants can be correctly secreted by cells and cause disease due to catalytic defects in the enzyme active site, inappropriate post-translational modification or a potential inability to bind essential co-factors. The classification of disease pathogenesis presented here provides a molecular framework for appropriate targeting of future Krabbe disease therapies.

## Introduction

Krabbe disease, also known as globoid cell leukodystrophy, is a rare, autosomal recessive disorder in which extensive demyelination is accompanied by rapid, fatal neurodegeneration. Krabbe disease is caused by defects in the lysosomal enzyme galactocerebrosidase (GALC) that is essential for the catabolism of galactosphingolipids, including the primary lipid component of myelin, galactocerebroside. The molecular understanding of Krabbe disease has been aided by the sequence characterization of the human gene, structure determination of the GALC protein and via the study of various animal models of the disease (1-5). The GALC protein is produced and glycosylated in the endoplasmic reticulum (ER) and trans-Golgi network (TGN) where it is post-translationally modified by N-linked glycosylation at four sites (3). GALC then becomes cargo for the cation-independent mannose-6-phosphate receptor (M6PR) which targets it to lysosomes either directly from the trans-Golgi network (TGN) or indirectly via secretion and re-uptake (6). It is in the endolysosome/lysosome that GALC processes lipid substrates. Sphingolipid degradation by lysosomal hydrolases also requires non-enzymatic sphingolipid activator proteins known as saposins. In rare cases, Krabbe disease can be caused by lack of functional saposin A (7).

Despite these molecular insights into GALC function the cellular events that result in apoptosis of myelin-

forming oligodendrocytes and Schwann cells of the central and peripheral nervous systems remains poorly understood. Krabbe disease severity and progression are highly variable and the age of clinical onset can vary from the first few weeks of life (early infantile) through late-infantile and juvenile into adult presentations (late-onset). To date, 147 different mutations in the GALC gene have been catalogued for the Online Metabolism and Molecular Bases of Inherited Disease (OMMBID) (8).

Currently, the only available treatment for Krabbe disease is haematopoietic stem cell transplantation but as this cannot repair damage that has already occurred it must be initiated before the onset of symptoms. For this reason early diagnosis is crucial and has led to the establishment of several newborn screening programmes (9). For mutations of known severity these screening programmes aid the rapid implementation of treatment. However, for newly identified mutations the clinical severity remains uncertain.

Understanding the pathogenesis of specific mutations in GALC is not only important for predicting the severity of newly identified mutations but is also important for determining which patients will be appropriate candidates for new treatments and combination therapies currently being developed (10-13). One such therapy being pursued for Krabbe disease and a number of related lysosomal storage diseases is pharmacological chaperone therapy (14-18). The goal of this approach is to treat patients with small molecules that can bind and stabilize partially misfolded enzyme that would otherwise remain trapped in the ER and be degraded by cellular quality-control pathways. This approach will not be appropriate for patients with large deletions in the GALC gene or for variants that are catalytically defective but must instead be targeted towards missense mutations that cause protein misfolding. For these reasons, it is important that we understand the molecular defects caused by specific mutations in GALC. Here we present the characterization of a series of Krabbe disease variants and highlight different molecular mechanisms underlying the pathogenesis of this disease.

## Results

### Missense mutations alter GALC secretion and processing

In order to understand the different effects of missense mutations on GALC processing and trafficking a series of clinically-relevant variants were selected and expressed in HEK293T cells. Two numbering schemes are commonly used for GALC based on either the first or second methionine in the ER signal sequence. Here we are using the numbering based on the second methionine as this is consistent with the majority of the literature and the available structural data. For

clarity we have included a table showing the equivalent numbering for the alternative start site (Table S1). FLAG-tagged GALC variants were detected in cell lysates and conditioned media by Western blot. Use of the anti-FLAG antibody in this case ensures that all GALC variants are detected using an identical epitope reducing any differences in binding of the antibody due specifically to the encoded mutation. All GALC variants are equally expressed in cells and individual bands are detected in cell lysates representing the addition of four N-linked glycans as GALC is produced and glycosylated in the ER and TGN (Fig. 1A and Fig. 5B). However, only a subset of these variants is secreted and detected in the conditioned media suggesting a defect in trafficking beyond the ER-TGN (Fig. 1A). For those variants that are secreted, only one band is detectable in the conditioned media identifying that these proteins possess uniform glycosylation and are therefore likely to have undergone correct processing before being secreted from cells. To confirm this, the sensitivity of GALC glycans to PNGaseF and EndoH was tested (Fig. 1B). The partially processed protein in the cell lysates was susceptible to both treatments as expected, while the protein secreted into the media was primarily EndoH resistant confirming that the glycans had been correctly modified during their trafficking through the ER-TGN.

The secretion profile illustrated here for a range of missense mutations provides an initial classification of GALC variants into those that are correctly processed and secreted and those that may cause disease due to misfolded protein being trapped in the ER. To confirm the delivery, or lack thereof, to the lysosomal compartment there is an additional post-translational processing step that occurs in the lysosome. Upon delivery to the lysosome, a loop on the surface of GALC is cleaved by a lysosomal protease. This event does not affect activity and is not likely to have any effect on the structure but can be detected by SDS-PAGE as 50- and 30-kDa bands (6, 19, 20). In the previous assays (Fig. 1A and B) there was no detectable band for the lower molecular weight cleavage product expected upon delivery of GALC to the lysosome. However, as detection of GALC involved using an antibody against the epitope tag it was possible that in the proteolytic environment of the endolysosomal system this tag is cleaved from the GALC protein restricting detection of the processed form. To overcome this, we generated a new panel of monoclonal antibodies raised against the full-length human GALC protein. We carried out an immunoprecipitation (IP) of GALC from cell lysates using either the FLAG antibody or two different mouse monoclonal antibodies (mAb1, mAb2). Following IP, bands were detected by Western blot using a rabbit polyclonal antibody raised against a peptide within the smaller C-terminal 30-kDa fragment (Fig. 2A). For all samples the full-length uncleaved GALC is detectable at approx. 80 kDa. In both the IPs carried out with monoclonal antibodies an additional band is present at 30-kDa consistent with the fragment size following cleavage upon delivery to the lysosome. This band is not

detectable in the IP using the anti-FLAG antibody highlighting the critical importance of appropriate antibodies for assay development. To confirm that the cleavage event occurs post-Golgi and thus is a measure of successful trafficking to lysosomes we carried out an uptake assay whereby conditioned media containing full-length uncleaved GALC was applied to untransfected HEK293T cells (Fig. 2B). The appearance of the 30-kDa fragment in untransfected cells treated with full-length GALC identifies that the detection of cleaved GALC can provide a measure of lysosomal delivery of wild-type (WT) and mutant forms of GALC. Cell lysates from GALC missense mutations that are not secreted into the conditioned media, such as R515H and T513M (Fig. 1A), have no detectable 30-kDa cleavage product (Fig. 2C) consistent with these proteins being trapped in the ER due to misfolding and therefore are not delivered to lysosomes. Missense mutations that are secreted into the conditioned media, such as R380W and E215K, show reduced but still detectable levels of GALC cleavage product.

#### **Non-secreting, unprocessed missense mutations of GALC are trapped in the ER**

Absence of GALC in the conditioned media and the lack of cleaved product in cell lysates together provide compelling evidence that a particular GALC variant causes disease due to compromised protein folding and is thus retained in the ER. However, both these measures are indirect and so several GALC variants were further tested by direct monitoring of their cellular localization. GALC variants were expressed in HeLa cells and detected by immunofluorescence microscopy using monoclonal antibodies raised against full-length human GALC and compared with markers for the ER (calreticulin, calnexin and REEP5) and lysosome (cathepsin D). Immunofluorescence microscopy clearly identifies that WT GALC co-localizes with cathepsin D identifying its correct trafficking to lysosomes (Fig. 3A). Missense mutations that were shown to inhibit secretion of GALC and prevent normal processing in the lysosome were tested for their cellular localization (Fig. 3B, Fig. S1 and S2). T513M and L618S show clear reticular staining consistent with ER architecture and very poor co-localization with the lysosomal marker (Fig. 3C). The co-localization of these variants with the ER markers confirmed their ER localization (Fig. 3B, 3C, Fig. S1). Additional missense mutations I583S, R515H and Y319C were also identified to co-localize with the ER markers calreticulin and calnexin (Fig. S2).

Missense mutations of GALC that are secreted by cells and show cleavage consistent with lysosomal localization were also tested for their cellular localization. The residue R380 has been shown by us previously to be critical for enzyme activity as it directly binds substrate in the active site (3, 21). In this case, we would expect this mutation not to interfere with folding and to traffic to the lysosome where the disease-causing defect is primarily catalytic. In agreement with this, immunostaining of the R380W mutant identifies that it co-localizes with

lysosomes to the same extent as WT GALC (Fig. 4A and 4C). GALC containing the missense mutation E215K can be secreted into the media and shows cellular localization similar to WT GALC (Fig. 4B). Consistent with the secretion profile, E215K is not primarily co-localized with the ER marker calreticulin suggesting that the disease mechanism in this case is not due to a misfolding defect (Fig. 4C).

#### **N279T introduces a new glycosylation site that interferes with protein trafficking**

The missense mutation N279T is detected in conditioned media but runs at a slightly higher molecular weight compared with WT and other missense variants (Fig. 1A). Analysis of the sequence surrounding this residue identifies that this mutation introduces a new consensus site for N-linked glycosylation of residue N277 (Fig. 5A). This additional glycan can be detected in cell lysates and can be abolished by introduction of the second mutation N277D to block glycosylation of this residue (Fig. 5B). Unlike other secreted GALC variants, N279T does not localize to lysosomes (Fig. 5C). Introduction of an additional glycan at residue N277 would be predicted to destabilize the fold of this GALC variant as this residue is deeply buried at the interface between two domains (Fig. 5D). Although introduction of the second mutation N277D blocks addition of the extra glycan, the double mutant does not rescue trafficking to the lysosome (Fig. 5C). Despite the conservative nature of the introduced N277D mutation, this change alone causes misfolding of GALC such that it is not trafficked to the lysosome (Fig. 5C). The observation that N279T can be secreted but not enter the endocytic pathway suggests that some protein can escape ER-associated quality control but is defective for binding to the mannose 6-phosphate receptor. To test this, an uptake assay was carried out whereby conditioned media containing the N279T variant was applied to untransfected cells in a manner similar to that shown for WT GALC (Fig. 5E). Unlike WT GALC, the N279T variant cannot be taken up by cells suggesting that the additional glycan interferes with endocytosis.

#### **The polymorphism at position 546 has subtle effects on enzyme activity, stability and trafficking**

The missense mutations described here have been tested in the most common polymorphic background I546. However the polymorphism T546 has been reported to be present within the population at allele frequency of 40-45% (22). Clinical assays for GALC activity in peripheral blood have suggested that the T546 variant has half the activity of the I546 variant, contributing to the highly variable "normal" activity levels within the general population (22-24). In order to test the effect of this polymorphism on GALC activity we generated stable expression constructs of each of these polymorphisms and selected high-yield clones in order to purify milligram quantities of GALC for steady-state kinetic studies (Fig. 6A and Table S2).

Surprisingly, the two variants do not have significantly different catalytic properties *in vitro*: T546 has 89% activity of I546 suggesting that the observed differences in activity from blood samples may be due to differential stability, expression or trafficking in cells. Thermal denaturation of secreted GALC monitored by differential scanning fluorimetry revealed a subtle destabilization of the T546 variant *in vitro* (Fig. 6B;  $T_m$ , I546 =  $51.8 \pm 0.2$  °C vs  $T_m$ , T546 =  $50.6 \pm 0.2$  °C). To determine whether this degree of destabilization was sufficient to impair trafficking of GALC the presence of T546 in conditioned media was monitored. Interestingly, this variant was not detectable in the media suggesting it is trapped in the ER (Fig. 7A). However, if this polymorphism completely blocked trafficking of GALC to the lysosome this would represent a disease-causing mutation. Cellular localization of T546 by immunofluorescence microscopy confirms that although this variant is not detected in conditioned media it can be trafficked to the lysosome (Fig. 7B). The combination of multiple, subtle deleterious effects on activity, stabilization and trafficking of the T546 variant may explain the observed reduction in enzyme activity from peripheral blood samples.

Although the I546T polymorphism on its own has only subtle effects on GALC function, when combined with additional missense mutations the cumulative effects can become severe. Indeed, it has been suggested previously that certain mutations are only disease-causing in specific polymorphic backgrounds (25). The G270D variant is responsible for a late-onset form of Krabbe disease and is usually found in *cis* with the I546T polymorphism (26). In the I546 background, the G270D variant possesses WT levels of protein secreted in to the media and WT localization to lysosomes (Fig. 7A, C). However, in the T546 background, the G270D mutation is severely compromised in its cellular trafficking and is not localized to lysosomes (Fig. 7A and 7C, Fig. S3). The cumulative effect of these two sequence variations is not due to a single localized structural perturbation as the two residues, G270 and I546, lie in distant parts of the GALC structure (Fig. 7D). Specifically, G270 is situated near the disulphide bond that stabilizes a critical active-site loop structure while residue 546 lies in the C-terminal lectin domain.

The missense mutation Y551S causes infantile Krabbe disease and has been identified in the T546 polymorphic background (26). These two residues are much closer to each other in the GALC structure and so might be expected to have a severe cumulative effect on folding (Fig. 7D). However, the secretion profile of this mutation shows that even in the I546 background the Y551S mutation has a severe detrimental effect on protein folding and secretion (Fig. 7A). In agreement with this, the Y551S variant in both polymorphic backgrounds does not traffic correctly to lysosomes (Fig. 7C and Fig. S3).



## Discussion

In order to better understand the underlying molecular mechanisms causing GALC dysfunction in Krabbe disease, clinically-relevant variants of GALC were expressed in cells and monitored for their capacity to traffic beyond the ER-TGN and to localize to their site of action in lysosomes. Using a series of cell-based assays we have identified the underlying pathogenesis for a number of these variants. Initial screening of the secretion profile of GALC variants provides a good first indication of whether a mutation may result in protein misfolding. The protein quality control pathways present in the ER should result in the retention and degradation of misfolded protein. In this case misfolded GALC will not traffic beyond the ER and thus will not be detected in the conditioned media. Using this approach we identified seven variants (Y319C, L618S, G41S, T513M, I583S, R515H and R63H) that are likely to cause disease due to protein misfolding. Another subset of GALC variants (R380W, E215K, N279T, P302R and R380L) were present in the conditioned media identifying that the underlying pathology in these cases is not primarily protein misfolding. A further assay was used to monitor trafficking beyond the ER-TGN by exploiting the observation that upon delivery of GALC to lysosomes, a 30-kDa cleavage product can be identified following SDS-PAGE and immunoblotting. In support of their predicted defect in protein folding, two of the missense mutations identified as non-secretors, R515H and T513M, were not processed into this smaller fragment. Consistent with the secretion assay, two of the secreted variants (R380W and E215K) were shown to undergo proteolytic processing similar to that seen for WT GALC.

Lack of protein secretion and the inability to be processed into the 50- and 30-kDa fragments are indicative of a misfolded GALC variant suggesting it is retained in the ER. To further test this, the cellular localization of GALC variants was directly monitored by immunofluorescence microscopy using newly-generated, specific monoclonal antibodies raised against full-length human GALC protein. Using this approach, the GALC variants, T513M, L618S, Y319C, I583S and R515H were not localized to lysosomes and were trapped in the ER (Fig. 3, Fig. S1 and S2). These GALC variants can be classified as protein misfolding mutants due to their lack of secretion, inability to be proteolytically processed and their retention in the ER.

Those disease variants that do not result in significant misfolding, based on their secretion, processing and cellular localization, fall into a number of categories. The residue R380 has been shown by us previously to play a critical role in the catalytic processing of substrates (3, 21). In support of this, R380W and R380L are secreted similarly to WT GALC and R380W was shown to localize to lysosomes. E215K was also normally secreted and trafficked within the cell but is not likely to be catalytically defective as this mutation does not lie near the enzyme active site. This residue is on the surface of GALC and the charge inversion caused

by this mutation may interfere with important interactions with partner proteins such as saposin A. The mutation N279T was shown to introduce a new glycosylation consensus sequence resulting in the addition of an N-linked glycan to residue N277. This variant is capable of being secreted by cells and so can pass ER quality control but is unable to be taken up by cells or traffic to lysosomes suggesting that the new glycan may interfere with endocytosis by altering binding to the M6PR.

A number of polymorphisms have been identified in the GALC gene and have been shown to have a negative effect on GALC activity. In this study we produced both the I546 and T546 variants of GALC and compared their activity and stability *in vitro*. Surprisingly this polymorphism has only a subtle effect on GALC enzyme kinetics and thermal stability but did have a significant effect on its trafficking beyond the ER-TGN. This suggests that the lower activity reported for this polymorphism is a reflection of its reduced trafficking rather than its inability to process substrate. It has been shown previously that the polymorphic background that a missense mutation occurs in can critically change the disease pathogenesis (25, 27). In agreement with this, we show that G270D only results in protein misfolding, as measured by ER co-localization, when present in the T546 background. Thus in this case, it is the combination of two subtle defects that results in disease. However, despite the Y551S variant only being identified in the T546 background, it results in protein misfolding in both polymorphic backgrounds identifying that this mutation alone is sufficient to cause significant protein misfolding.

The availability of structural data for GALC allows prediction of the functional effects of missense mutations: the greater proportion of a side chain that is buried in the structure the greater the likelihood that a mutation of that residue will disrupt the fold (3). These predictions can now be compared with the functional effects identified here (Table S3). The side chains for residues G41, Y319, T513, I583 and L618 are completely buried in the structure suggesting that mutation of these would result in misfolding and in agreement with this prediction, all were not secreted by cells and/or were localized to the ER. However, G270 is also completely buried but the G270D variant is only localized to the ER when present in the T546 background. The side chains of R63, R515 and Y551 are also primarily (~80%) buried in the structure and are here identified as non-secretors identifying that even only partially buried residues can significantly disrupt folding. However, based on buried surface area, P302 would also fall into this category but is shown here to be secreted in disagreement with the structure-based prediction. E215 and R380 are relatively exposed on the surface of GALC (only 57% and 48% buried) suggesting that residues that are <60% buried are unlikely to result in misfolding. Based on this analysis, the structure provides a good molecular framework for making predictions but verification in cell-based assays is still required.

In this study we have used three different assays to monitor trafficking defects in GALC variants. Each assay has different levels of complexity and sensitivity allowing a confidence level to be associated with the designation of the molecular defect caused by a specific mutation. The importance of using multiple assays is best highlighted by those variants that have subtle effects such as the I546T polymorphism. Although this polymorphism is not disease-causing, in our secretion assay it is not detectable suggesting that this assay alone is not sufficiently sensitive to distinguish subtle GALC defects. However, monitoring cellular localization by immunostaining appears to be much more sensitive as in this case the polymorphic variant can be localized to lysosomes. Thus although the secretion assay is a fast and convenient readout it is important that conclusions regarding the underlying pathogenesis of subtle variants be confirmed using the more sensitive co-localization assay.

Extrapolation from disease mechanism to disease severity remains highly challenging. The disease-causing variant L618S has a complex effect in patients and has been identified in patients with age of onset ranging from late-infantile to adult (28-30). This variation in disease severity is likely influenced by many factors including the polymorphic background and compound heterozygosity both of which have been documented for this variant. Recently, Shin et al also showed that L618S is severely compromised in its ability to be correctly processed into the 50- and 30-kDa fragments supporting a defect in folding (31). However, unlike our localization studies identifying L618S as trapped in the ER, Shin et al found some evidence that this variant co-stains with LAMP2 suggesting it maintains some capacity to be delivered to late endosomes/lysosomes. For this reason one must be cautious to assign degree of disease severity based on these assays but should instead use the identification of the nature of the defect as a critical factor for subsequent therapeutic targeting. Mutations that cause misfolding, giving rise to lack of secretion, processing and primarily ER co-localization, may be good targets for pharmacological chaperone therapies while those that retain significant capacity for correct trafficking and processing will require future enzyme replacement approaches. However, the extent of misfolding caused by specific mutations will alter

the effectiveness of pharmacological chaperones. Specifically, the misfolding caused by some mutations may prove too severe to respond to these approaches and will therefore require enzyme replacement strategies.

Studies in related lysosomal storage disorders have examined the effects of missense mutations on enzyme processing and trafficking. Retention of misfolded protein in the ER is an important factor in specific variants that cause Gaucher disease (32-35), Pompe disease (36) and GM1 gangliosidosis (37). In several cases, this insight has been used to examine if pharmacological chaperones, ERAD inhibitors and proteostasis regulators can rescue these misfolded variants and restore trafficking to the lysosome. Although these studies have provided some potential leads for therapeutic development, they have also highlighted that these approaches are often only successful with a limited subset of mutations potentially due to the severity of the misfolding defect (mentioned above). One of the current limitations in drug development for Krabbe disease is the lack of reliable high-throughput (HT) approaches for the screening of new therapeutics against the wide range of mutations present in the patient population. Although activity measurements can be carried out in a HT manner (38), previous work has shown that GALC activity does not correlate well with disease severity and may not provide a good measure for treatment efficacy (25, 31). One of the primary reasons for this may be attributed to the fact that the activity assays are often conducted on whole cell lysates and so cannot distinguish between protein that is localized in the ER or the endocytic pathway leaving the question as to whether the active protein is still mislocalized unanswered. Recent work measuring the activity of samples enriched for the lysosomal fraction show a better correlation with disease severity providing a good first step towards more reliable activity measurements (31). Future development of the assays described here into a HT format would provide a complementary readout to these activity assays and confirm those conditions where GALC is successfully trafficked beyond the ER-TGN and thus corrected for misfolding defects.



## Materials and Methods

**Cloning and mutagenesis of human GALC constructs.** For large-scale protein expression in mammalian cells wild-type human GALC (hGALC) polymorphisms I546 and T546 were cloned into pSecTag2B vector with an N-terminal His6 tag and stable HEK293T cell lines were established and cultured as described previously (3). For transient expression of mutants in mammalian tissue culture mutant hGALC constructs were generated in the pSecTag2b vector by Quikchange site directed mutagenesis of wild type hGALC in either the I546 or T546 background (as indicated). Constructs were then sub-cloned into the pHLsec vector resulting in removal of the His tag and addition of a C-terminal FLAG epitope tag.

**Expression and purification of human GALC protein.** Large-scale expression and IMAC-affinity purification of hGALC proteins from HEK293T cells for antibody (Ab) generation, activity assays and thermal denaturation experiments was carried out as described previously for the equivalent mouse GALC constructs (21, 39).

**Monoclonal antibody generation.** Monoclonal mouse anti-human GALC antibodies were raised against full-length secreted human GALC protein. Mice were immunized subcutaneously 2 times with 25 µg of purified human GALC, mixed with GERBU® adjuvant. Three days before the fusion, the mice received an i.v. injection with 10 µg of antigen together with adrenaalin. Spleen cells and SP2 myeloma cells were used for fusion. Positive clones were selected and cloned by repeated screening against the GALC protein by ELISA. The performance and specificity of several antibody clones were validated by Western blot and immunofluorescence microscopy against transfected cells.

**Cell culture.** HEK293T and HeLa cells were maintained in Dulbecco's Modified Eagle Medium supplemented with 10% foetal bovine serum, 50 units/ml Penicillin, 50 µg/ml Streptomycin and 4 mM L-glutamine. Cells were maintained at 37°C in humidified air with 5% CO<sub>2</sub>. HEK293T cells used for immunoblot studies were grown in Freestyle 293 expression medium and transiently transfected using Lipofectamine-2000 transfection reagent. HeLa cells used for immunostaining were transiently transfected at 40% confluence using the TransIT-HeLa Monster transfection reagent. For GALC uptake assays, serum free conditioned media from transfected or untransfected cells was harvested and filtered 72 hours post-transfection and then applied to untransfected cells for 48 hours. Cells were lysed in 1% Triton X-100 and subjected to immunoprecipitation using monoclonal mouse anti-GALC antibody in the presence of protein A sepharose beads. GALC cellular uptake was detected by immunoblot, as described below.

**Immunoblotting and immunoprecipitation of GALC.** Secreted protein was detected from

conditioned media harvested 72 hours post-transfection. Cells were lysed in either 1% Triton X-100 or 1% SDS in Tris-buffered saline (TBS, 20 mM Tris-HCl, 150 mM NaCl, 5 mM MgCl<sub>2</sub>, 1 mM EDTA) in the presence of 10 mM N-ethylmaleimide (NEM), 1 mM phenylmethylsulfonyl fluoride (PMSF) and protease inhibitors. Where appropriate, protein concentration was quantified via BCA assay. For immunoprecipitation, 1000 µg cell lysate was precleared using IgG/protein A-sepharose beads. Pre-cleared lysate was incubated with anti-FLAG or anti-GALC antibodies in the presence of A-sepharose beads and washed in TBS supplemented with 0.1% Triton X-100. All samples were denatured and reduced in SDS loading buffer containing 20 mM DTT prior to SDS-PAGE separation. For immunoblot studies, proteins were transferred to methanol-activated Immobilon-FL PVDF membrane. After blocking, membranes were incubated with primary antibody followed by incubation with infrared fluorophore-conjugated secondary antibody. Immuno-reactive bands were visualised with an Odyssey infrared imager (LI-COR Biosciences).

### Confocal immunofluorescence microscopy.

Cells were grown on glass coverslips, washed with PBS and fixed in 4% PFA. Fixed cells were washed in PBS, quenched with 15mM glycine and permeabilised in 0.1% saponin in PBS. Fixed cells were blocked in 1% BSA, 0.01% saponin in PBS before incubation with primary antibody diluted in 1% BSA with 0.01% saponin followed by fluorescently-conjugated secondary antibody. Coverslips were mounted onto slides using mounting medium with DAPI (*Prolong Antifade*, Invitrogen Molecular Probes). Images were acquired on a Zeiss LSM880 confocal microscope with Zeiss zen software. To analyse co-localization, intensities were measured for ≥20 individual cells from ≥3 independent experiments and Pearson's correlation coefficients were calculated using the Volocity 5.2 software (Perkin Elmer). Automatic Costes thresholds were applied to all images (40). Significance was calculated using two-tailed t-tests implemented in Prism5 (Graphpad), error bars represent SEM. Endogenous GALC is not detectable by immunostaining of untransfected HeLa cells (Fig. S4A and B). HeLa cells transfected with the N-terminally His-tagged construct of GALC show equivalent cellular localization as the C-terminally FLAG-tagged construct (Fig. S4).

**Antibodies.** The antibodies used for immunostaining were: monoclonal mouse anti-GALC clone mAb6 (described above), polyclonal rabbit anti-cathepsin D (Calbiochem), polyclonal rabbit calreticulin (Pierce), polyclonal rabbit calnexin (Abcam), polyclonal rabbit anti-REEP5 (proteintech), polyclonal goat AlexaFluor488-conjugated anti-mouse IgG (Life technologies) and polyclonal goat AlexaFluor555-conjugated anti-rabbit IgG (Life technologies). Antibodies used for immunoblotting were: monoclonal mouse anti-FLAG (Sigma), monoclonal mouse anti-GALC clones mAb1 and mAb2 (described above), polyclonal rabbit anti-GALC (Abcam), polyclonal rabbit anti-

actin (Sigma), polyclonal goat IRdye680-conjugated anti-rabbit IgG (LiCor) and polyclonal goat IRdye800-conjugated anti-mouse IgG (LiCor).

**Deglycosylation of GALC.** Conditioned media and cells lysed in 1% SDS (as described above) were denatured at 100°C for 10 mins and were either subjected to EndoH or PngaseF treatment (NEB) for 4 hours at 37°C.

**Enzyme activity assays.** Activity assays were carried out as described previously (21). In brief, activity assays were carried out at 37°C using purified hGALC in 20 mM sodium acetate, 150 mM NaCl, 0.1% v/v NP-40, pH 4.6 and chromogenic substrate 4-nitrophenyl- $\beta$ -D-galactopyranoside (4N $\beta$ DG). Product 4-nitrophenol was monitored by absorbance at 410nm following addition of stopping buffer (360 mM NaOH, 280 mM glycine, pH 10.6).  $K_m$  and  $V_{max}$  were obtained from plots of initial velocity against substrate concentration by non-linear curve-fitting to the Michaelis-Menten equation using Prism5 (GraphPad).  $k_{cat}$  was determined as the gradient of the linear plot of  $V_{max}$  against enzyme concentration.

**Differential scanning fluorimetry.** DSF experiments were carried out as described previously (39). In brief, 5  $\mu$ g of purified hGALC in PBS, pH 7.4 was combined with 5x SyPRO Orange dye prior to thermal denaturation performed using a Bio-Rad MiniOpticon RT-PCR thermal cycler. The melting temperature ( $T_m$ ) was the inflexion point of the sigmoidal curve obtained by curve fitting using DSF analysis scripts and Prism5 (GraphPad).

## Acknowledgements

We thank Scottie Robinson, Paul Luzio and Lena Wartosch for helpful discussions. We thank Colin Crump for providing ER-specific antibodies. S.J.S. is funded by an MRC PhD studentship, C.H.H. was supported by a Wellcome Trust PhD studentship, J.E.D. is supported by a Royal Society University Research Fellowship (UF100371). We are also grateful for funding from The Lundbeck Foundation to A.H.V. The Cambridge Institute for Medical Research is supported by Wellcome Trust Strategic Award 100140.

## References

- Chen YQ, Rafi MA, de Gala G, Wenger DA. Cloning and expression of cDNA encoding human galactocerebrosidase, the enzyme deficient in globoid cell leukodystrophy. *Hum Mol Genet* 1993;2:1841-1845.
- Sakai N, Inui K, Fujii N, Fukushima H, Nishimoto J, Yanagihara I, Isegawa Y, Iwamatsu A, Okada S. Krabbe disease: isolation and characterization of a full-length cDNA for human galactocerebrosidase. *Biochem Biophys Res Commun* 1994;198:485-491.
- Deane JE, Graham SC, Kim NN, Stein PE, McNair R, Cachon-Gonzalez MB, Cox TM, Read RJ. Insights into Krabbe disease from structures of galactocerebrosidase. *Proc Natl Acad Sci U S A* 2011;108:15169-15173.
- Duchen LW, Eicher EM, Jacobs JM, Scaravilli F, Teixeira F. Hereditary leucodystrophy in the mouse: the new mutant twitcher. *Brain* 1980;103:695-710.
- Sakai N, Inui K, Tatsumi N, Fukushima H, Nishigaki T, Taniike M, Nishimoto J, Tsukamoto H, Yanagihara I, Ozono K, Okada S. Molecular cloning and expression of cDNA for murine galactocerebrosidase and mutation analysis of the twitcher mouse, a model of Krabbe's disease. *J Neurochem* 1996;66:1118-1124.
- Nagano S, Yamada T, Shinnoh N, Furuya H, Taniwaki T, Kira J. Expression and processing of recombinant human galactosylceramidase. *Clin Chim Acta* 1998;276:53-61.
- Spiegel R, Bach G, Sury V, Mengistu G, Meidan B, Shalev S, Shneor Y, Mandel H, Zeigler M. A mutation in the saposin A coding region of the prosaposin gene in an infant presenting as Krabbe disease: first report of saposin A deficiency in humans. *Mol Genet Metab* 2005;84:160-166.
- Wenger DA, Escolar ML, Luzi P, Rafi MA. Krabbe disease (globoid cell leukodystrophy). *The Online Metabolism & Molecular Bases of Inherited Disease*: McGraw Hill; 2013.
- Duffner PK, Caggana M, Orsini JJ, Wenger DA, Patterson MC, Crosley CJ, Kurtzberg J, Arnold GL, Escolar ML, Adams DJ, Andriola MR, Aron AM, Ciafaloni E, Djukic A, Erbe RW, *et al.* Newborn screening for Krabbe disease: the New York State model. *Pediatr Neurol* 2009;40:245-252; discussion 253-245.
- Hawkins-Salsbury JA, Shea L, Jiang X, Hunter DA, Guzman AM, Reddy AS, Qin EY, Li Y, Gray SJ, Ory DS, Sands MS. Mechanism-based combination treatment dramatically increases therapeutic efficacy in murine globoid cell leukodystrophy. *J Neurosci* 2015;35:6495-6505.
- Rafi MA, Rao HZ, Luzi P, Curtis MT, Wenger DA. Extended normal life after AAVrh10-mediated gene therapy in the mouse model of Krabbe disease. *Molecular therapy : the journal of the American Society of Gene Therapy* 2012;20:2031-2042.
- Nagabhushan Kalburgi S, Khan NN, Gray SJ. Recent gene therapy advancements for neurological diseases. *Discov Med* 2013;15:111-119.
- Hawkins-Salsbury JA, Reddy AS, Sands MS. Combination therapies for lysosomal storage disease: is the whole greater than the sum of its parts? *Hum Mol Genet* 2011;20:R54-60.
- Lee WC, Kang D, Causevic E, Herdt AR, Eckman EA, Eckman CB. Molecular characterization of mutations that cause globoid cell leukodystrophy and pharmacological rescue using small molecule chemical chaperones. *J Neurosci* 2010;30:5489-5497.
- Berardi AS, Pannuzzo G, Graziano A, Costantino-Ceccarini E, Piomboni P, Luddi A. Pharmacological chaperones increase residual beta-galactocerebrosidase activity in fibroblasts from Krabbe patients. *Mol Genet Metab* 2014;112:294-301.
- Hossain MA, Higaki K, Saito S, Ohno K, Sakuraba H, Nanba E, Suzuki Y, Ozono K, Sakai N. Chaperone therapy for Krabbe disease: potential for late-onset GALC mutations. *J Hum Genet* 2015;60:539-545.
- Valenzano KJ, Khanna R, Powe AC, Boyd R, Lee G, Flanagan JJ, Benjamin ER. Identification and characterization of pharmacological chaperones to correct enzyme deficiencies in lysosomal storage disorders. *Assay Drug Dev Technol* 2011;9:213-235.
- Parenti G. Treating lysosomal storage diseases with pharmacological chaperones: from concept to clinics. *EMBO Mol Med* 2009;1:268-279.
- Sakai N, Inui K, Midorikawa M, Okuno Y, Ueda S, Iwamatsu A, Okada S. Purification and characterization of galactocerebrosidase from human lymphocytes. *J Biochem* 1994;116:615-620.
- Chen YQ, Wenger DA. Galactocerebrosidase from human urine: purification and partial characterization. *Biochim Biophys Acta* 1993;1170:53-61.
- Hill CH, Graham SC, Read RJ, Deane JE. Structural snapshots illustrate the catalytic cycle of beta-galactocerebrosidase, the defective enzyme in Krabbe disease. *Proc Natl Acad Sci U S A* 2013;110:20479-20484.
- Wenger DA, Rafi MA, Luzi P, Datto J, Costantino-Ceccarini E. Krabbe disease: genetic aspects and progress toward therapy. *Mol Genet Metab* 2000;70:1-9.
- De Gasperi R, Gama Sosa MA, Sartorato EL, Battistini S, MacFarlane H, Gusella JF, Krivit W, Kolodny EH. Molecular heterogeneity of late-onset forms of globoid-cell leukodystrophy. *Am J Hum Genet* 1996;59:1233-1242.
- Raghavan S, Zeng B, Torres PA, Pastores GM, Kolodny EH, Kurtzberg J, Krivit W. Globoid cell leukodystrophy (Krabbe disease): normal umbilical cord blood galactocerebrosidase activity and polymorphic mutations. *J Inherit Metab Dis* 2005;28:1005-1009.
- Wenger DA, Luzi P, Rafi MA. Krabbe disease: are certain mutations disease-causing only when specific polymorphisms are present or when inherited in trans with specific second mutations? *Mol Genet Metab* 2014;111:307-308.
- Wenger DA, Rafi MA, Luzi P. Molecular genetics of Krabbe disease (globoid cell leukodystrophy): diagnostic and clinical implications. *Hum Mutat* 1997;10:268-279.
- Furuya H, Kukita Y, Nagano S, Sakai Y, Yamashita Y, Fukuyama H, Inatomi Y, Saito Y,

- Koike R, Tsuji S, Fukumaki Y, Hayashi K, Kobayashi T. Adult onset globoid cell leukodystrophy (Krabbe disease): analysis of galactosylceramidase cDNA from four Japanese patients. *Hum Genet* 1997;100:450-456.
28. Satoh JI, Tokumoto H, Kurohara K, Yukitake M, Matsui M, Kuroda Y, Yamamoto T, Furuya H, Shinnoh N, Kobayashi T, Kukita Y, Hayashi K. Adult-onset Krabbe disease with homozygous T1853C mutation in the galactocerebrosidase gene. Unusual MRI findings of corticospinal tract demyelination. *Neurology* 1997;49:1392-1399.
29. Xu C, Sakai N, Taniike M, Inui K, Ozono K. Six novel mutations detected in the GALC gene in 17 Japanese patients with Krabbe disease, and new genotype-phenotype correlation. *J Hum Genet* 2006;51:548-554.
30. Hossain MA, Otomo T, Saito S, Ohno K, Sakuraba H, Hamada Y, Ozono K, Sakai N. Late-onset Krabbe disease is predominant in Japan and its mutant precursor protein undergoes more effective processing than the infantile-onset form. *Gene* 2014;534:144-154.
31. Shin D, Feltri ML, Wrabetz L. Altered trafficking and processing of GALC mutants correlates with globoid cell leukodystrophy severity. *J Neurosci* 2016;36:1858-1870.
32. Wang F, Song W, Brancati G, Segatori L. Inhibition of endoplasmic reticulum-associated degradation rescues native folding in loss of function protein misfolding diseases. *J Biol Chem* 2011;286:43454-43464.
33. Schmitz M, Alfalah M, Aerts JM, Naim HY, Zimmer KP. Impaired trafficking of mutants of lysosomal glucocerebrosidase in Gaucher's disease. *Int J Biochem Cell Biol* 2005;37:2310-2320.
34. Sawkar AR, Schmitz M, Zimmer KP, Reczek D, Edmunds T, Balch WE, Kelly JW. Chemical chaperones and permissive temperatures alter localization of Gaucher disease associated glucocerebrosidase variants. *ACS Chem Biol* 2006;1:235-251.
35. Ron I, Horowitz M. ER retention and degradation as the molecular basis underlying Gaucher disease heterogeneity. *Hum Mol Genet* 2005;14:2387-2398.
36. Parenti G, Zuppaldi A, Gabriela Pittis M, Rosaria Tuzzi M, Annunziata I, Meroni G, Porto C, Donaudy F, Rossi B, Rossi M, Filocamo M, Donati A, Bembi B, Ballabio A, Andria G. Pharmacological enhancement of mutated alpha-glucosidase activity in fibroblasts from patients with Pompe disease. *Molecular therapy : the journal of the American Society of Gene Therapy* 2007;15:508-514.
37. Zhang S, Bagshaw R, Hilson W, Oho Y, Hinek A, Clarke JT, Callahan JW. Characterization of beta-galactosidase mutations Asp332-->Asn and Arg148-->Ser, and a polymorphism, Ser532-->Gly, in a case of GM1 gangliosidosis. *Biochem J* 2000;348 Pt 3:621-632.
38. Ribbens J, Whiteley G, Furuya H, Southall N, Hu X, Marugan J, Ferrer M, Maegawa GH. A high-throughput screening assay using Krabbe disease patient cells. *Anal Biochem* 2013;434:15-25.
39. Hill CH, Viuff AH, Spratley SJ, Salamone S, Christensen SH, Read RJ, Moriarty NW, Jensen HH, Deane JE. Azasugar inhibitors as pharmacological chaperones for Krabbe disease. *Chem Sci* 2015;6:3075-3086.
40. Costes SV, Daelemans D, Cho EH, Dobbin Z, Pavlakis G, Lockett S. Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophys J* 2004;86:3993-4003.

## Figure Legends

**Figure 1. Missense mutations alter GALC secretion from cells.** HEK293T cells were transiently transfected with FLAG-tagged wild-type (WT) GALC and a panel of clinically relevant Krabbe disease mutations. Conditioned media was harvested at 72 hours and cells lysed in 1% SDS followed by SDS-PAGE and immunoblotting. A) Secreted and total GALC expression were detected by Western blot using an anti-FLAG antibody. Actin served as a loading control. B) Conditioned media and cell lysates were digested with either EndoH or with PNGaseF before immunoblotting with anti-FLAG to detect GALC.

**Figure 2. Proteolytic cleavage of WT and mutant GALC.** A) Protein lysates from either untransfected HEK293T cells or from cells transfected with WT GALC were subjected to immunoprecipitation with monoclonal antibodies against full-length human GALC (mAb1 and mAb2) or the FLAG epitope tag. GALC was detected by Western blot using a polyclonal rabbit GALC antibody, cleavage product at ~30-kDa is marked with an asterisk. B) Conditioned media (CM) from HEK293T cells transfected with WT GALC contains full-length GALC (lane 1). GALC-containing CM (+) and CM from untransfected cells (-) was harvested after 72 hours and applied to untransfected HEK293T cells for 48 hours. Treated cells were subjected to immunoprecipitation using the GALC monoclonal antibody mAb2 and detected as in A. C) HEK293T cells were transiently transfected with WT GALC and Krabbe mutations R380W, E215K, R515H and T513M and immunoprecipitated using the mAb2 antibody and detected as in A.

**Figure 3. Missense mutations of GALC that interfere with secretion and cleavage are trapped in the ER.** Representative confocal microscopy images of HeLa cells transiently transfected with either WT GALC (A), or the Krabbe mutations T513M and L618S (B). Cells were plated onto glass coverslips, fixed and immunostained using monoclonal antibody against GALC (green), the lysosomal marker cathepsin D (red) or the ER marker calreticulin (red). Nuclei were stained with DNA-binding dye, DAPI (blue). Scale bar 10  $\mu$ m. (C) To quantify co-localization, Pearson's correlation coefficients were calculated for each variant with both cathepsin D and calreticulin. Mean  $\pm$  SEM for at least 20 individual cells from  $\geq 3$  independent experiments are shown. \*\*\* $p$  < 0.0001, calculated using a two-tailed unpaired t-test.

**Figure 4. Missense mutations R380W and E215K are trafficked to the lysosome.** Representative confocal microscopy images of HeLa cells transiently transfected with Krabbe mutations R380W (A) and E215K (B). Transfected cells were plated onto glass coverslips, fixed and immunostained for GALC (green), the lysosomal marker cathepsin D (red) or the ER marker calreticulin (red). Scale bar 10  $\mu$ m. (C) Pearson's correlation coefficients were calculated for each of the mutations with both cathepsin D and calreticulin. Mean  $\pm$  SEM for at least 20 individual cells from  $\geq 3$  independent experiments are shown. \*\* $p$  < 0.005, \* $p$  < 0.05, ns = not significant, calculated using a two-tailed unpaired t-test.

**Figure 5. N279T is a hyperglycosylation mutant with defective trafficking.** A) Protein sequence of WT and N279T GALC illustrating the introduction of a new consensus glycosylation site (NXT). B) HEK293T cells transiently transfected with WT GALC, Krabbe mutation N279T or double mutant N279T/N277D were lysed with 1% SDS, separated by SDS-PAGE and immunoblotted. Total GALC expression was analysed by Western blot against the FLAG epitope tag. C) HeLa cells were transiently transfected with Krabbe mutation N279T, the double mutation N279T/N277D or N277D alone, plated and fixed on glass coverslips and immunostained for GALC (green) and the lysosomal marker cathepsin D (red). Scale bar 10  $\mu$ m. D) The newly glycosylated residue, N277, is highlighted on the structure of GALC (PDB ID: 3ZR5). The structure is coloured according to domain (TIM barrel in blue,  $\beta$ -sandwich in red and lectin domain in green). Zoomed view (inset) shows the N277 sidechain as sticks. E) Conditioned media from either untransfected HEK293T cells (-) or cells transfected with WT or Krabbe mutation N279T (+) was harvested after 72 hours and applied to untransfected HEK293T cells for 48 hours. Treated cells were subjected to immunoprecipitation using the GALC monoclonal antibody mAb2. GALC was detected by Western blot using a polyclonal rabbit GALC antibody.

**Figure 6. Effect of the I546T polymorphism on enzyme activity and stability.** A) Michaelis–Menten plots of initial velocity vs. substrate concentration at GALC concentrations of 0.420 nM ( $\circ$ ), 0.842, ( $\diamond$ ), 1.68, ( $\blacktriangledown$ ), 3.37 nM, ( $\blacktriangle$ ), 6.73 nM, ( $\square$ ), and 13.4 nM ( $\bullet$ ). (Inset) Plot of  $V_{\max}$  vs. GALC concentration showing  $k_{\text{cat}}$  as the gradient (11.9 s $^{-1}$  for I546 and 10.6 s $^{-1}$  for T546). Experiments were performed in triplicate and SEM error bars are shown. B) Differential scanning fluorimetry of GALC polymorphisms I546 ( $\blacksquare$ ) and T546 ( $\circ$ ) demonstrate similar thermal denaturation profiles.



**Figure 7. Polymorphic background alters the trafficking of some missense mutations.** A) HEK293T cells were transiently transfected with WT GALC (I546), T546, Y551S-I546, Y551S-T546, G270D-I546, G270D-T546. Conditioned media was harvested at 72 hours and cells lysed in 1% SDS followed by SDS-PAGE and immunoblotting. Secreted and total GALC expression was detected by Western blot against the FLAG epitope tag and loading analysed by blotting against actin. B) Representative confocal microscopy images of HeLa cells transiently transfected with WT GALC containing the T546 polymorphism. Cells were plated onto glass coverslips, fixed and immunostained for GALC (green), the lysosomal marker cathepsin D (red). Nuclei were stained with DNA-binding dye, DAPI (blue). Scale bar 10  $\mu$ m. To quantify co-localization Pearson's correlation coefficients were calculated for cathepsin D. Mean  $\pm$  SEM for at least 20 individual cells from  $\geq 3$  independent experiments are shown. C) Equivalent confocal microscopy images as in panel B for Y551S-I546, Y551S-T546, G270D-I546 and G270D-T546. D) The positions of I546T, G270D and Y551S are highlighted on the structure of GALC (PDB ID: 3ZR5). The structure is coloured as in Fig 5. For each variant, the zoomed view (inset) shows the relevant side chain as sticks and the surrounding region of the structure that would be affected by the mutation. The single disulphide bond (yellow) present in GALC is illustrated as spheres.

**SUPPORTING INFORMATION**

Supplementary Table S1: Missense mutations of GALC described in the text.

This table is first associated with Figure 1 but is relevant to the whole paper and describes the alternative numbering used in the literature allowing for greater clarity when comparing with other published data using this numbering scheme.

Supplementary Table S2: Enzyme activity of WT GALC I546 and the T546 polymorphism.

This table is associated with Figure 6 and provides the details of the enzyme activity values  $K_m$ ,  $V_{max}$  and  $k_{cat}$  calculated from the data in Fig. 6.

Supplementary Table S3: Proportion of side chain buried in the structure, prediction and identified defect

This table is associated with the Discussion and summarises the correlation between the buried surface area of a side chain as calculated from the structure and the predicted effect this would have on the protein compared with the molecular mechanism identified in this work.

Supplementary Figure S1: Co-localization of T513M and L618S missense mutations with additional ER markers.

This figure provides additional representative confocal images of T513M and L618S co-localized with the ER markers calnexin and REEP5. These images are associated with Figure 3.

Supplementary Figure S2: Co-localization of additional missense mutations of GALC with ER markers.

This figure provides representative confocal images of additional missense mutations I583S, R515H and Y319C showing co-localization with the ER markers calreticulin and calnexin. These images are associated with Figure 3.

Supplementary Figure S3: Effect of polymorphic background on trafficking of G270D and Y551S.

This figure provides additional representative confocal images of G270D and Y551S in both polymorphic backgrounds co-localized with the ER marker REEP5. These images are associated with Figure 7.

Supplementary Figure S4. Representative confocal images from untransfected HeLa cells and cells expressing N-terminally tagged GALC constructs.

This figure provides evidence that endogenous GALC is not detectable in untransfected HeLa cells and that a construct of GALC possessing a tag at the N-terminus displays similar cellular localization as the C-terminally tagged construct. This figure is associated with the Materials and Methods section.

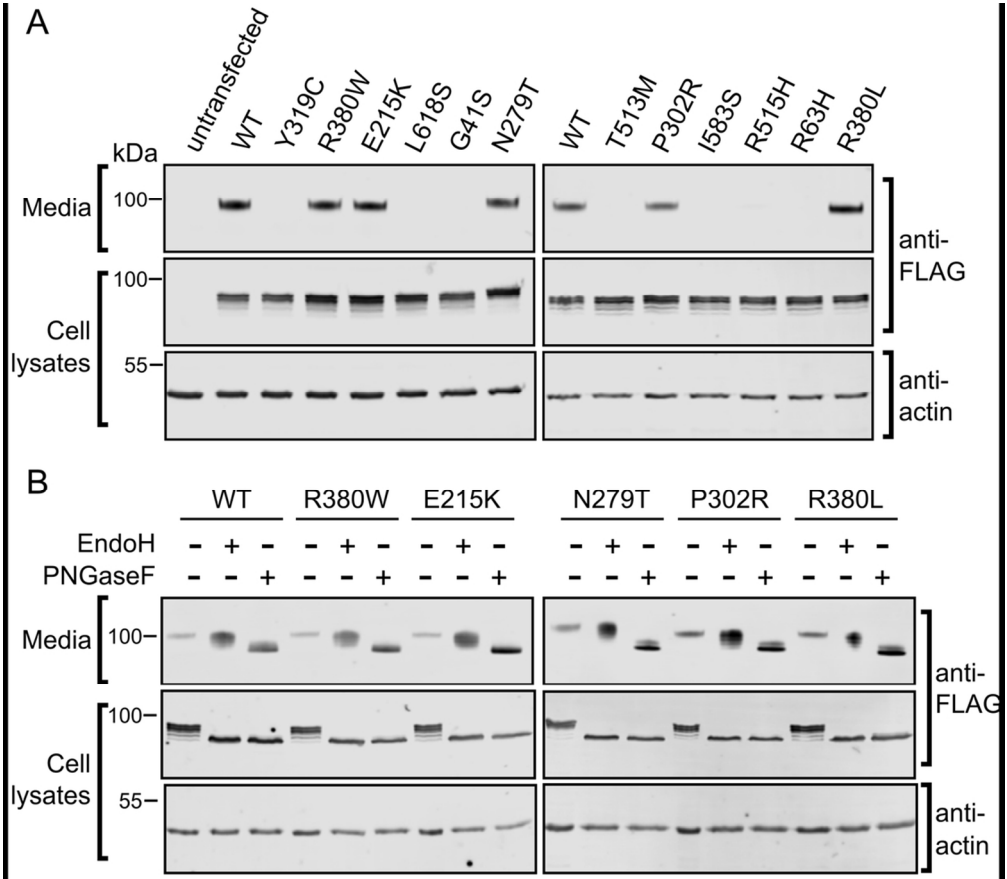


Figure 1. Missense mutations alter GALC secretion from cells. HEK293T cells were transiently transfected with FLAG-tagged wild-type (WT) GALC and a panel of clinically relevant Krabbe disease mutations. Conditioned media was harvested at 72 hours and cells lysed in 1% SDS followed by SDS-PAGE and immunoblotting. A) Secreted and total GALC expression were detected by Western blot using an anti-FLAG antibody. Actin served as a loading control. B) Conditioned media and cell lysates were digested with either EndoH or with PNGaseF before immunoblotting with anti-FLAG to detect GALC.

104x91mm (300 x 300 DPI)

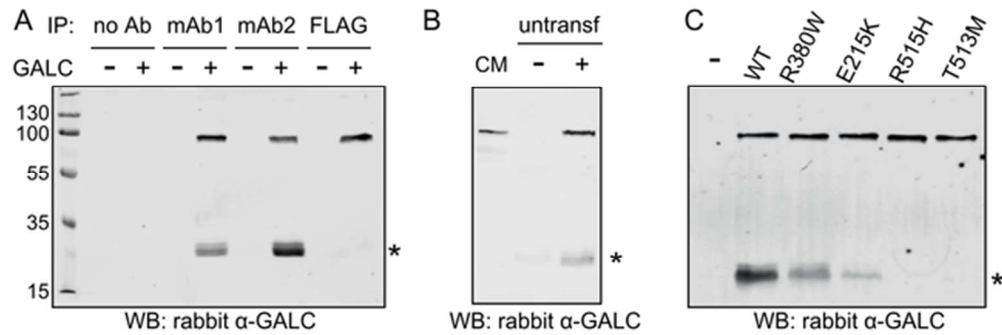


Figure 2. Proteolytic cleavage of WT and mutant GALC. A) Protein lysates from either untransfected HEK293T cells or from cells transfected with WT GALC were subjected to immunoprecipitation with monoclonal antibodies against full-length human GALC (mAb1 and mAb2) or the FLAG epitope tag. GALC was detected by Western blot using a polyclonal rabbit GALC antibody, cleavage product at ~30-kDa is marked with an asterisk. B) Conditioned media (CM) from HEK293T cells transfected with WT GALC contains full-length GALC (lane 1). GALC-containing CM (+) and CM from untransfected cells (-) was harvested after 72 hours and applied to untransfected HEK293T cells for 48 hours. Treated cells were subjected to immunoprecipitation using the GALC monoclonal antibody mAb2 and detected as in A. C) HEK293T cells were transiently transfected with WT GALC and Krabbe mutations R380W, E215K, R515H and T513M and immunoprecipitated using the mAb2 antibody and detected as in A.  
 54x18mm (300 x 300 DPI)

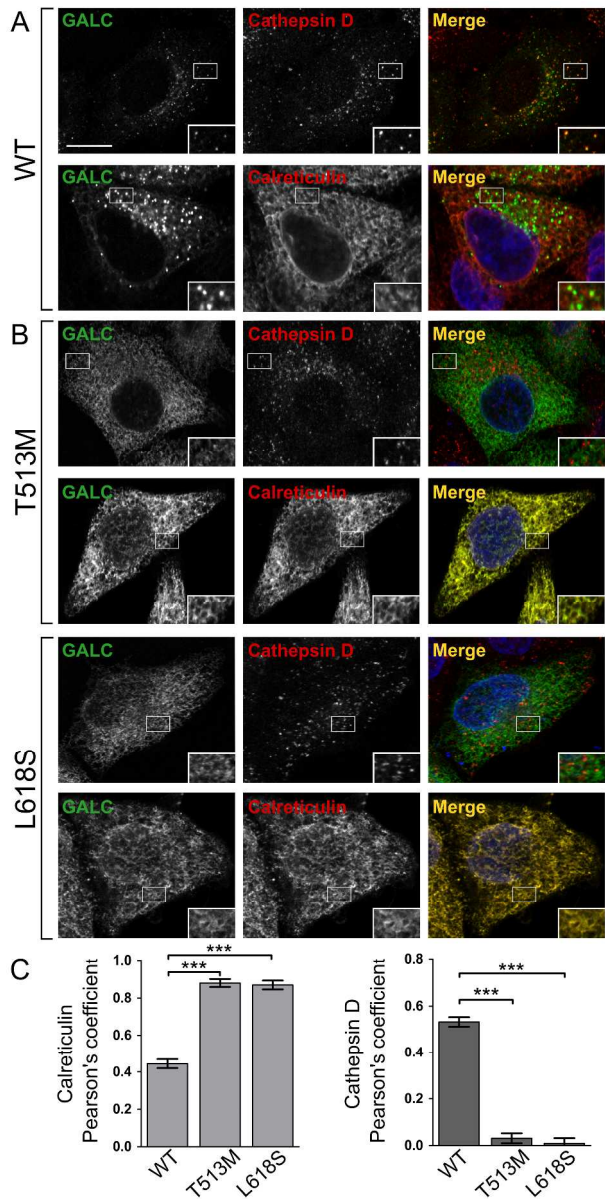


Figure 3. Missense mutations of GALC that interfere with secretion and cleavage are trapped in the ER. Representative confocal microscopy images of HeLa cells transiently transfected with either WT GALC (A), or the Krabbe mutations T513M and L618S (B). Cells were plated onto glass coverslips, fixed and immunostained using monoclonal antibody against GALC (green), the lysosomal marker cathepsin D (red) or the ER marker calreticulin (red). Nuclei were stained with DNA-binding dye, DAPI (blue). Scale bar 10  $\mu$ m. (C) To quantify co-localization, Pearson's correlation coefficients were calculated for each variant with both cathepsin D and calreticulin. Mean  $\pm$  SEM for at least 20 individual cells from  $\geq 3$  independent experiments are shown. \*\*\* $p$  = < 0.0001, calculated using a two-tailed unpaired t-test.

210x400mm (600 x 600 DPI)



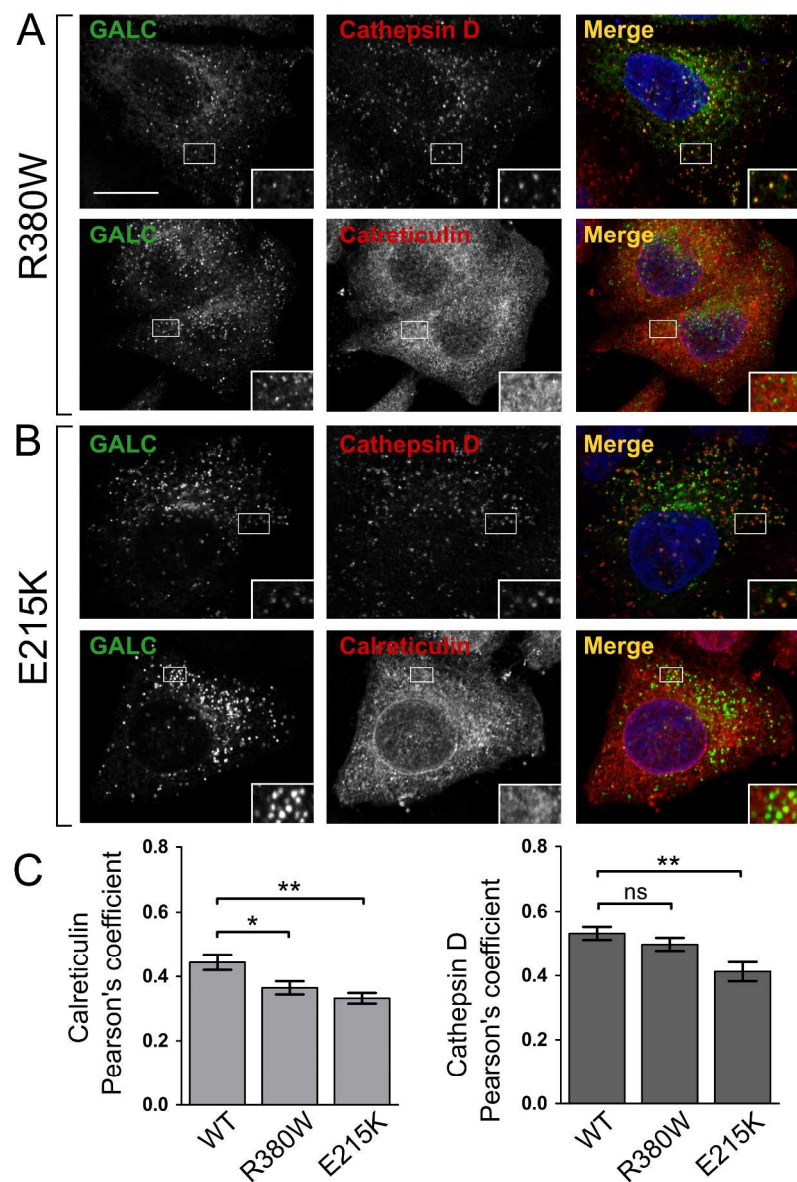


Figure 4. Missense mutations R380W and E215K are trafficked to the lysosome. Representative confocal microscopy images of HeLa cells transiently transfected with Krabbe mutations R380W (A) and E215K (B). Transfected cells were plated onto glass coverslips, fixed and immunostained for GALC (green), the lysosomal marker cathepsin D (red) or the ER marker calreticulin (red). Scale bar 10  $\mu$ m. (C) Pearson's correlation coefficients were calculated for each of the mutations with both cathepsin D and calreticulin. Mean  $\pm$  SEM for at least 20 individual cells from  $\geq 3$  independent experiments are shown. \*\*p = <0.005, \*p = <0.05, ns = not significant, calculated using a two-tailed unpaired t-test.

175x255mm (600 x 600 DPI)

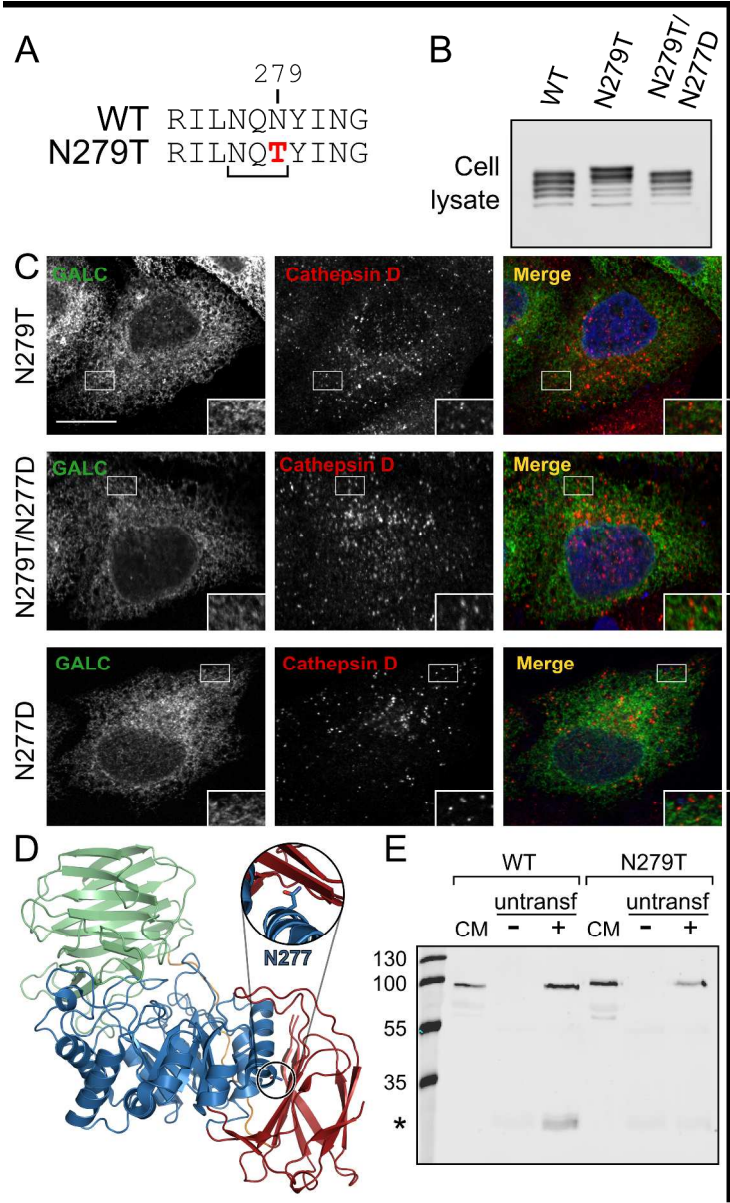


Figure 5. N279T is a hyperglycosylation mutant with defective trafficking. A) Protein sequence of WT and N279T GALC illustrating the introduction of a new consensus glycosylation site (NXT). B) HEK293T cells transiently transfected with WT GALC, Krabbe mutation N279T or double mutant N279T/N277D were lysed with 1% SDS, separated by SDS-PAGE and immunoblotted. Total GALC expression was analysed by Western blot against the FLAG epitope tag. C) HeLa cells were transiently transfected with Krabbe mutation N279T, the double mutation N279T/N277D or N277D alone, plated and fixed on glass coverslips and immunostained for GALC (green) and the lysosomal marker cathepsin D (red). Scale bar 10  $\mu$ m. D) The newly glycosylated residue, N277, is highlighted on the structure of GALC (PDB ID: 3ZR5). The structure is coloured according to domain (TIM barrel in blue,  $\beta$ -sandwich in red and lectin domain in green). Zoomed view (inset) shows the N277 sidechain as sticks. E) Conditioned media from either untransfected HEK293T cells (-) or cells transfected with WT or Krabbe mutation N279T (+) was harvested after 72 hours and applied to untransfected HEK293T cells for 48 hours. Treated cells were subjected to immunoprecipitation using the GALC monoclonal antibody mAb2. GALC was detected by Western blot using a polyclonal rabbit GALC

antibody.  
197x326mm (600 x 600 DPI)

For Peer Review

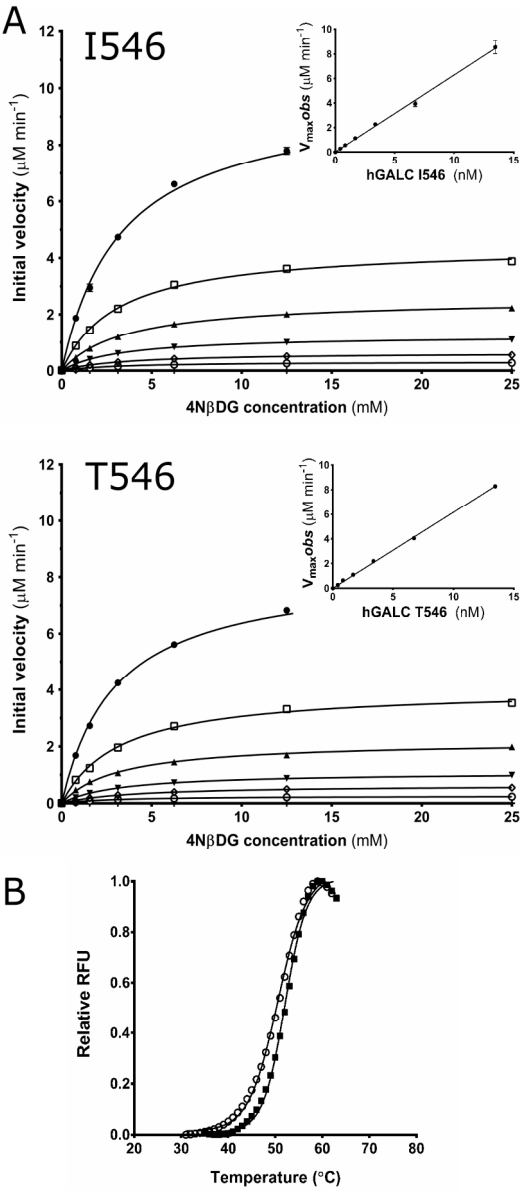


Figure 6. Effect of the I546T polymorphism on enzyme activity and stability. A) Michaelis–Menten plots of initial velocity vs. substrate concentration at GALC concentrations of 0.420 nM ( $\circ$ ), 0.842, ( $\diamond$ ), 1.68, ( $\blacktriangledown$ ), 3.37 nM, ( $\blacktriangle$ ), 6.73 nM, ( $\square$ ), and 13.4 nM ( $\bullet$ ). (Inset) Plot of  $V_{\text{max,obs}}$  vs. GALC concentration showing  $k_{\text{cat}}$  as the gradient (11.9 s $^{-1}$  for I546 and 10.6 s $^{-1}$  for T546). Experiments were performed in triplicate and SEM error bars are shown. B) Differential scanning fluorimetry of GALC polymorphisms I546 ( $\blacksquare$ ) and T546 ( $\circ$ ) demonstrate similar thermal denaturation profiles.

175x382mm (300 x 300 DPI)

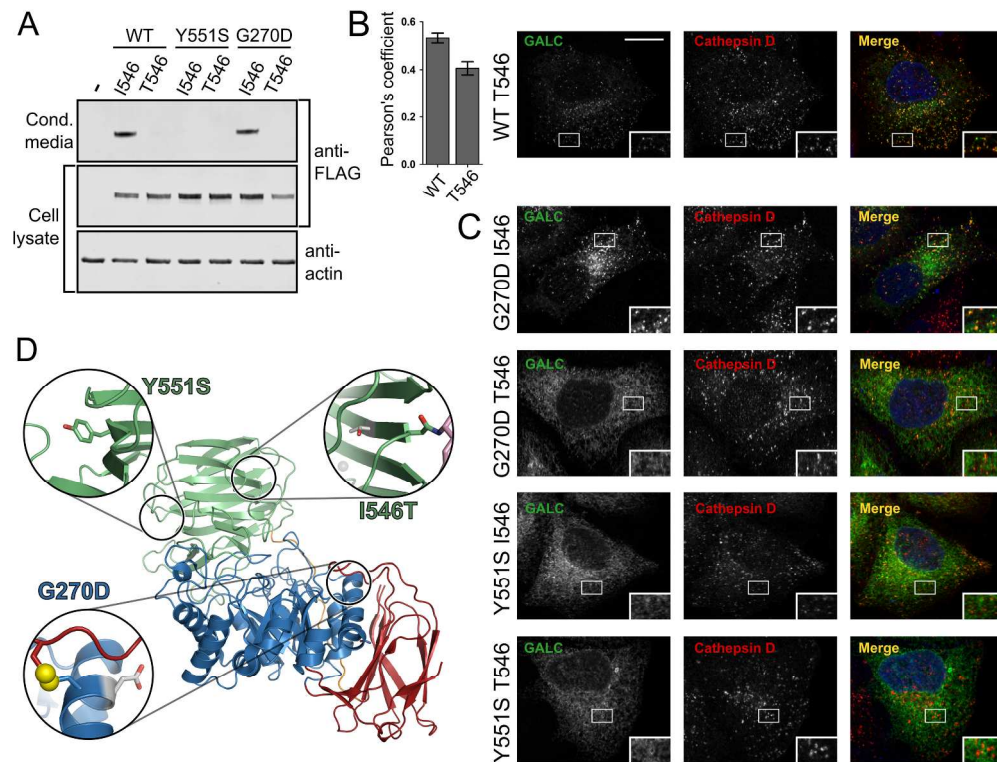


Figure 7. Polymorphic background alters the trafficking of some missense mutations. A) HEK293T cells were transiently transfected with WT GALC (I546), T546, Y551S-I546, Y551S-T546, G270D-I546, G270D-T546.

Conditioned media was harvested at 72 hours and cells lysed in 1% SDS followed by SDS-PAGE and immunoblotting. Secreted and total GALC expression was detected by Western blot against the FLAG epitope tag and loading analysed by blotting against actin. B) Representative confocal microscopy images of HeLa cells transiently transfected with WT GALC containing the T546 polymorphism. Cells were plated onto glass coverslips, fixed and immunostained for GALC (green), the lysosomal marker cathepsin D (red). Nuclei were stained with DNA-binding dye, DAPI (blue). Scale bar 10  $\mu$ m. To quantify co-localization Pearson's correlation coefficients were calculated for cathepsin D. Mean  $\pm$  SEM for at least 20 individual cells from  $\geq 3$  independent experiments are shown. C) Equivalent confocal microscopy images as in panel B for Y551S-I546, Y551S-T546, G270D-I546 and G270D-T546. D) The positions of I546T, G270D and Y551S are highlighted on the structure of GALC (PDB ID: 3ZR5). The structure is coloured as in Fig 5. For each variant, the zoomed view (inset) shows the relevant side chain as sticks and the surrounding region of the structure that would be affected by the mutation. The single disulphide bond (yellow) present in GALC is illustrated as spheres.

123x95mm (600 x 600 DPI)



SUPPORTING INFORMATION

Supplementary Table S1. Missense mutations of GALC described in the text

Mutation	Alternate numbering *
G41S	G57S
R63H	R79H
E215K	E231K
G270D	G286D
N279T	N295T
P302R	P318R
Y319C	Y335C
R380W/L	R396W/L
T513M	T529M
R515H	R531H
I546T §	I562T
Y551S	Y567S
I583S	I599S
L618S	L634S

\* The first start site in exon 1 is used for numbering in Uniprot entry P54803 and is has also been adopted by some publications. For clarity, these designations are also provided.

§ Polymorphism, also annotated as 1637T>C

**Table S2. Enzyme activity of WT GALC I546 and the T546 polymorphism**

Enzyme	$K_m$ (mM)	$V_{max}$ (nmol min <sup>-1</sup> μg <sup>-1</sup> )	$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_m$ (s <sup>-1</sup> M <sup>-1</sup> )
GALC – I546	3.3 ± 0.06	9.87 ± 0.2	11.9	3,600
GALC – T546	3.6 ± 0.13	8.74 ± 0.3	10.6	2,900

For Peer Review

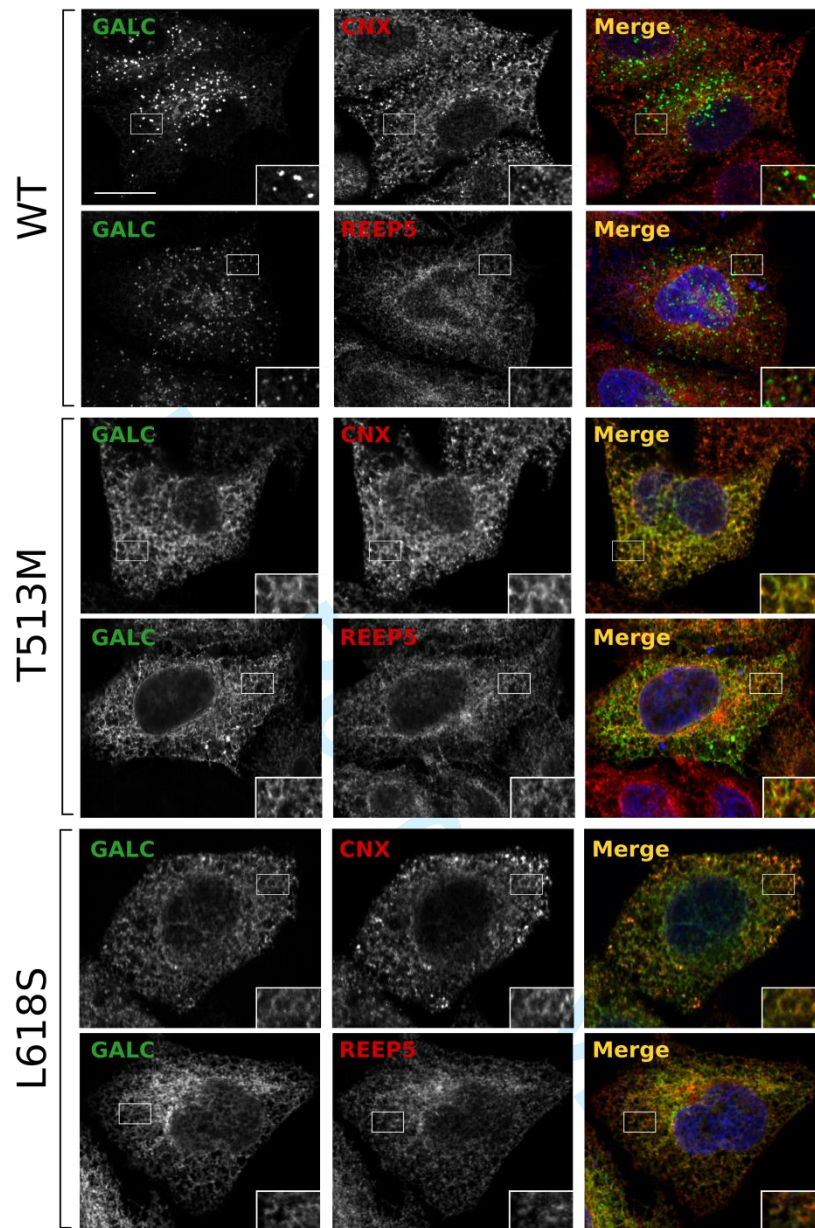
**Table S3. Proportion of side chain buried in the structure, prediction and identified defect**

Mutation	% Buried ‡	Structure-based Prediction	Secreted	Localization	Defect
G41S	100	Misfolded	No		Misfolded
G270D	100	Misfolded	I546 dependent	I546 dependent	Subtle misfolded
T513M	100	Misfolded	No	ER	Misfolded
I546T §	100	Misfolded	No	Lysosomal	Subtle misfolded
I583S	100	Misfolded	No	ER	Misfolded
L618S	100	Misfolded	No	ER	Misfolded
N279T	99	Misfolded	Yes	ER	Hyper-glycosylated Not endocytosed
Y319C	99	Misfolded	No	ER	Misfolded
P302R	92	Misfolded	Yes		Not misfolded
R515H	82	Uncertain	No	ER	Misfolded
R63H	79	Uncertain	No		Misfolded
Y551S	78	Uncertain	No	ER	Misfolded
E215K	57	Not misfolded	Yes	Lysosomal	Not misfolded
R380W/L	48	Not misfolded	Yes	Lysosomal	Not misfolded Catalytic defect

‡ Using the mouse GALC structure (PDB 3ZR5) the proportion of each sidechain that is buried in the structure is reported as a percentage of that residue's entire surface area

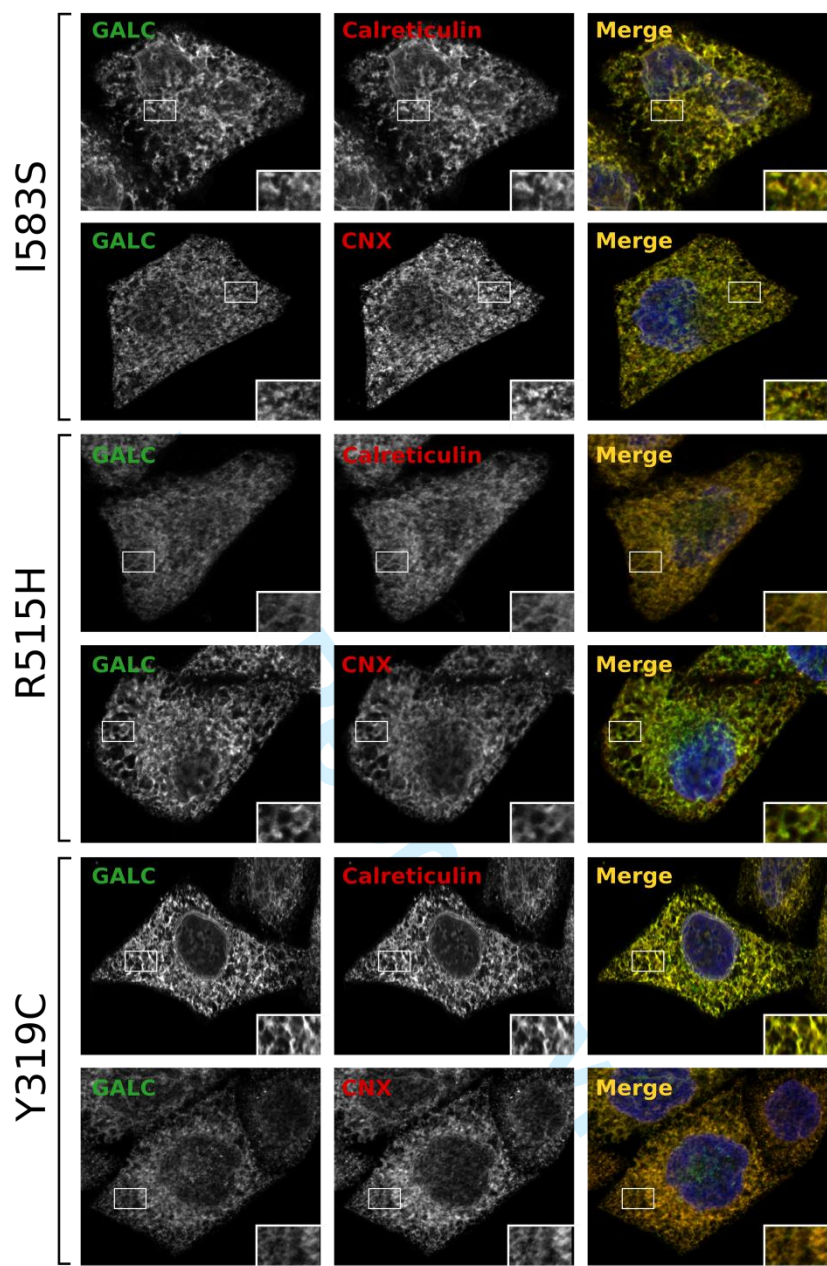
§ Polymorphism, not disease-causing

Supplementary Figure S1



**Figure S1. Co-localization of T513M and L618S missense mutations with additional ER markers.** Representative confocal images of HeLa cells transfected with wild-type and Krabbe disease mutations T513M and L618S. Cells were plated onto glass coverslips, fixed and immunostained using monoclonal antibody against GALC (green) and the ER markers calnexin (CNX, red) or REEP5 (red). Nuclei were stained with DNA-binding dye, DAPI (blue). Scale bar 10  $\mu$ m.

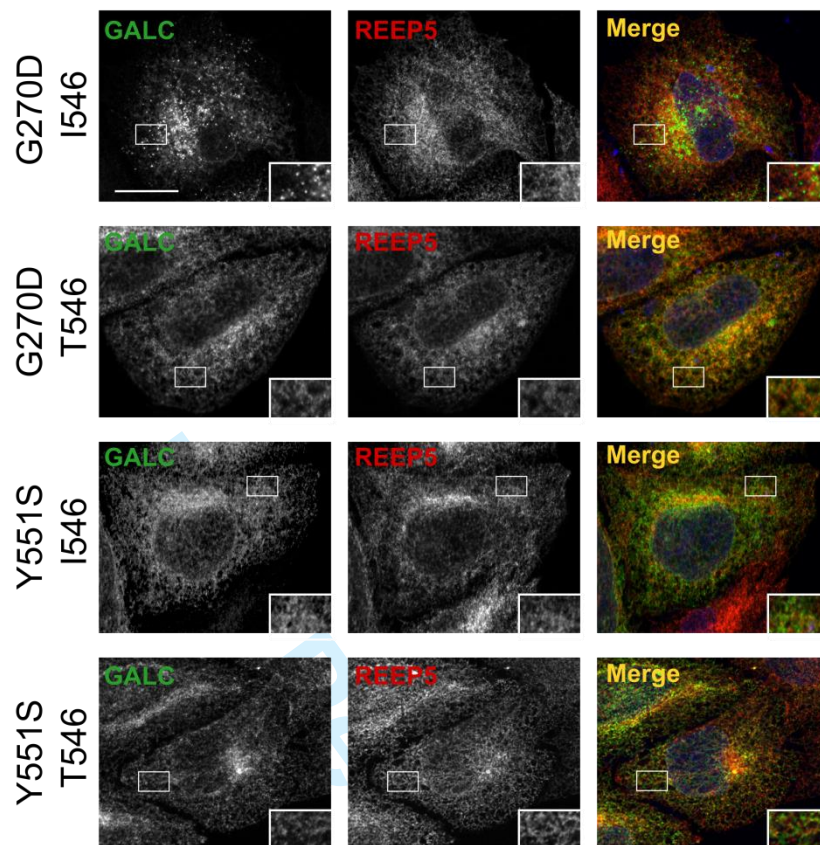
Supplementary Figure S2



**Figure S2. Co-localization of additional missense mutations of GALC with ER markers.** Representative confocal images of HeLa cells transfected with wild-type and Krabbe disease mutations I583S, R515H and Y319C. Cells were plated onto glass coverslips, fixed and immunostained using monoclonal antibody against GALC (green) and the ER markers calreticulin (red) or calnexin (CNX, red). Nuclei were stained with DNA-binding dye, DAPI (blue). Scale bar 10  $\mu$ m.

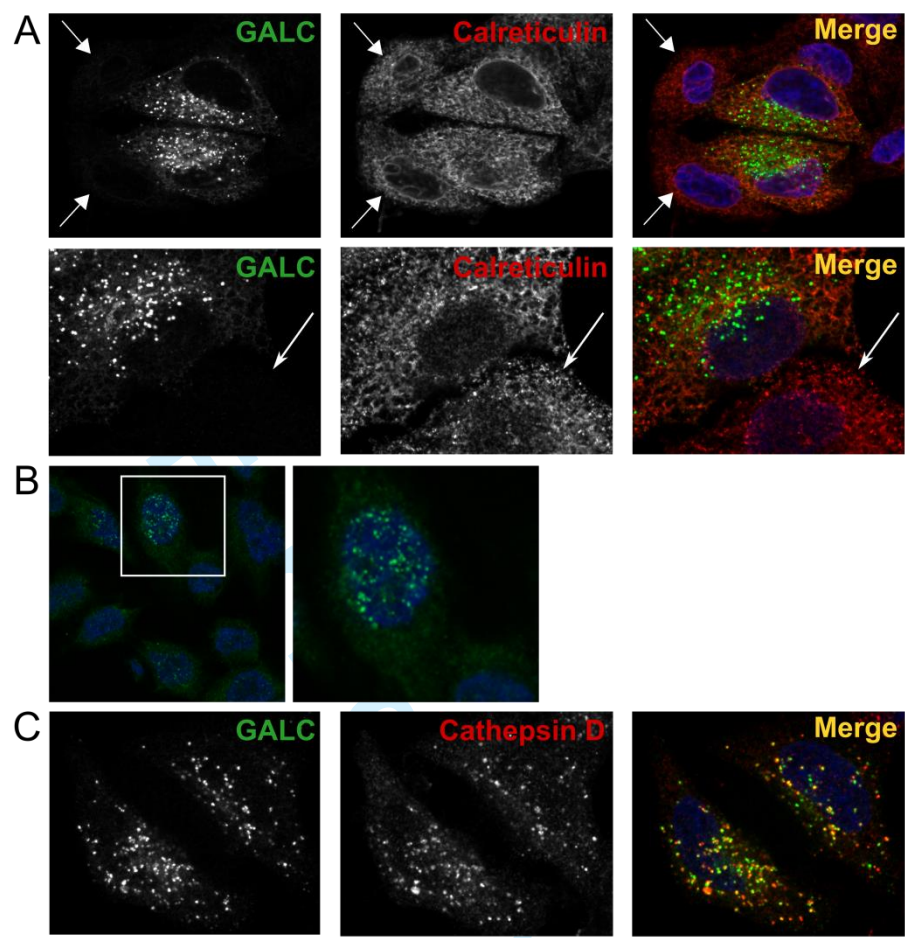


Supplementary Figure S3

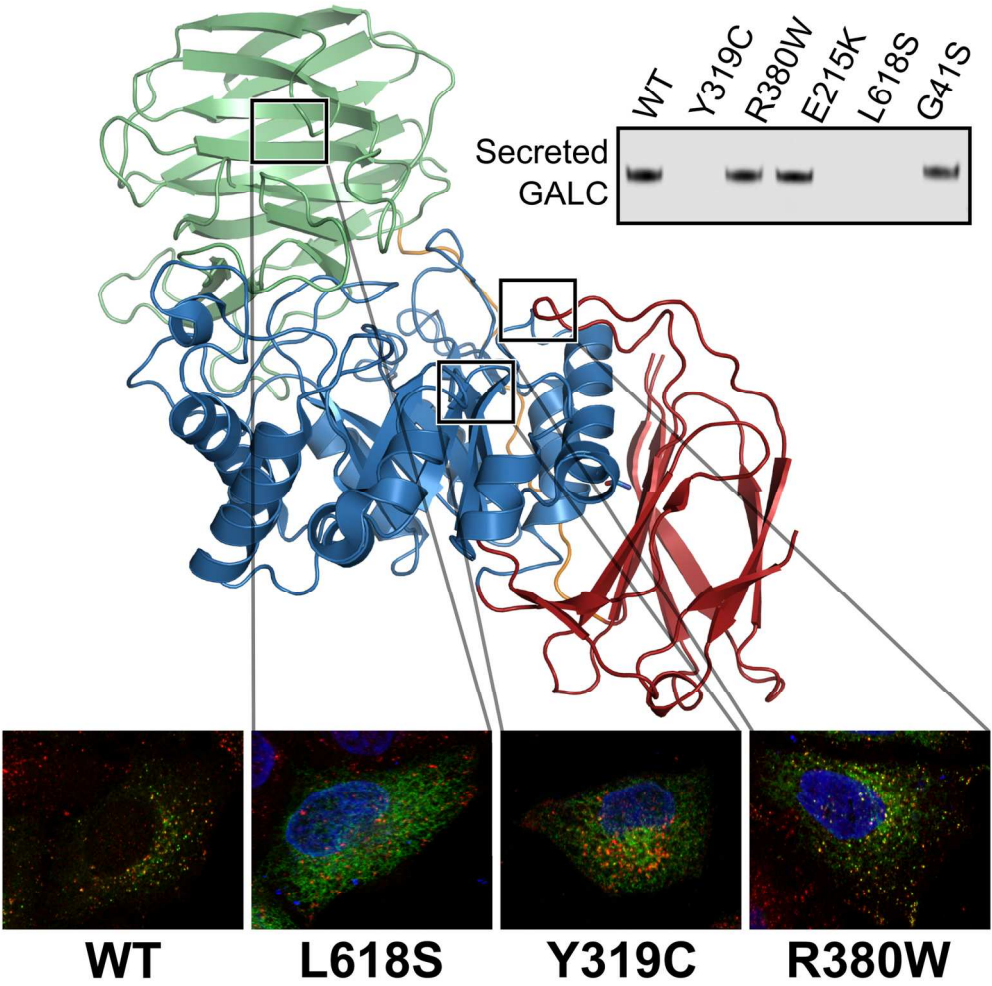


**Figure S3. Effect of polymorphic background on trafficking of G270D and Y551S.** Representative confocal images of HeLa cells transfected with Krabbe mutations G270D and Y551S in polymorphic backgrounds I546 and T546. Cells were plated onto glass coverslips, fixed and immunostained for GALC (green), the ER marker REEP5 (red). Nuclei were stained with DNA-binding dye, DAPI (blue). Scale bar 10 μm.

Supplementary Figure S4



**Figure S4. Representative confocal images from untransfected HeLa cells and cells expressing N-terminally tagged GALC constructs.** A) Examples of confocal images containing transfected and untransfected cells illustrating undetectable endogenous GALC staining. Cells were plated onto glass coverslips, fixed and immunostained using monoclonal antibody against GALC (green) and the ER marker calreticulin (red). B) Representative confocal image of untransfected cells with data collected at high laser intensity to illustrate non-specific staining only. C) Representative confocal image of HeLa cells transfected with an N-terminally tagged construct of GALC to illustrate equivalent co-localization with the lysosomal marker cathepsin D (red).



80x80mm (600 x 600 DPI)