Fabry disease: Identification of 50 novel α-galactosidase A mutations causing the classic phenotype and three-dimensional structural analysis of 29 missense mutations

*Junaid Shabbeer,[1] Makiko Yasuda,[1] Stacy D. Benson[2] and Robert J. Desnick[1]**

[1] Department of Human Genetics, Mount Sinai School of Medicine of New York University, New York, NY 10029, USA
[2] Department of Chemistry, Oklahoma State University, Stillwater, OK 74078, USA
* *Correspondence to*: Tel: +1 212 659 6700; Fax: +1 212 360 1809; E-mail: robert.desnick@mssm.edu

Date received (in revised form): 16th December 2005

Abstract

Fabry disease, an X-linked recessive inborn error of glycosphingolipid catabolism, results from the deficient activity of the lysosomal exoglycohydrolase, α-galactosidase A (EC 3.2.1.22; α-Gal A). The molecular lesions in the *α-Gal A* gene causing the classic phenotype of Fabry disease in 66 unrelated families were determined. In 49 families, 50 new mutations were identified, including: 29 missense mutations (N34K, T41I, D93V, R112S, L166G, G171D, M187T, S201Y, S201F, D234E, W236R, D264Y, M267R, V269M, G271S, G271V, S276G, Q283P, A285P, A285D, M290I, P293T, Q312H, Q321R, G328V, E338K, A348P, E358A, Q386P); nine nonsense mutations (C56X, E79X, K127X, Y151X, Y173X, L177X, W262X, Q306X, E338X); five splicing defects (IVS4-1G>A, IVS5-2A>G, IVS5 + 3A>G, IVS5 + 4A>G, IVS6-1G>C); four small deletions (18delA, 457delGAC, 567delG, 1096delACCAT); one small insertion (996insC); one 3.1 kilobase Alu-Alu deletion (which included exon 2); and one complex mutation (K374R, 1124delGAG). In 18 families, 17 previously reported mutations were identified, with R112C occurring in two families. In two classically affected families, affected males were identified with two mutations: one with two novel mutations, D264Y and V269M and the other with one novel (Q312H) and one previously reported (A143T) mutation. Transient expression of the individual mutations revealed that D264Y and Q312H were localised in the endoplasmic reticulum and had no detectable or markedly reduced activity, whereas V269M and A143T were localised in lysosomes and had approximately 10 per cent and approximately 35 per cent of expressed wild-type activity, respectively. Structural analyses based on the enzyme's three-dimensional structure predicted the effect of the 29 novel missense mutations on the mutant glycoprotein's structure. Of note, three novel mutations (approximately 10 per cent) were predicted not to significantly alter the glycoprotein's structure; however, they were disease causing. These studies further define the molecular heterogeneity of the *α-Gal A* mutations in classical Fabry disease, permit precise heterozygote detection and prenatal diagnosis, and provide insights into the structural alterations of the mutant enzymes that cause the classic phenotype.

Keywords: *Fabry disease, α-galactosidase A, α-Gal A, mutations, molecular modelling, transient expression*

Introduction

Fabry disease is an X-linked recessive inborn error of glycosphingolipid catabolism resulting from the deficient activity of the lysosomal exoglycohydrolase, α-galactosidase A (EC 3.2.1.22; α-Gal A).[1] The enzymatic defect causes the progressive accumulation of globotriaosylceramide (GL-3) and related glycosphingolipids with terminal α-linked galactosyl moieties in the plasma and in tissue lysosomes throughout the body. In classically affected males who have little, if any, α-Gal A activity, onset of disease manifestations occurs in childhood or adolescence and is characterised by severe acroparesthesias, angiokeratoma, corneal and lenticular opacities and hypohidrosis. With advancing age, the progressive glycosphingolipid deposition, particularly in vascular endothelial lysosomes, leads to vascular disease of the heart, kidneys and brain, resulting in early demise, typically in the fourth or fifth decade of life. In contrast to the classic phenotype, later-onset variants have been identified which have residual α-Gal A activity. These patients typically present

with cardiac and/or renal disease and lack the major classical manifestations, including angiokeratoma, acroparesthesias, hypohidrosis and ocular abnormalities.[2–5]

To date, a variety of mutations have been identified which cause the classic phenotype, including missense, nonsense and splice-site mutations, as well as partial gene rearrangements, including small and large intragenic deletions and insertions[1,6,7] (the Human Gene Mutation Database). Most mutations have been private (ie unique to one or a few families), with the exception of certain mutations found in unrelated individuals that occurred at CpG dinucleotides, known hot spots for mutation.[8,9] Of note, non-coding sequence variants have been identified in *α-Gal A* alleles from normal individuals and patients with Fabry disease[10,11] (the SNP database), with the exception of *D313Y*, an exonic sequence variant that encodes an enzyme with 60–70 per cent of wild-type activity, but does not cause disease in males or females with this variant.[12,13]

Mutation detection in Fabry disease is important for several reasons. First, heterozygote detection by enzyme assay of carriers for this X-linked recessive disease is unreliable because obligate carriers can have normal activity due to random X-chromosomal inactivation.[14–16] Secondly, safe and effective treatment for Fabry disease by α-Gal A replacement therapy has recently become available worldwide.[17–20] Because affected males should be treated early to prevent the serious complications of the disease, and because carriers may be asymptomatic or have very mild manifestations,[1,21,22] it is important to identify affected males and carriers for medical monitoring and early treatment.[23] Thirdly, recent *in vitro* studies[24] have demonstrated that certain α-Gal A missense mutations result in mutant proteins that are misfolded and degraded in the endoplasmic reticulum (ER). Some of these misfolded mutant proteins, particularly those that have residual enzymatic activity (>1 per cent), have been rescued by pharmaceutical chaperones, such as galactose and deoxygalactonojirimycin, which are reversible competitive inhibitors of α-Gal A.[25–28] In fact, intravenous administration of galactose (1 g/kg) to a 56-year-old male with the cardiac variant phenotype resulted in marked improvement of his cardiac manifestations.[29] Thus, it is of interest to predict which of the missense mutations that cause the classic Fabry phenotype might be rescuable based on the enzyme's 3D structure.[30,31] These mutations can then be studied *in vitro* to determine their rescuability.

In this paper, 50 new and 17 previously reported *α-Gal A* mutations were identified in 66 unrelated classically affected families. Two unrelated patients with the classic phenotype had mutant alleles with two missense mutations, each of which was expressed *in vitro*. In addition, the structural alterations resulting from the 29 novel missense mutations were predicted based on the 3D structure of the wild-type enzyme homodimeric glycoprotein. Several of these mutations resulted from misfolding and are candidates to determine their rescuability by pharmacological chaperones.

Materials and methods

Patient specimens

Peripheral blood was collected from the probands of 66 unrelated families with the classic phenotype of Fabry disease. The α-Gal A activity was determined in the plasma and/or lymphocytes as previously described.[32] Genomic DNA was extracted using the Puregene isolation kit according to manufacturer's instructions (Gentra Systems, Minneapolis, MN, USA). All specimens were obtained with informed consent and the approval of the Institutional Review Board of the Mount Sinai School of Medicine of New York University.

Mutation analysis

Mutation analysis was performed as previously described.[33] Briefly, each of the *α-Gal A* exons and adjacent flanking and/or intronic sequences was amplified by means of polymerase chain reaction (PCR) from genomic DNA. Each amplicon was then analysed by denaturing high-performance liquid chromatography, and abnormally running fragments were sequenced using an ABI Prism 3700 Capillary Array Sequencer with the ABI Prism™ BigDye™ Terminator Ready Reaction Mix (Perkin-Elmer-Cetus, Norwalk, CT, USA). Each mutation was confirmed by repeat PCR amplification and sequencing of the opposite strand, and/or by co-segregation of the lesion and disease phenotype in other members of each family. In addition, 100 normal chromosomes were examined to rule out a polymorphism for each missense mutation.

Southern blot analysis

Genomic DNA (10 μg), extracted from normal controls and the proband, was digested with 100 units (U) of HindIII and PvuII for 16 hours. After electrophoresis on a 1 per cent agarose gel, DNA was transferred onto a Hybond-N + membrane (Amersham GE Healthcare, Piscataway, NJ, USA) in 0.4M NaOH solution using standard procedures. The filter was ultraviolet-irradiated and then hybridised with a randomly primed ^{32}P-labelled PCR-generated probe spanning exons 1 and 2 of the *α-Gal A* gene. After overnight hybridisation using PerfectHyb Plus hybridisation buffer (Sigma Aldrich), membranes were washed in diluted standard saline citrate buffers with 0.1 per cent sodium dodecyl sulphate. Bands were visualised after exposure and analysed on a Molecular Dynamics STORM 860 phosphoimager (GE Healthcare).

Microsatellite studies

The two probands with the R112C missense mutation were studied to determine if they were related or if the mutations occurred independently. Their genomic DNAs were haplotyped with the microsatellite markers close to the *α-Gal A* locus, including DXS1231, DXS8020, DXS8034, DXS8089,

DXS8100, DXS8063 and DXS8096. Forward primers were fluorescent dye-labelled (Invitrogen Life Technologies, Carlsbad, CA, USA). Genomic DNA was amplified in 10 μl volumes with 50 ng of genomic DNA, 2 mM $MgCl_2$, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 200 nM of each primer, 0.2 mM deoxyribonucleoside triphosphates and 2 U *Taq* DNA polymerase (AmpliTaq Gold, Applied Biosystems, Foster City, CA, USA). The reaction mixtures were initially incubated at 95 °C for 10 minutes, and then underwent 27 cycles of amplifications with denaturation at 94 °C for 30 seconds, annealing at 58 °C for 30 seconds, extension at 72 °C for 30 seconds and a final extension step at 72 °C for 7 minutes. Microsatellites were analysed with an ABI Prism 3100 Genetic Analyzer using GeneScan Analysis Software (Version 3.1.2) and Genotyper Software (Version 2.5) (Perkin-Elmer-Cetus, Norwalk, CT, USA).

Conservation of missense mutations

Each of the missense lesions was analysed to determine the relative conservation of the substituted amino acid by comparison with four mammalian and 22 non-mammalian eukaryotic α-Gal A orthologues and eight α-N-acetylgalactosaminidase (α-Gal B) orthologues in the GenBank database. These searches were performed using the MacVector program (Oxford Molecular Group). Highly conserved residues were defined as those that were present in three of the four (75 per cent) mammalian orthologues in at least 17 (77 per cent) of the eukaryotic orthologues and in six (75 per cent) of the eukaryotic orthologues of the related *α-Gal B* gene.[34]

In vitro expression studies

The full-length wild-type *α-Gal A* cDNA was cloned into the pAsc8 vector.[35] Mutant constructs, carrying the individual D264Y, V269M, Q312H or A143T mutations or the double mutations (D264Y/V269M and A143T/Q312H), were generated by site-directed mutagenesis PCR (Stratagene, La Jolla, CA, USA) using the primers: GGACCAGGGGGTTG-GAATT ACCCAGATATGTTAGTG (D264Y sense), CAC-TAACATATCTGGGTAATTCCAACCC CCTGGTCC (D264Y antisense), GACCCAGATATGTTAATGATTGGC-AACTTTGG (V269M sense), CCAAAGTTGCCA-ATCATTAACATATCTGGGTC (V269M antisense), GCTGGAAATAAAACCTGCACAGGCTTCCCTGGGAG (A143T sense), CTCCCAGGGA AGCCTGTGCAGGTTT-TATTTCCAGC (A143T antisense), CAAGCCAAAGCTC-TCCTTC ATGATAAGGACGTAACTGCCA (Q312H sense) and TGGCAGTTACGTCCTTATCATG-AAGGAGAGCTTTGGCTTG (Q312H antisense).

All constructs were confirmed by re-sequencing and plasmid preparations were made using the Qiagen plasmid midi kit (Valencia, CA, USA). The wild-type and each mutant construct were individually transfected into COS-7 cells and analysed for intracellular α-Gal A activity and subcellular localisation, as previously described.[13]

Structural analysis of α-Gal A missense mutations

The 3.25 Å X-ray structure of human α-Gal A[30] was the basis of the structural analysis. Side-chain positions were compared with an independently derived human α-Gal A model based on the chicken α-Gal B X-ray structure.[36] Mutations were modelled and visualised in the program O[37] and energy minimised with the CNS program.[38]

Results

Mutation detection

Table 1 summarises the 50 novel and 17 previously reported mutations detected in 66 unrelated patients with classic Fabry disease. PCR amplification of the *α-Gal A* exons and adjacent intronic or flanking sequences from genomic DNA, and electrophoresis of the amplicons, did not reveal any gene rearrangements (>50 base pairs [bp]), except one in which exon 2 did not amplify. Southern blot analysis indicated a large deletion of ~3 kilobases (kb), which included exon 2. To determine the precise breakpoints of the deletion, several random primer pairs in introns 1 and 2 were used to amplify the proband's genomic DNA. Using a sense primer (GCTA-ATGGCAAGACCCTG) located at g2765 in intron 1 and an antisense primer (AAATCCCCCAGTTCTGCTGAGCTA) at g7218 in intron 2, an approximately 4.5 kb PCR fragment was expected; however, these primers amplified a 1.3 kb fragment. Sequencing of this PCR fragment identified the breakpoints in Alu repetitive sequences at g3260 and at g6410, resulting in an Alu–Alu rearrangement that deleted 3,152 bp, including the entire exon 2 (Figure 1).

In the remaining 65 unrelated probands, sequencing the respective *α-Gal A* amplicons detected single mutations in each, with the exception of two probands — one having two novel missense alterations (D264Y and V269M) and the other having one novel and one previously reported mutation (Q312H and A143T, respectively) in their respective *α-Gal* A alleles (Table 1). Also, one proband had a complex mutation resulting in a missense mutation K374R and an adjacent 3 bp deletion. The other novel lesions included 26 missense mutations (N34K, T41I, D93V, R112S, L166G, G171D, M187T, S201Y, S201F, D234E, W236R, M267R, G271S, G271V, S276G, Q283P, A285P, A285D, M290I, P293T, Q321R, G328V, E338K, A348P, E358A and Q386P), nine nonsense mutations (C56X, E79X, K127X, Y151X, Y173X, L177X, W262X, Q306X and E338X), four small deletions (18delA, 457delGAC, 567delG and 1096delACCAT), one small insertion (996insC) and five splice site mutations (IVS4-1G>A, IVS5-2A>G, IVS5 + 3A>G, IVS5 + 4A>G and IVS6-1G>C). Reverse transcriptase-PCR studies demonstrated that the IVS5 + 3A>G splice mutation resulted in two abnormal transcripts, one of which included 66 bp of intron 5, whereas the other included the entire intron 5. No normal transcript was detected.

Table 1. α-Gal A mutations in 66 unrelated probands with classic Fabry disease.

Exon	cDNA mutation	Predicted change	*α-Gal A* activity	
			Plasma	Leukocytes
Novel mutations:				
Missense				
1	102T > G	N34K	–	–
1	122C > A	T41I	6.6	1.0
2	278A > T	D93V	–	–
2	334C > A	R112S	1.7	0.3
3	496_497CT > GG	L166G	–	–
3	512G > A	G171D	0.4	0.3
4	560T > C	M187T	0.6	0.2
4	602C > A	S201Y	0.8	1.7
4	602C > T	S201F	0.2	0.1
5	702T > G	D234E	–	–
5	706T > C	W236R	–	0.4
5	790G > T	D264Y*	1.5	0.7
5	800T > G	M267R	2.8	0.9
6	805G > A	V269M*	–	–
6	811G > A	G271S	1.9	–
6	812G > T	G271V	0.7	0.2
6	826A > G	S276G	–	0.9
6	848 A > C	Q283P	–	–
6	853G > C	A285P	–	0.3
6	854C > A	A285D	0.2	0.2
6	870G > A	M290I	–	–
6	877C > A	P293T	3.2	0.8
6	936G > T	Q312H*	–	–
6	962A > G	Q321R	–	–
6	983G > T	G328V	–	0.1
7	1012G > A	E338K	–	–
7	1042G > C	A348P	–	–
7	1073A > C	E358A	0.6	0.9
7	1157A > C	Q386P	0.9	1.2

(*continued*)

Table 1. *Continued.*

Exon	cDNA mutation	Predicted change	*α-Gal A* activity	
			Plasma	Leukocytes
Nonsense				
1	168C > A	C56X	–	0.6
2	235G > T	E79X	–	–
3	379A > T	K127X	0.5	<0.1
3	453C > G	Y151X	0.1	0.2
3	519C > A	Y173X	–	–
3	530T > A	L177X	< 0.1	0.3
5	785G > A	W262X	–	–
6	916C > T	Q306X	0.3	0.6
7	1012G > T	E338X	0.4	0.1
Deletions				
1	18delA	E7 fs	0.7	0
2		Ex2del	–	0.1
3	457_459delGAC	D153del	–	–
4	567delG	A190fs	<0.1	0.1
7	1096_1100delACCAT	T366fs	4.6	0.8
Insertion				
6	996insC		–	–
Complex				
7	1121A > G,1124delGAG	K374R,G375 fs	–	–
Splicing defects				
IVS4 acceptor	640-1G > A	IVS4-1G > A	–	–
IVS5 donor	801+3A > G	IVS5+3A > G	0.5	0.2
IVS5 donor	801+4A > G	IVS5+4A > G	–	–
IVS5 acceptor	802-2A > G	IVS5-2A > G	–	–
IVS6 acceptor	1000-1G > C	IVS6-1G > C	–	–
Previously reported mutations:				
Missense				
1	92C > T	A31V	1.7	3.3
2	334C > T	R112C	0.8	1.0
3	401A > C	Y134S	<0.1	0.3
3	427G > A	A143T*	2.5	17.0

(*continued*)

Table 1. *Continued.*

Exon	cDNA mutation	Predicted change	*α-Gal A* activity	
			Plasma	Leukocytes
5	776C > G	P259R	0.5	1.4
6	983G > C	G328A	1.0	0.8
7	1241T > C	L414S	–	–
Nonsense				
3	456C > A	Y152X	0.6	–
3	469C > T	Q157X	0.1	0.5
5	661C > T	Q221X	–	–
5	677G > A	W226X	0.5	2.9
5	679C > T	R227X	0.4	–
7	1024C > T	R342X	0.2	–
Deletions				
1	26delA	H9 fs	–	–
7	1031_1032delTC	S345 fs	–	–
7	1209_1211delAAG	R404del	–	–
7	1235_1236delCT	T412 fs	1.6	0.6

*D264Y and V269M were present together in a single proband, as were Q312H and A143T. A143T has been reported alone in other Fabry patients with a late-onset phenotype.

The previously reported lesions included seven missense mutations (A31V, R112C, Y134S, A143T, P259R, G328A and L414S), six nonsense mutations (Y152X, Q157X, Q221X, W226X, R227X and R342X) and four small deletions (26delA, 1031delTC, 1209delAAG and 1235delCT). The R112C substitution was identified in two probands that were found to be unrelated from microsatellite studies. Of note, the R112S, R112C, Y152X, R227X and R342X mutations occurred at CpG dinucleotides, known mutational hot spots.[9] Although previously reported sequence variants were detected in *α-Gal A* alleles from normal individuals and patients with Fabry disease[10–13] (the SNP Database), no new non-pathological sequence variants were detected in this study.

Expression and subcellular location of the double missense mutations

To determine the functional effects of each substitution in the two affected males whose *α-Gal A* alleles had two missense lesions, expression studies in COS-7 cells were carried out (Table 2). The *V269M* allele had approximately 10 per cent of the mean expressed wild-type activity and the mutant enzyme protein was localised immunohistologically to the lysosomes, whereas the *D264Y* allele had no detectable activity and the mutant protein remained in the ER. The *A143T* allele had approximately 35 per cent of expressed wild-type activity and localised to the lysosome, whereas the *Q312H* allele had approximately 5 per cent of expressed wild-type activity and was detected predominantly in the ER. Constructs with the double

Intron 1: ·ACTCCACCTCCCGGGTTTAAGCAGTTCTCCTG*···TCGTAGTCTCCTGAGTAGCTGGGATTACAGGCACACC···g3291

Patient: ·ACTCCACCTCCCGGGTTTAAGCAGTTCTCCTG*·*CCTCAGCCTCCCAAGTAGCTGGGACTACAGGCGCACA···

Intron 2: ·CCTCCACCTCTTGGGTTCAAGTGATTCTCCTG···*CCTCAGCCTCCCAAGTAGCTGGGACTACAGGCGCACA···g6447

Figure 1. *α-Gal A* Alu-Alu recombination causing deletion of exon 2 in a classically affected male. Wild-type intron 1 and intron 2 sequences are shown above and below the rearrangement in the patient. Bases starred are at the breakpoint between the intron 1 and 2 sequences, while bases underlined indicate those that are identical in introns 1 and 2.

Table 2. *In vitro* expression of the double missense mutations in COS-7 cells.

α-Gal A mutation	Enzymatic activity*	Immunohistochemical localisation
V269M	10	Lysosomes
D264Y	<1	ER
V269M/D264Y	<1	ER
A143T	35	Lysosomes
Q312H	5	ER
A143T/Q312H	<1	ER

*Expressed as % of mean wildtype activity. Abbreviation: ER = endoplasmic reticulum.

mutations**,** D264Y/V269M or A143T/Q312H**,** expressed no detectable enzymatic activity and were retained in the ER, consistent with being detected in patients with the classic phenotype.

Molecular modelling of the α-Gal A missense mutations

Based on a refined model of the α-Gal A crystal structure (Benson, S.D. *et al.*, unpublished results), the locations and predicted structural alterations of each of the 29 novel missense mutations were determined (Figure 2 and Table 3). The degree of conservation of the residues involved in the novel missense mutations is shown in Table 4. **N34K:** Asparagine-34 is located in the N-terminal loop, where its side chain forms hydrogen bonds with the side-chain of asparagine-224. The substituted lysine is too large for this position and its positive charge could interfere with the proper glycosylation at asparagine-192. **T41I:** Threonine-41 is part of the first β-strand of the β/α-barrel, and the β-strands align the active site. The threonine side-chain points into a small pocket away from the active site. Although there is room for the side-chain of the substituted isoleucine, hydrogen bonding between threonine-41 and histidine-225 would be lost, most likely resulting in protein misfolding. **D93V:** Aspartate-93 is on the rim of the active site and provides the negative charge environment surrounding the galactosyl residue that will be cleaved. Although it is not one of the active aspartate residues, it interacts with the hydroxyl group at the C6 of the terminal galactose and orientates the substrate for cleavage. The substitution of a valine would destroy the interaction with the C6 hydroxyl group and, presumably, would prevent proper binding and orientation of the substrate for cleavage. **R112S:** Arginine-112 is located on the α2-helix of the N-terminal β/α-barrel, some distance from the active-site pocket. Its side-chain points toward the disulphide bond between cysteine-52 and cysteine-94, possibly stabilising the bond. A substitution to serine is predicted to destabilise both the disulphide bond and the pocket that the arginine occupies, leading to protein misfolding. **L166G:** Leucine-166 is on the β4-strand of the N-terminal β/α-barrel that lines the edge of the active-site pocket. A substitution of a glycine is predicted to enlarge the active-site pocket, probably causing lower specificity and efficiency because it would be more difficult to orientate the substrate. Glycine is also more flexible, adding to the inability of the active pocket to stabilise the substrate. **G171D:** Glycine-171 is located at the end of the β4-strand along the edge of the active site and next to the enzymatically active aspartate-170. Glycine is a small and flexible amino acid suit-

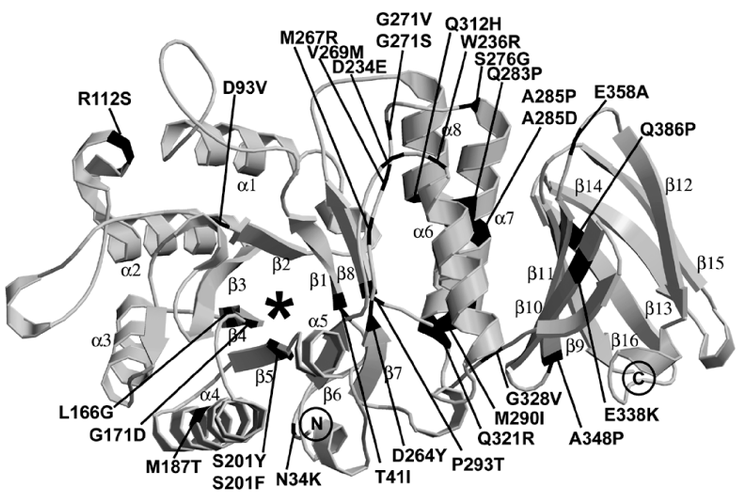

Figure 2. The structure of the human *α-Gal A* monomer based on its X-ray crystal structure. The N- and C-termini are labelled and the active site indicated by a star. The locations of the mutations are shown.

Table 3. Predicted structural alterations caused by the novel missense mutation.

Residue	Fabry mutation	Importance in *α-Gal A*	Residue/side-chain accessible surface (Å^2)	Solvent exposure of residue
N34	K	Disrupts hydrogen bonding	42.83/18.23	Buried
T41	I	Buried, lines active site	8.87/2.18	Buried
D93	V	Active site; prevents substrate binding	2.38/2.38	Buried
R112	S	Misfolding	19.69/15.48	Buried
L166	G	Active site; prevents substrate binding	4.19/4.19	Buried
G171	D	Active site; prevents substrate binding	3.04/0.00	Buried
M187	T	Misfolding	0.05/0.05	Buried
S201	Y	Active site, prevents substrate binding	7.71/7.71	Buried
	F	Active site; prevents substrate binding		
D234	E	Dimer interface; disrupts dimer association	1.44/0.84	Buried
W236	R	Dimer interface; buries a charge	20.31/18.66	Buried
D264	Y	Active site; prevents substrate binding	12.79/11.00	Buried
M267	R	Misfolding; buries charge	4.87/3.78	Buried
V269	M	Misfolding	1.76/0.00	Buried
G271	V	Misfolding; phi/psi constraints	15.13/0.00	Intermediate
	S	Misfolding; phi/psi constraints		
S276	G	Misfolding; loss of hydrogen bonding	5.32/3.27	Buried
Q283	P	Misfolding	0.00/0.00	Buried
A285	P	Misfolding	0.01/0.01	Buried
	D	Misfolding, buries charge		
M290	I	Misfolding	3.29/3.10	Buried
P293	T	Misfolding	1.71/0.01	Buried
Q312	H	Unknown	40.14/28.62	Intermediate
Q321	R	Misfolding, disrupts hydrogen bonding	39.66/33.15	Intermediate
G328	V	Misfolding	0.83/0.00	Buried
E338	K	Unknown	5.92/5.42	Buried
A348	P	Misfolding	7.98/7.45	Buried
E358	A	Misfolding; loss of charge	33.13/28.70	Intermediate
Q386	P	Misfolding	12.93/12.93	Buried

able for this position. The substitution of an aspartate would cause a charged side-chain to be buried into a small hydrophobic pocket, which would disrupt the active site and misplace the active aspartate-70. This would prevent substrate binding. **M187T:** Methionine-187 is located on the outer edge in the middle of the α4-helix of the β/α-barrel and its side-chain is buried. Although the substitution of a threonine should be tolerated, it will not occupy the same amount of space and would probably disrupt proper folding of the enzyme. **S201Y/F:** Serine-201 is located on the β5-strand of

Table 4. Conservation of novel missense mutations in *α-Gal A* and *α-Gal B* orthologues.

	Eukaryotic (22)	Mammalian (4)	*α-Gal B* orthologues (8)	Conservation score*
N34K	13	3	7	++
T41I	1	3	0	–
D93V	18	3	8	+++
R112S	5	3	7	+
L166G	4	3	0	–
G171D	7	3	8	+
M187T	16	4	5	++
S201F/Y	15	4	8	++
D234E	13	3	6	++
W236R	13	3	7	++
D264Y	20	3	8	+++
M267R	14	3	7	++
V269M	3	3	1	–
G271V/S	21	3	7	+++
S276G	11	3	7	++
Q283P	7	3	7	+
A285P/D	9	3	7	+
M290I	10	3	1	+
P293T	17	3	8	+++
Q312H	5	3	6	+
Q321R	20	3	8	+++
G328V	9	3	7	+
E338K	14	3	6	++
A348P	1	3	6	+
E358A	3	3	1	–
Q386P	3	3	6	+

*For the degree of evolutionary conservation: –, minimal, if any conservation; +, conserved in three of four α-*Gal A* mammalian orthologues and six of eight α-*Gal B*, or five of 22 α-*Gal A* eukaryotic orthologues; ++, additionally, conserved among 11–16 (50–73%) of 22 α-*Gal A* eukaryotic orthologues; +++, additionally, conserved among at least 17 (77%) of 22 α-*Gal A* eukaryotic orthologues.

the N-terminal β/α-barrel and has its side-chain directed into the lower part of the active site cavity. The substitution of either a tyrosine or phenylalanine in this position would decrease the size of the active site. Additionally, a tyrosine or phenylalanine would displace lysine-168, which appears to be critical in properly orientating the substrate into the active site for cleavage, thus abolishing enzymatic activity. **D234E:** Aspartate-234 is located on the outer edge of the β/α-barrel in the loop connecting the β6-strand and α6-helix, and is at the dimer interface. It interacts only with the backbone nitrogens of phenylalanine-273 and glycine-274 of its polypeptide and not with the other subunit. Although a change to the same charged, but slightly larger, glutamate is not a drastic change, it presumably interferes with dimerisation and renders the enzyme non-functional. **W236R:** Tryptophan-236 is in the loop between the β6-strand and α6-helix on the outer

edge of the β/α-barrel and fills a large cavity close to the dimer interface where there is a pocket. This cavity has hydrophobic residues along one side towards the core of the protein and hydrophilic residues on the side closest to the solvent. An arginine substitution with a positive charge on its side-chain could position its charge near the hydrophilic residues, whereas the hydrophobic middle of the side-chain could interact with the hydrophobic side of the cavity. Arginine would not occupy the large pocket of the tryptophan, however, leading to destabilisation of the enzyme. Tryptophan residues are probably instrumental in early establishment of a protein's hydrophobic core during protein folding, so that this substitution would also interfere with folding. **D264Y:** Aspartate-264 is on the β7-strand of the β/α-barrel and has its side-chain in the active site. A change to tyrosine would constrict the active site and remove the negative charge that assists in orientating the substrate. This substitution would cause either misfolding or markedly impair substrate binding. **M267R:** Methionine-267 is on a loop between the β7-strand and α7-helix that aligns the entrance to the active site and has its side-chain pointing into the active site. A change to arginine would constrict the active site and add an additional positive charge. It would also interfere with lysine-168, which appears to help to align the substrate in the active pocket. Thus, this substitution would markedly interfere with substrate binding. **V269M:** Valine-269 is on a loop between the β7-strand and α7-helix that surrounds the entrance to the active site, but is slightly more removed from the active site than methionine-267. It is in a small hydrophobic pocket and a substitution with methionine would cause some constriction of this area leading to misfolding and impairment of substrate binding. **G271V/S:** Glycine-271 is located in a turn between the β7-strand and α7-helix in the N-terminal β/α barrel. The phi/psi angles are disallowed for the other 19 amino acids. Glycine-271 is in a buried hydrophobic area surrounded by polar residues. A side-chain of either valine or serine could be accommodated at this position, but the rigidity of this turn requires a residue with the flexibility of glycine. Therefore, these changes would lead to misfolding. **S276G:** Serine-276 is on the outer edge of the β/α-barrel at the start of the α7-helix near the dimer interface. Its side-chain is involved in extensive hydrogen bonding with the backbone and side-chain nitrogens of glutamine-279 and with the backbone carbonyl oxygens of phenylalanine-273 and leucine-275. Because this hydrogen bonding network stabilises this part of the enzyme, a change to glycine would disrupt proper folding. **Q283P:** Glutamine-283 is part of the α7-helix in the N-terminal β/α-barrel. A proline at this position would disrupt proper folding of the enzyme. **A285P/D:** Alanine-285 is part of the α7-helix in the β/α-barrel and lies buried in the interface between the N- and C-terminal domains. The proline substitution would disrupt the helix, interfering with the proper folding of the enzyme. The helix also is beside the C-terminal domain, and this interaction would be disturbed. Substitution of an aspartate residue, which is negatively charged, would be extremely unstable and would disrupt folding because there are no surrounding residues to counter this charge. **M290I:** Methionine-290 is at the end of the α7-helix, where it occupies a large hydrophobic pocket. An isoleucine at this position would not occupy the same volume and would destabilise this section of the enzyme, thus disrupting proper folding of the enzyme. **P293T:** Proline-293 occurs just before the β8-strand of the N-terminal β/α barrel and is buried in a central portion of the enzyme some distance from the active site. Prolines often provide rigidity to the protein that promotes proper folding. A threonine substitution at this position should be tolerated, but there will probably be a cost to the folding dynamics of the protein. Any protein that does form should have enzymatic activity, but there may be little, if any, stable enzyme. **Q312H:** The last α-helix (α8) of the N-terminal β/α barrel is actually formed from two helices separated by two residues (glutamine-312 and aspartate-313) that are not in the more extended conformation. The glutamine-312 side-chain is exposed to solvent on one side and the side-chain of tryptophan-81 on the other. The structure should be able to accommodate the substituted histidine, whose side-chain could even make a more favourable interaction with the side-chain of tryptophan-81. This mutation is also distant from the active site. Structurally, it is difficult to determine why this mutation would be deleterious; however, the pKa of histidine suggests that it would be protonated in the lysosome and thus have a positive charge, which could interfere with the organisation in this area of the protein. **Q321R:** Glutamine-321 occurs in the last α-helix (α8) of the N-terminal β/α-barrel. Its side-chain is mostly exposed, but it hydrogen bonds with the side-chain of threonine-39. An arginine substitution would add a positive charge to this area, and its longer side-chain would prevent the interaction with threonine-39, possibly destabilising this area and preventing proper folding. **G328V:** Glycine-328 is located in the loop between the α8-helix of the N-terminal domain and the β9-strand of the C-terminal domain. Glycine is ideal for fitting into tight areas of a protein's structure because it has no side−chain, and this is the case for glycine-328. When changed to valine, there is no room for the bulkier side-chain−and the substitution would cause some disruption in the way that the N- and C-termini were packed together. Most likely, the enzyme would not fold properly and would be degraded. **E338K:** Glutamate-338 is located in the β10-strand of the C-terminal domain and is hydrogen bonded with tryptophan-340 and arginine-356. Arginine-356 makes a strong salt link with aspartate-244 that helps to stabilise the interaction between the N- and C-terminal domains. The substitution of lysine would compete with arginine-356 and destabilise the enzyme's structure, leading to misfolding. **A348P:** Alanine-348 is at the start of the β11-strand in the C-terminal domain. A proline substitution would be difficult to accommodate and would disrupt the preferred secondary structure. Because it

occurs right after a loop, however, one would predict that a small amount of protein would fold properly and be functional. **E358A**: Glutamate-358 is in a loop after the β11-strand of the C-terminal domain; it is completely solvent-exposed and an alanine substitution could easily fit into this position. Glutamate-358 does hydrogen bond to tryptophan-236 and lysine-240, however, stabilising this loop region, which is near the dimer interface. Alanine would not support these interactions and lead to misfolding. **Q386P**: Glutamine-386 is located in the β13-strand of the C-terminal domain. A proline replacement does not contain an amide hydrogen, so it will not maintain the hydrogen bonding of the β-sheets. This mutation would interfere with the proper folding of the C-terminal domain.

Discussion

Mutation analysis of the *α-Gal A* gene in 66 unrelated probands with Fabry disease identified 50 new mutations, demonstrating the extensive molecular genetic heterogeneity underlying this lysosomal storage disease. Of the new mutations, several were notable, including a 3.1 kb deletion (only the fifth large deletion detected in this 'Alu-rich' gene [Human Gene Mutation Database]), an allele with two adjacent base substitutions (L166G) and two alleles, each with two missense mutations (D264Y/V269M, and Q312H/A143T).

Although all types of mutations have been found to cause Fabry disease,[1] there have been relatively few large gene deletions, considering the fact that the *α-Gal A* gene is an Alu-rich gene with about one Alu per kb. Previously, only four deletions over 1 kb were reported among the over 400 mutations causing this disease, and only one of these four was due to Alu–Alu recombination.[1] The large deletion reported here most probably resulted from unequal, but homologous, recombination between the highly homologous and similarly orientated Alu sequences in introns 1 and 2, thereby resulting in the approximately 3.1 kb loss, including all of exon 2 (Fig. 1).

The L166G mutation is unusual, in that it involves two base changes, a double transversion of CT to GG at cDNA positions 496 and 497. A single event most likely resulted in the other complex mutation that gave the missense mutation K374R, located 3 bp upstream of a GAG deletion.

The D264Y and V269M mutations were found in a classically affected male. Both mutations were located in exon 6. Structural analysis predicted that V269M would constrict the active site, but that the protein would fold properly and retain some, albeit reduced, activity. By contrast, the aspartate at position 264 lines the active-site pocket, and the change to tyrosine predicts a marked constriction of the active site. Also, the negative charge that probably assists in orientating the substrate would be lost, markedly altering folding and enzyme activity. These structural predictions were confirmed by *in vitro* expression assays which demonstrated that the *V269M* allele had about 10 per cent of the mean wild-type expressed activity, which was detectable in the lysosomes (Figure 2), whereas the *D264Y* allele expressed no detectable enzyme activity, and the enzyme protein was detected immunologically in the ER.

Most mutations causing Fabry disease are private, occurring in a single or few families;[1,7] however, several mutations at CpG dinucleotides, known mutational hot spots, occur more often in unrelated families. These include R112S, R112C, A143T, Y152X, R227X and R342X in unrelated classically affected families. Here, the R112S mutation is first reported in a classically affected male. In addition, other residues are encoded at CpG dinucleotides, including T39, R49, R118, C142, D153, R220, R301, D315, V316, R356, R363, I367 and A368. Mutations have been reported at all of these CpG sites, except at codons 39, 118, 315, 316, 367 and 368. Four of the mutations reported here (8 per cent) were detected only in the affected proband and were not present in either of the probands' parents. Three of these *de novo* mutations occurred at CpG sites (R112C, Y152X and Q157X). One other *de novo* mutation occurred at S201F.

In this paper, we have mapped the novel missense mutations onto the native, properly folded enzyme to better understand their 3D locations and proximity to the active site. It is easy to understand why mutations at residues aligning the active-site pocket, such as aspartate-93, aspartate-264 or methionine-267, perturb enzyme function because they are critical in binding and orientating the substrate for hydrolysis.[30] Other missense mutations interfere with the proper folding of the enzyme, leading to retention in the ER. The enzyme exists as a homodimer, but there is no indication of functional cooperativity between the two subunits. Two of the mutations occur near the dimer interface (aspartate-234 and tryptophan-236), although they do not interact with the other subunit. The structural analyses predict that three of the substitutions (M187T, Q312H and A348P) would be tolerated or have residual activity. These mutations all resulted in the classic phenotype, however, and only four mutations had significant residual activity in the plasma or leukocytes. The structural studies are based on the properly folded enzyme and cannot provide all of the information needed to determine the folding mechanism of this complex glycoprotein. When the structural analyses are coupled with other *in vitro* expression studies, however, they can give a clearer insight into genotype–phenotype correlations. The mutations that are predicted to be accommodated by structural analysis might be functional if allowed to fold and are candidates for studies with pharmacological chaperones.

In summary, these studies identify 50 additional mutations causing Fabry disease, bringing the total number of reported mutations to over 450. The mutations continue to be found in individual families and only a few have been found in other unrelated families — most of these occurring at CpG dinucleotides, which are mutational hot spots. The studies reported here predict most of the structural alterations resulting from the amino acid substitutions. Most disrupt the

native structure and cause protein misfolding. Those located around the active site interfere with orientation and/or binding with the substrate, thereby altering enzyme function. A few located on or near the dimer interface alter proper dimerisation of the glycoprotein. None of the mutations described here replaced any of the three asparagines at N-glycosylation sites, but N34K presumably could affect glycosylation at N192, where asparagines-34 interacts with the side-chains of asparagines-192. Several of the normally substituted amino acids were predicted to be structurally accommodated and therefore tolerated (ie T41I, M187T, G271V, G271S, P293T, Q312H and A348P); however, these replacements all caused severe loss of enzyme function or stability because they resulted in the classic phenotype. Thus, predicting the phenotype based on the structural alterations of the mutations may underestimate the severity of the substitution on the protein's ability to fold into a functional configuration. Clearly, most of the novel missense mutations described here altered folding and presumably led to the glycopolypeptide's aggregation/retention in the ER and subsequent proteosomal degradation in the cytosol, consistent with the classic phenotype of Fabry disease.

Acknowledgments

This work was supported, in part, by grants from the National Institutes of Health, including a research grant (R37 DK34045, Merit Award), a grant (5 MO1 RR00071) for the Mount Sinai General Clinical Research Center Program from the National Center of Research Resources and a basic research grant from the Genzyme Corporation. M.Y. is the recipient of an NIH postdoctoral training fellowship in Mental Retardation and Developmental Disabilities (5T32 HD07105).

Electronic database information

Human Gene Mutation Database: http://www.hgmd.org.
GenBank: http://www.ncbi.nlm.nih.gov.
Single Nucleotide Polymorphism (SNP) Database:
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=snp.

References

1. Desnick, R.J., Ioannou, Y.A., Eng, C.M. *et al.* (2001), 'α-Galactosidase A deficiency: Fabry disease', in Scriver, C.R., Beaudet, A.L., Sly, W.S. *et al.* (eds), '*The Metabolic and Molecular Bases of Inherited Disease*', 8th edn, McGraw-Hill, New York, NY, pp. 3733–3774.
2. Elleder, M., Bradova, V., Smid, F. *et al.* (1990), 'Cardiocyte storage and hypertrophy as a sole manifestation of Fabry's disease. Report on a case simulating hypertrophic non-obstructive cardiomyopathy', *Virchows Arch. A Pathol. Anat. Histopathol.* Vol. 417, pp. 449–455.
3. Nakao, S., Kodama, C., Takenaka, T. *et al.* (2003), 'Fabry disease: Detection of undiagnosed hemodialysis patients and identification of a "renal variant" phenotype', *Kidney Int.* Vol. 64, pp. 801–807.
4. Nakao, S., Takenaka, T., Maeda, M. *et al.* (1995), 'An atypical variant of Fabry's disease in men with left ventricular hypertrophy', *N. Engl. J. Med.* Vol. 333, pp. 288–293.
5. von Scheidt, W., Eng, C.M., Fitzmaurice, T.F. *et al.* (1991), 'An atypical variant of Fabry's disease with manifestations confined to the myocardium', *N. Engl. J. Med.* Vol. 324, pp. 395–399.
6. Ashley, G.A., Shabbeer, J., Yasuda, M. *et al.* (2001), 'Fabry disease: Twenty novel α-galactosidase A mutations causing the classical phenotype', *J. Hum. Genet.* Vol. 46, pp. 192–196.
7. Shabbeer, J., Yasuda, M., Luca, E. and Desnick, R.J. (2002), 'Fabry disease: 45 novel mutations in the a-galactosidase. A gene causing the classical phenotype', *Mol. Genet. Metab.* Vol. 76, pp. 23–30.
8. Barker, D.F., Schafer, M. and White, R. (1984), 'Restriction sites containing CpG show a higher frequency of polymorphism in human DNA', *Cell* Vol. 36, pp. 131–138.
9. Cooper, C. and Youssoufian, H. (1988), 'The CpG dinucleotide and human genetic disease', *Hum. Genet.* Vol. 78, pp. 151–155.
10. Davies, J.P., Winchester, B.G. and Malcolm, S. (1993), 'Sequence variations in the first exon of alpha-galactosidase A', *J. Med. Genet.* Vol. 30, pp. 658–663.
11. Fitzmaurice, T.F., Desnick, R.J. and Bishop, D.F. (1997), 'Human α-galactosidase A: High plasma activity expressed by the -30G → A allele', *J. Inherit. Metab. Dis.* Vol. 20, pp. 643–657.
12. Froissart, R., Guffon, N., Vanier, M.T. *et al.* (2003), 'Fabry disease: D313Y is an alpha-galactosidase A sequence variant that causes pseudo-deficient activity in plasma', *Mol. Genet. Metab.* Vol. 80, pp. 307–314.
13. Yasuda, M., Shabbeer, J., Benson, S.D. *et al.* (2003), 'Fabry disease: Characterization of alpha-galactosidase A double mutations and the D313Y plasma enzyme pseudodeficiency allele', *Hum. Mutat.* Vol. 22, pp. 486–492.
14. Brown, R.M. and Brown, G.K. (1993), 'X-chromosome inactivation and the diagnosis of X-linked disease in females', *J. Med. Genet.* Vol. 30, pp. 177–184.
15. Lyon, M. (1961), 'Gene action in the X-chromosome of the mouse (*Mus musculus L.*)', *Nature* Vol. 190, pp. 372–373.
16. Willard, H.F. (2001), 'The sex chromosome and X chromosome inactivation', in Scriver, C.R., Beaudet, A.L., Sly, W.S. *et al.* (eds), '*The Metabolic and Molecular Bases of Inherited Disease*', McGraw-Hill, New York, NY, pp. 1191–1212.
17. Eng, C.M., Banikazemi, M., Gordon, R. *et al.* (2001), 'A phase 1/2 clinical trial of enzyme replacement in Fabry disease: Pharmacokinetic, substrate clearance, and safety studies', *Am. J. Hum. Genet.* Vol. 68, pp. 711–722.
18. Eng, C.M., Guffon, N., Wilcox, W.R. *et al.* (2001), 'Safety and efficacy of recombinant human α-galactosidase A replacement therapy in Fabry's disease', *N. Eng. J. Med.* Vol. 345, pp. 9–16.
19. Schiffmann, R., Kopp, J.B., Austin, H.A. *et al.* (2001), 'Enzyme replacement therapy in Fabry disease: A randomized controlled trial', *JAMA* Vol. 285, pp. 2743–2749.
20. Wilcox, W.R., Banikazemi, M., Guffon, N. *et al.* (2004), 'Long-term safety and efficacy of enzyme replacement therapy for Fabry disease', *Am. J. Hum. Genet.* Vol. 75, pp. 65–74.
21. Galanos, J., Nicholls, K., Grigg, L. *et al.* (2002), 'Clinical features of Fabry's disease in Australian patients', *Intern. Med. J.* Vol. 32, pp. 575–584.
22. MacDermot, K.D., Holmes, A. and Miners, A.H. (2001), 'Anderson-Fabry disease: Clinical manifestations and impact of disease in a cohort of 60 obligate carrier females', *J. Med. Genet.* Vol. 38, pp. 769–775.
23. Desnick, R.J., Brady, R., Barranger, J. *et al.* (2003), 'Fabry disease, an under-recognized multisystemic disorder: Expert recommendations for diagnosis, management, and enzyme replacement therapy', *Ann. Intern. Med.* Vol. 138, pp. 338–346.
24. Fan, J.Q. (2003), 'A contradictory treatment for lysosomal storage disorders: Inhibitors enhance mutant enzyme activity', *Trends Pharmacol. Sci.* Vol. 24, pp. 355–360.
25. Asano, N., Ishii, S., Kizu, H. *et al.* (2000), '*In vitro* inhibition and intracellular enhancement of lysosomal α-galactosidase A activity in Fabry lymphoblasts by 1-deoxygalacto-nojirimycin and its derivatives', *Eur. J. Biochem.* Vol. 267, pp. 4179–4186.
26. Desnick, R.J. and Schuchman, E.H. (2002), 'Enzyme replacement and enhancement therapies: Lessons from lysosomal disorders', *Nat. Rev. Genet.* Vol. 3, pp. 954–966.
27. Fan, J.Q., Ishii, S., Asano, N. and Suzuki, Y. (1999), 'Accelerated transport and maturation of lysosomal α-galactosidase A in Fabry lymphoblasts by an enzyme inhibitor', *Nat. Med.* Vol. 5, pp. 112–115.

28. Ishii, S., Yoshioka, H., Mannen, K. *et al.* (2004), 'Transgenic mouse expressing human mutant a-galactosidase A in an endogenous enzyme deficient background: A biochemical animal model for studying active-site specific chaperone therapy for Fabry disease', *Biochim. Biophys. Acta* Vol. 1690, pp. 250–257.
29. Frustaci, A., Chimenti, C., Ricci, R. *et al.* (2001), 'Improvement in cardiac function in the cardiac variant of Fabry's disease with galactose-infusion therapy', *N. Eng. J. Med.* Vol. 345, pp. 25–32.
30. Garman, S.C. and Garboczi, D.N. (2004), 'The molecular defect leading to Fabry disease: Structure of human alpha-galactosidase', *J. Mol. Biol.* Vol. 337, pp. 319–335.
31. Matsuzawa, F., Aikawa, S.I., Doi, H. *et al.*, (2005), 'Fabry disease: Correlation between structural changes in α-galactosidase, and clinical and biochemical phenotypes', *Hum. Genet.* Vol. 117, pp. 317–328.
32. Desnick, R.J., Allen, K.Y., Desnick, S.J. *et al.* (1973), 'Fabry's disease: Enzymatic diagnosis of hemizygotes and heterozygotes. α-Galactosidase activities in plasma, serum, urine, and leukocytes', *J. Lab. Clin. Med.* Vol. 81, pp. 157–171.
33. Shabbeer, J., Robinson, M. and Desnick, R.J. (2005), 'Detection of α-galactosidase A mutations causing Fabry disease by denaturing high performance liquid chromatography', *Hum. Mutat.* Vol. 25, pp. 299–305.
34. Wang, A., Bishop, D. and Desnick, R. (1990), 'Human α-*N*-acetylgalactosaminidase: Molecular cloning, nucleotide sequence, and expression of a full-length cDNA', *J. Biol. Chem.* Vol. 265, pp. 21859–21866.
35. Higgins, M.E., Davies, J.P., Chen, F.W. and Ioannou, Y.A. (1999), 'Niemann-Pick C1 is a late endosome-resident protein that transiently associates with lysosomes and the trans-Golgi network', *Mol. Genet. Metab.* Vol. 68, pp. 1–13.
36. Garman, S.C., Hannick, L., Zhu, A. and Garboczi, D.N. (2002), 'The 1.9 A structure of α-N-acetylgalactosaminidase: Molecular basis of glycosidase deficiency diseases', *Structure (Camb.)* Vol. 10, pp. 425–434.
37. Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard, M. (1991), 'Improved methods for building protein models in electron density maps and the location of errors in these models', *Acta Crystallogr.* Vol. A47, pp. 110–119.
38. Brunger, A.T., Adams, P.D., Clore, G.M. *et al.* (1998), 'Crystallography & NMR system: A new software suite for macromolecular structure determination', *Acta Crystallogr.* Vol. D54, pp. 905–921.