Am. J. Hum. Genet. 73:162–173, 2003

Fabry Disease: Novel α-Galactosidase A 3′-Terminal Mutations Result in Multiple Transcripts Due to Aberrant 3′-End Formation

Makiko Yasuda,[1,2] Junaid Shabbeer,[1] Makiko Osawa,[2] and Robert J. Desnick[1]

[1]Department of Human Genetics, Mount Sinai School of Medicine, New York; and [2]Department of Pediatrics, Tokyo Women's Medical University, Tokyo

Mutations in the gene that encodes the lysosomal exoglycohydrolase, α-galactosidase A (α-GalA), cause Fabry disease, an X-linked recessive inborn error of glycosphingolipid catabolism. Human α-GalA is one of the rare mammalian genes that has its polyadenylation signal in the coding sequence and lacks a 3′ untranslated region (UTR). We identified two novel frameshift mutations, 1277delAA (del2) and 1284delACTT (del4), in unrelated men with classical Fabry disease. Both mutations occurred in the 3′ terminus of the coding region and obliterated the termination codon, and del2 also altered the polyadenylation signal. To characterize these mutations, 3′ rapid amplification of cDNA ends (RACE) and polymerase chain reactions (PCR) were performed, and the amplicons were subcloned and sequenced. Both mutations generated multiple transcripts with various lengths of 3′ terminal sequences, some elongating ~1 kb. Mutant transcripts were classified as follows: type I transcripts had terminal in-frame thymidines that created termination codons when polyadenylated, type II had downstream termination codons within the elongated α-GalA sequence, and type III, the most abundant, lacked termination codons at their 3′ ends. To determine if the type III transcripts were degraded by the recently described cytosolic messenger RNA degradation pathway for messages lacking termination codons, northern blot analysis was performed. However, the finding of similar levels of nuclear and cytoplasmic α-GalA mRNA in normal and patient lymphoblasts suggested that mRNA degradation did not result from either mutation. Expression of representative transcript types revealed differences in intracellular localization and/or protein stability and catalytic activity, with most mutant proteins being nonfunctional. Characterization of these 3′ mutations identified a novel molecular mechanism causing classical Fabry disease.

Introduction

Fabry disease (MIM 301500) is an X-linked recessive inborn error of glycosphingolipid catabolism resulting from the deficient activity of the lysosomal exoglycohydrolase, α-galactosidase A (α-GalA) (EC 3.2.1.22). The enzymatic defect results in the progressive accumulation of neutral glycosphingolipids with terminal α-linked galactosyl moieties (primarily globotriaosylceramide [GL-3]) in the plasma and in the lysosomes of cells throughout the body (Desnick et al. 2001). In classically affected males with little, if any, α-GalA activity, the accumulation of GL-3, particularly in the vascular endothelium, leads to the major clinical manifestations, including acroparesthesias, angiokeratoma, hypohidrosis, and the characteristic corneal and lenticular opacities. Onset usually occurs in childhood or adolescence, and, with advancing age, vascular complications of the heart, kidney, and brain lead to early demise. In contrast, atypical variants, who have low levels of residual α-GalA activity, lack the typical manifestations and present later in life, with major symptoms limited to the heart (von Scheidt et al. 1991; Nakao et al. 1995; Desnick et al. 2001).

The complete genomic and cDNA sequences of the human α-GalA gene have been determined (Bishop et al. 1986; Kornreich et al. 1989). This gene is unique among eukaryotic genes, because it lacks a 3′ UTR, except for rare variant mRNAs having short 3′ UTRs of 6 or 7 nt (Bishop et al. 1988). The polyadenylation signal (pAS) is in the coding sequence, and the termination signal is in the last codon. To date, >300 disease-causing α-GalA mutations have been identified, including missense mutations, small deletions/insertions, splice mutations, and large gene rearrangements (Desnick et al. 2001). Despite the marked molecular heterogeneity of this disease, mutations that affect 3′-end formation have not been described.

In most eukaryotic mRNAs, 3′-end formation is generated through endonucleolytic cleavage and addition

Received March 5, 2003; accepted for publication April 29, 2003; electronically published June 6, 2003.

Address for correspondence and reprints: Dr. Robert J. Desnick, Department of Human Genetics, Mount Sinai School of Medicine, 1425 Madison Avenue, Box 14-98, New York, NY 10029. E-mail: rjdesnick@mssm.edu

0002-9297/2003/7301-0015$15.00

of the poly(A) tail. In mammalian cells, this reaction depends on three elements that define the core pAS: the consensus AAUAAA hexonucleotide or some functional variant, a degenerate U-rich or GU-rich sequence 10–30 nt downstream of the cleavage site (the downstream element [DSE]), and the cleavage site itself, which becomes the point of poly(A) addition, that is, the poly(A) site (Zhao et al. 1999). The poly(A) tail presumably enhances the translation and stability of mRNAs (Sachs et al. 1997; Preiss and Hentze 1998), as well as contributing to their transport from the nucleus to the cytoplasm (Eckner et al. 1991; Huang and Carmichael 1996; Hilleren et al. 2001). Defects in 3′-end formation can therefore have profound effects on gene expression and may cause disease. Only a few human pAS mutations have been reported elsewhere; they include mutations in the α- and β-globin (Orkin et al. 1985; Rund et al. 1992; Harteveld et al. 1994), the IL2RG (Hsu et al. 2000), and the FOXP3 genes (Bennett et al. 2001), all four of which have 3′ UTRs. Most of these pAS mutations caused disease by reducing the levels of their respective mature mRNAs.

Here, we describe two novel α-GalA mutations that altered 3′-end formation in unrelated men with classical Fabry disease. Both lesions were small deletions that caused frameshift mutations at the 3′ terminus of the gene, thereby obliterating the termination codon. Surprisingly, both mutations resulted in multiple transcripts with various 3′ lengths. Of note, many of the mutant transcripts were cleaved at upstream and distant downstream sites, which had alternative pAS motifs. To further characterize these disease-causing mutations, the stability of the mutant transcripts in each patient's cultured lymphoblasts was analyzed, and the expressed activity and subcellular localization of representative mutant proteins were determined in COS-7 cells.

Subjects, Material, and Methods

Probands

Proband 1 was a 51-year-old Swedish man who had onset of acroparesthesias at 10 years of age and subsequently had gastrointestinal manifestations, including abdominal pain and chronic diarrhea. He developed hypertrophic cardiomyopathy and, because of an atrioventricular block II–III, required a pacemaker. His renal function was normal, with only a trace of urinary protein. His α-GalA activities in plasma and cultured lymphoblasts were 0.4 nmol/h/ml (normal mean ± SD 21.6 ± 6.4 nmol/h/ml) and 1.8 nmol/h/mg protein (normal mean ± SD 86.3 ± 15.8 nmol/h/mg protein), respectively.

Proband 2 was a 51-year-old Brazilian man who had childhood onset of acroparesthesias, angiokeratoma, hypohidrosis, and corneal opacities. He had microalbuminuria, which may have been secondary to his diabetes mellitus, but retained normal renal function. He had no evidence of cardiac or cerebral involvement. His α-GalA activities in plasma and cultured lymphoblasts were 0.4 nmol/h/ml and 1.7 nmol/h/mg, respectively.

Mutation Analysis

Peripheral blood was collected in EDTA from the probands after informed consent was obtained. Genomic DNA was isolated and amplified, and mutation analysis was performed as described elsewhere (Shabbeer et al. 2002). In brief, each of the α-GalA exons and flanking intronic sequences was PCR amplified from genomic DNA. Each amplicon was then sequenced with an ABI Prism 3700 Capillary Array Sequencer, using the ABI Prism BigDye Terminator Ready Reaction Mix (Perkin-Elmer-Cetus). Each mutation was confirmed by repeat PCR amplification and sequencing of the opposite strand and by cosegregation of the lesion and disease in other family members.

RNA Isolation and Characterization of Mutant Transcripts

Total RNA was isolated from control and patient lymphoblast cell lines, using TRI reagent (Molecular Research Center). Polyadenylated RNA (poly[A]$^+$ RNA) was isolated using the Oligotex mRNA kit (Qiagen), and, for each sample, 1.0 μg of poly(A)$^+$ RNA was used to perform 3′ rapid amplification of cDNA ends rapid amplification of cDNA ends (RACE) with the SMART RACE cDNA amplification kit (Clontech). A gene-specific primer near the end of exon 6 (5′-GACGTAATTGCCATCAATCAGGACC-3′) and a nested universal primer within the oligo(dT) anchor sequence (5′-CCTCAAGCTATGGTATCAACGCAGAGT-3′) were used to amplify a fragment of ~380 bp for the normal α-GalA cDNA, with the precise length depending on the length of the poly(A) tail. The PCR products were cloned into PCRII-TOPO vectors, using the Topo TA cloning kit (Invitrogen), and 50 clones were randomly selected to be sequenced for each mutation. The different del2 and del4 subclones were designated transcripts "A1–A20" and "B1–B16," respectively.

Subcellular Fractionation and Northern Analysis

Cultured control and patient lymphoblasts were treated or untreated with 100μg/ml cycloheximide for 5 h, and subcellular fractionation of control and patient lymphoblasts was performed as described elsewhere (Frischmeyer et al. 2002), with the following modifications: cultured lymphoblasts were washed in cold PBS, pelleted by centrifugation at 200 *g* for 5 min at

4°C, and resuspended in hypotonic lysis buffer (10 mmol/L NaCl, 3 mmol/L $MgCl_2$, 10 mmol/L Tris-HCl [pH 7.4], 0.5% NP-40, 1 mmol/L dithiothreitol, and 400 U/ml RNase inhibitor [Roche Molecular Biochemicals]). After vortexing and incubation on ice for 5 min, the nuclei were pelleted by centrifugation at 800 *g* for 5 min at 4°C. mRNA was isolated immediately from the nuclear and cytoplasmic fractions, using TRI reagent (Molecular Research Center). For each cellular fraction, 5.0 μg of mRNA was loaded onto a gel containing 1.0% agarose and 5.0% formaldehyde. After electrophoresis, RNA was transferred onto a Hybond XL membrane (Amersham Pharmacia Biotech) in 20 × standard saline citrate buffer. The filter was incubated at 80°C for 2 h and was UV irradiated and hybridized with a randomly primed ^{32}[P]-labeled full-length α-GalA cDNA probe. To determine the integrity of the RNA, the membrane was rehybridized with a ^{32}[P]-labeled cDNA for human β-actin. The intensity of each signal was estimated using the National Institutes of Health Image program.

Generation of Mutant α-GalA Constructs

The full-length wild-type α-GalA cDNA was obtained from normal lymphoblasts by 3′ RACE, using the sense primer 3′-AGGTTAATCTTAAAAGCCCAGGTTAC-5′ and the nested universal primer (described above). After Topo TA cloning, the insert was subcloned into the pcDNA 3.0 expression vector (Invitrogen), and this construct was designated "pcWT." To generate the full-length mutant α-GalA constructs—pcA2, pcA3, pcA6, pcA7, pcA8, pcA12, pcA14, pcA16, pcA19 (del2), pcB1, pcB3, pcB4, pcB6, pcB7, pcB10, and pcB14–pcB16 (del4)—a 1-kb fragment containing most of the wild-type α-GalA gene from the 5′ end was digested out of the pcWT and was spliced into the respective mutant pCR-II TOPO plasmids, using *Kpn*I restriction sites. The full-length mutant α-GalA cDNAs were then subcloned into the pcDNA 3.0 expression vector. All constructs were confirmed by sequencing. Note that the reverse primer used for 3′ RACE contained three termination codons, each in a different reading frame (fig. 1*A*). Thus, all cloned cDNAs were designed to terminate within this primer sequence, regardless of their reading frame.

Cell Culture and Transient Transfections

Patient lymphoblasts were cultured as described elsewhere (Anderson and Gusella 1984). COS-7 cells from the American Type Culture Collection were maintained in Dulbecco's modified Eagle medium (Invitrogen) supplemented with 10% fetal bovine serum and antibiotics at 37°C with 5% CO_2. COS-7 cells were plated onto 35-mm six-well culture plates with or without coverslips for immunofluorescence microscopy and α-GalA enzyme assays, respectively. Plasmid DNA (1.0 μg) was transfected into each well, using FuGene6 transfection reagent (Roche Molecular Biochemicals), and incubated at 37°C for 48 h.

Protein Quantification and α-GalA Enzyme Assays

α-GalA activity was determined using the fluorogenic substrate, 4-methylumbelliferyl-α-*D*-galactopyranoside (Diagnostic Chemicals), as described elsewhere (Desnick et al. 1973). *N*-acetylgalactosamine (0.1 mmol) (Amersham Pharmacia Biotech) was used as an inhibitor of α-galactosidase B activity. For enzyme assays in COS-7 cells, transient transfections were performed as described above, and cells were harvested, lysed, and assayed for α-GalA activity. Protein concentrations of the samples were determined by use of the DC Protein Assay kit (Bio-Rad). Enzymatic activity was expressed as nanomoles of substrate protein hydrolyzed per hour per milligram. All assays were done in triplicate from three separate transfections, and, after subtraction of background, the mean values were used to calculate the percent of wild-type activity.

Immunofluorescence Microscopy

COS-7 cells were grown on coverslips and transfected as described above. After 48 h of incubation, cells were washed with PBS and fixed with 50% methanol/50% acetone (vol/vol). Then the cells were washed extensively in PBS, blocked in 3% BSA/0.2% Tween-20/PBS for 30 min, and incubated overnight at 4°C with rabbit anti–human α-GalA antibodies (Ioannou et al. 1998). The α-GalA antibody solution was removed, and cells were incubated with mouse anti–human LAMP2 antibodies (Hybridoma Bank, University of Iowa) or mouse anti–human BiP antibodies (BD Biosciences Pharingen) for 1 h at room temperature. After three washes in 1% Triton-X/0.2% Tween-20/PBS, the cells were incubated with donkey anti–rabbit IgG-FITC (Jackson Immuno Research) and donkey anti–mouse IgG–Rhodamine Red (Jackson Immuno Research) for 30 min at room temperature. The coverslips were washed as described above and were mounted onto a microscope slide. The slides were observed under a Nikon Eclipse fluorescence microscope equipped with a charged coupling device camera.

Results

Identification and Characterization of the α-GalA Mutations

PCR amplification and direct sequencing of the α-GalA gene from genomic DNA isolated from probands 1 and 2 revealed a two-base deletion—1277delAA (designated "del2"), which altered the last nucleotide of the pAS AUUAAA to AUUAAG—and a four-base dele-

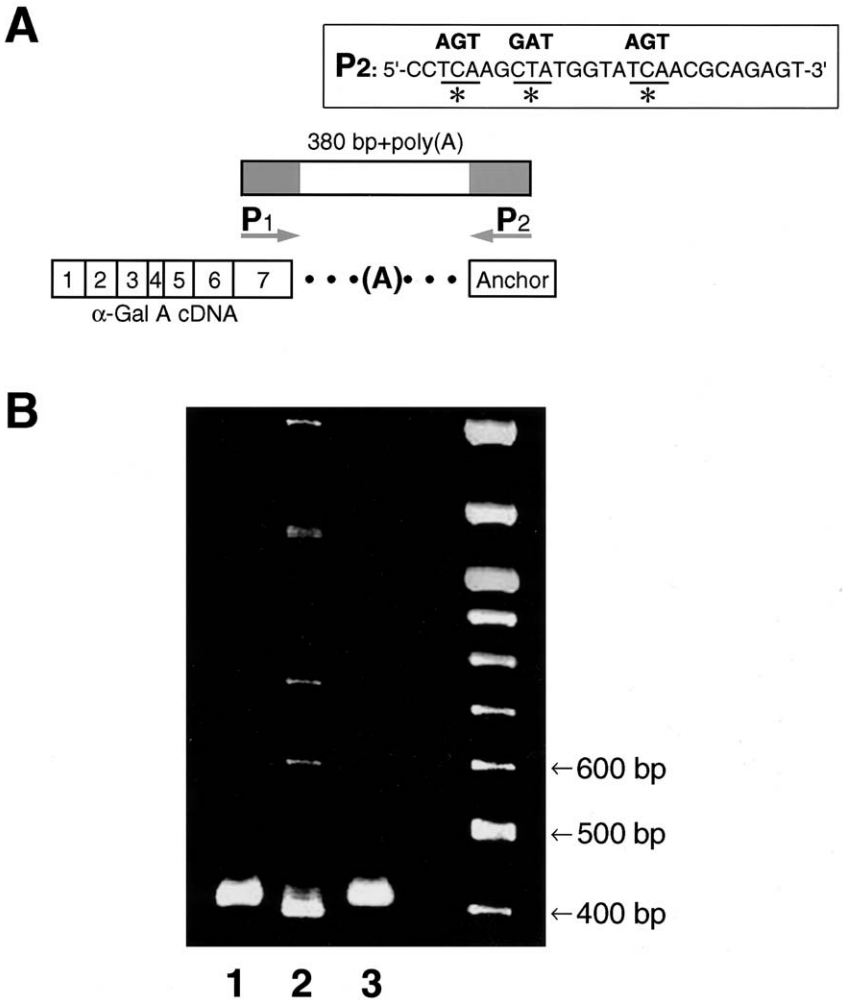

Figure 1 3′ RACE–PCR analysis, using normal and mutant lymphoblasts. *A*, Design of 3′ RACE–PCR. A gene-specific sense primer (P1) at the end of exon 6 and a nested primer within the oligo(dT) anchor sequence (P2) were used. The antisense primer contained three separate termination codons, all in different reading frames. *B*, 3′ RACE–PCR products analyzed on a 1.5% agarose gel. *Lane 1*, normal control; *lane 2*, 1277delAA; *lane 3*, 1284delACTT.

tion—1284delACTT (designated "del4"), which eliminated the four bases before the termination codon. Sequencing of RT-PCR products confirmed their respective lesions and did not identify any other upstream alterations. To characterize the mutant transcripts, 3′ RACE–PCR was performed using mRNA extracted from cultured lymphoblasts established from each proband. Electrophoresis showed that the del2 mutation had a major product that was slightly smaller than the wild-type cDNA and had several minor larger cDNAs, whereas the del4 product was similar in size to the wild-type cDNA (fig. 1*B*). These cDNAs were subcloned, and 50 randomly selected colonies were sequenced for each mutation. Surprisingly, both mutations had multiple cDNAs with various 3′ lengths, some being shorter than wild type and others being elongated (fig. 2). In contrast, sequencing of 10 colonies for the wild-type 3′ RACE cDNAs revealed seven subclones having no 3′ UTR and only three subclones having six or seven additional nucleotides before the poly(A) tail, as described elsewhere (Bishop et al. 1988). The different del2 and del4 subclones were designated transcripts "A1–A20" and "B1–B16," respectively.

The major del2 transcripts—A3, A4, and A6 (fig. 2*A*)—were 5–14 nt shorter than wild type and accounted for ~40% of the transcripts. Only A8, which accounted for 8% of del2 transcripts, used the same cleavage/polyadenylation site as wild-type α-GalA. Transcripts A1, A2, A5, and A7 were within the range of −17 to −1 nt of the wild type, whereas transcripts A9–A15 were elongated by 1–22 nt. Interestingly, ~20% of the cloned transcripts (A16–A20), were elongated by >180 nt, the

A

Normal α-Gal A: 1277del2

CTA GAA *AAT ACA* ATG CAG ATG TCA TTA AAA GAC TTA CTT **TAA** aat gtt tat ttt att gcc aac tac • • •

Mutant Transcripts:		Frequency	Type
A1	CTA GAA *AAT ACA* ATG CAG ATG TC	1	III
A2	• • • • • • • TCA T	1	I
A3	• • • • • • • TCA TT	12	III
A4	• • • • • • • TCA TTA AG	5	III
A5	• • • • • • • TCA TTA AGA CT	1	III
A6	• • • • • • • TCA TTA AGA CTT	5	III
A7	• • • • • • • TCA TTA AGA CTT ACT T	2	I
A8	• • • • • • • TCA TTA AGA CTT ACT TT	4	III
A9	• • • • • • • TCA TTA AGA CTT ACT TTc	1	III
A10	• • • • • • • TCA TTA AGA CTT ACT TTt	1	III
A11	• • • • • • • TCA TTA AGA CTT ACT **TTA** Aaa tgt	1	III
A12	• • • • • • • TCA TTA AGA CTT ACT **TTA** Aaa tgt tta ttt tat tgc	3	III
A13	• • • • • • • TCA TTA AGA CTT ACT **TTA** Aaa tgt tta ttt tat tgc g	1	III
A14	• • • • • • • TCA TTA AGA CTT ACT **TTA** Aaa tgt tta ttt tat tgc cg	1	III
A15	• • • • • • • TCA TTA AGA CTT ACT **TTA** Aaa tgt tta ttt tat tgc caa ct	1	III
A16	• • • • • • • TCA TTA AGA CTT ACT **TTA** Aaa • • •(+181 nt)	4	II
A17	• • • • • • • TCA TTA AGA CTT ACT **TTA** Aaa • • • •(+196 nt)	2	II
A18	• • • • • • • TCA TTA AGA CTT ACT **TTA** Aaa • • • • •(+335 nt)	1	II
A19	• • • • • • • TCA TTA AGA CTT ACT **TTA** Aaa • • • • • •(+706 nt)	1	II
A20	• • • • • • • TCA TTA AGA CTT ACT **TTA** Aaa • • • • • • •(+948 nt)	2	II

B

Normal α-Gal A: 1284del4

CTA GAA *AAT ACA* ATG CAG ATG TCA TTA AAA GAC TTA CTT **TAA** aat gtt tat ttt att gcc aac tac • • •

Mutant Transcripts:		Frequency	Type
B1	CTA G	1	III
B2	CTA GAA *AAT ACA* ATG CAG ATG TCA TT	2	III
B3	• • • • • • • TCA TTA AAA G	3	III
B4	• • • • • • • TCA TTA AAA GAC TT	1	III
B5	• • • • • • • TCA TTA AAA GAC TTT **AA**a atg	3	III
B6	• • • • • • • TCA TTA AAA GAC TTT **AA**a atg t	3	I
B7	• • • • • • • TCA TTA AAA GAC TTT **AA**a atg tt	19	III
B8	• • • • • • • TCA TTA AAA GAC TTT **AA**a atg tc	2	III
B9	• • • • • • • TCA TTA AAA GAC TTT **AA**a atg ct	1	III
B10	• • • • • • • TCA TTA AAA GAC TTT **AA**a atg ttt	6	III
B11	• • • • • • • TCA TTA AAA GAC TTT **AA**a atg tta g	1	III
B12	• • • • • • • TCA TTA AAA GAC TTT **AA**a atg tta aag	1	III
B13	• • • • • • • TCA TTA AAA GAC TTT **AA**a atg ttt att tta ttg cc	2	III
B14	• • • • • • • TCA TTA AAA GAC TTT **AA**a atg ttt att tta ttg cca act	2	III
B15	• • • • • • • TCA TTA AAA GAC TTT **AA**a atg ttt att tta ttg cca aaa t	1	I
B16	• • • • • • • TCA TTA AAA GAC TTT **AA**a • • • • • •(+335 bp)	2	II

Figure 2 Mutant transcripts detected in the 1277delAA (*A*) and 1284delACTT (*B*) mutations. The natural polyadenylation signal, AUUAAA, or mutated AUUAAG (del2) is indicated by an underline, and the alternative signal, AAUACA, is shown in italics. The original termination codon is shown in bold. Sequences upstream of the original termination site are capitalized, whereas nonconsensus and elongated sequences are in lowercase. The mutant transcripts were designated "A1–A20" (1277delAA) and "B1–B16" (1284delACTT), and their frequencies among the 50 colonies are indicated, as well as transcript type. A bullet (•) indicates that the base was identical to the one above.

longest (A20) by ~1 kb. Among these, A16, A18, and A20 were followed by A-rich sequences and therefore may have been derived through false priming of the 3′ RACE oligo(dT) primer (fig. 3). In contrast, transcripts A17 and A19 were not followed by A-rich motifs, suggesting that these cleavage sites must be used in vivo. Notably, transcripts A2 and A7 had terminal, in-frame thymidines, which, if followed by the poly(A) tail, would encode TAA, a de novo termination codon (designated "type I" transcripts). The longer transcripts, A16–A20, had termination codons that were ~55 nt downstream of the obliterated, wild-type poly(A) site (designated

Figure 3 The elongated del2 mutant transcripts are cleaved at different downstream sites. Uppercase sequences are the coding region of the human α-GalA gene, and the downward arrowhead (▼) indicates the natural poly(A) site. The poly(A) sites of the mutant transcripts are shown by insertion of the corresponding transcript numbers in parentheses. Consensus AAUAAA signals are shown in bold, and the DSEs are underlined.

"type II" transcripts) (fig. 4). However, most (~70%) of the del2 transcripts lacked termination codons at their 3′ ends (designated "type III" or "nonstop" transcripts).

In comparison with the variety of del2 transcripts, the del4 transcripts were less diverse (fig. 2*B*). The major del4 transcripts, B7 and B10, were 6 and 7 nt longer than the wild-type forms, respectively, and used the same cleavage sites as the wild-type variant forms (Bishop et al. 1988). Transcript B1, which was 36 nt shorter than wild-type transcripts, was the shortest among the del2 and del4 transcripts. B2–B4 were 5–14 nt shorter than wild type, whereas B5, B6, B8, B9, and B11–B15 were 4–23 nt longer. Transcript B1, as well as B16, which shared the same cleavage site as A18, may have been derived through false priming, since they were followed by A-rich sequences. Transcripts B6 and B15 both had a terminal, in-frame thymidine (type I), and B16 had a termination codon ~70 nt downstream of the obliterated termination site (type II) (fig. 4). The majority (~90%) of del4 transcripts lacked termination codons at their 3′ ends (type III).

Stability of the Mutant Transcripts

To determine whether the del2 and del4 type III transcripts that lacked a termination codon were degraded by the recently proposed "nonstop transcript decay" pathway, the mRNAs in the nuclear and cytoplasmic fractions of wild-type, del2, and del4 lymphoblasts were isolated and assessed by northern hybridization. The α-GalA mRNA levels (normalized to β-actin mRNA) in the nuclear and cytoplasmic fractions were comparable to—or somewhat higher than—those of the wild-type mRNA (fig. 5), even when the cells were treated with the translational inhibitor, cycloheximide (data not shown) (Frischmeyer et al. 2002. This indicated that the del2 and del4 mutant transcripts were stable and abundant. The sizes of the mutant α-GalA mRNAs were essentially the same as the normal transcript, and elongated transcripts were undetectable from either the del2 or del4 mRNAs.

Expression Analysis of Mutant Constructs

The intracellular α-GalA activity and subcellular localization of the mutant proteins encoded by representative del2 and del4 mutant cDNAs were determined by transient transfection in COS-7 cells. The mutant constructs had five different protein expression profiles, depending upon whether or not their cDNAs ended with a termination codon and the extent to which they were elongated (table 1). The mutant proteins that were encoded by transcripts with in-frame terminal thymidines (type I) and within the range of −15 to +5 nt of the wild-type sequence were sorted to the lysosomes (fig. 6*a*) and had wild-type catalytic activity (designated "type I-L"). These included the mutant proteins encoded by del2 transcripts A2 and A7 and del4 transcript B6 (table 1). In contrast, the mutant protein synthesized by transcript B15, which had an in-frame terminal thymidine but was elongated by 23 nt, had only partial activity and was retained in the endoplasmic reticulum (ER) (designated "type I-ER") (fig. 6*b*). Mutant proteins encoded by transcripts elongating >180 nt, which had de novo termi-

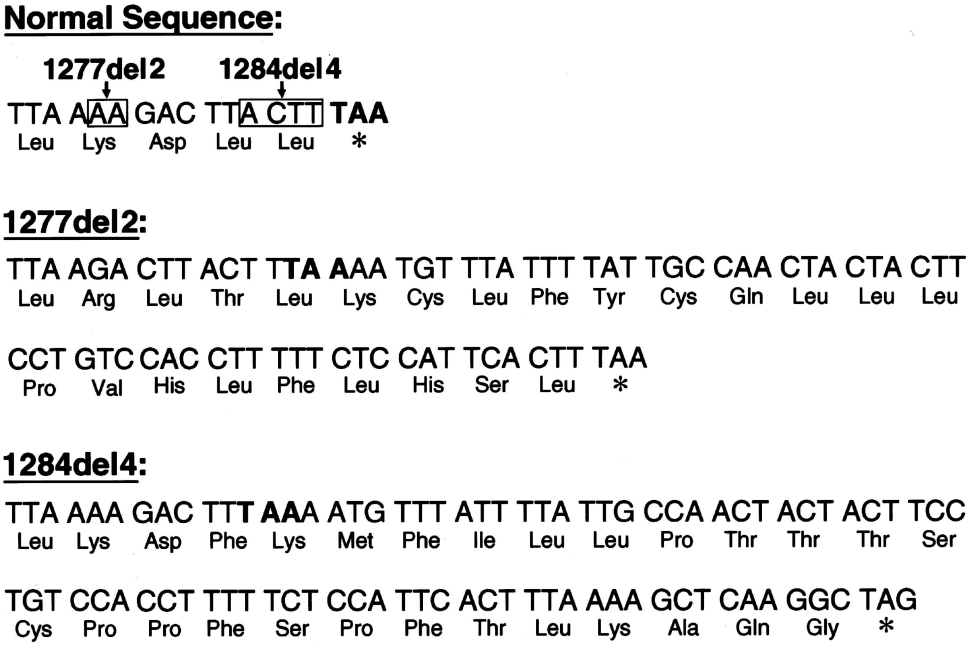

Figure 4 Elongation of the 1277delAA and 1284delACTT mutant proteins. Sequences are aligned by codons, and the corresponding amino acid is shown below each codon. Termination codons are indicated by asterisks (*).

nation codons within their elongated α-GalA sequences (type II), were retained in the ER, because of misfolding, and had no catalytic activity. These included del2 transcripts A16–A20 and del4 transcript B16. The mutant proteins synthesized by type III nonstop transcripts, whose lengths were between 36 nt shorter and 7 nt longer than those of the wild type, were transported to the lysosomes but retained only partial activity (designated "type III-L"). These included del2 transcripts A3, A6, and A8 and del4 transcripts B1, B3, B4, B7, and B10. In contrast, the proteins produced by the longer nonstop transcripts (17–22 nt longer than wild type) had no catalytic activity and were undetectable by immunofluorescence staining. These proteins were designated "type III-X" (fig. 6*c*) and included the mutant forms expressed by del2 transcripts A12 and A14 and del4 transcript B14. Of note, the mutant constructs were designed to have a termination codon within the reverse primer sequence following their poly(A) tails (fig. 1*A*). Thus, nonstop transcripts were predicted to encode multiple lysines, because their poly(A) tails were translated, and to terminate within the next 22 nt. To determine whether poly(A) tract length affected protein expression, nonstop transcripts that were identical to A6 and B7, except for their poly(A) tail length, were selected and expressed. The A6 and B7 transcripts originally used in the expression studies had poly(A) tracts of ~15 adenosines. Interestingly, when the poly(A) length was increased from ~15 to ~60 adenosines, the cataclytic activity was totally lost, and the A6- and B7-expressed proteins became undetectable by immunofluorescence (table 2).

Discussion

Among eukaryotic genes, the human α-GalA gene is unique, in that it has no 3′ UTR but retains the pAS in the coding sequence and a termination signal in the last codon. Therefore, the discovery of two deletions, del2 and del4, that cause classical Fabry disease was of interest, and studies were conducted to characterize these lesions. Both frameshift mutations resulted in multiple mutant transcripts with different 3′ lengths in cultured lymphoblasts. The del2 mutation, which occurred in the last nucleotide of the AUUAAA pAS consensus sequence, resulted in a more diverse population of transcripts, in comparison with the del4 mutation, which did not directly modify the pAS. To our knowledge, this is the first report of 3′ mutations that display such a wide spectrum of mutant transcripts.

This exceptional variety of mutant transcripts seems to result from at least two different mechanisms, alternative pAS usage and aberrant cleavage-site selection. For example, some of the mutant transcripts were cleaved through activation of the alternative pAS, AAUACA, which is located 28 nt upstream of the wild-type termination codon (fig. 2). Transcripts A1–A7 and B2–B4 presumably used this upstream signal, since their

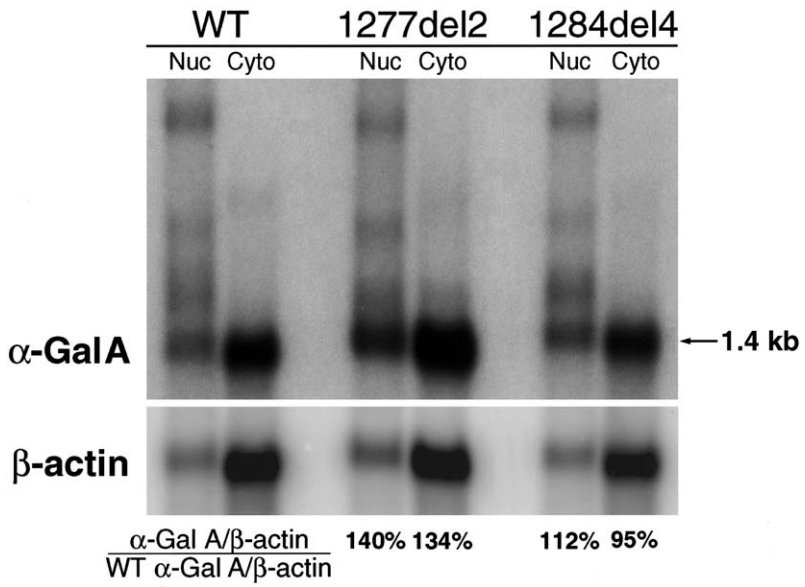

Figure 5 Northern blot analysis of nuclear (Nuc) and cytoplasmic (Cyto) fractions in normal and patient lymphoblast lines (1277delAA and 1284delACTT). α-GalA mRNA levels were normalized against β-actin, and the ratio of normalized mutant α-GalA to wild-type α-GalA was calculated for each cellular compartment.

cleavage sites were within reasonable distance (12–26 nt) from the AAUACA hexonucleotide but were too close (<10 nt) or were even further upstream of the mutated AUUAAG (del2) or wild-type AUUAAA sequence. Most of the remaining transcripts most likely used the mutated AUUAAG signal (in the case of A8–A15) or wild-type AUUAAA signal (in the case of B5–B15). However, since the two signals—AAUACA and AUUAAA/AUUAAG (del2)—were very close to one another, it was difficult, for some mutant transcripts, to pinpoint which of the two signals was actually used. Thus, the del2 mutation, whose major transcripts (A3, A4, and A6) used the upstream alternative AAUACA hexonucleotide, resulted in more diverse transcripts, compared with those resulting from the del4 mutation, whose major transcripts (B7 and B10) used the wild-type AUUAAA signal.

In addition to the upstream alternative pAS, the del2 mutation also seemed to have activated distant downstream pASs. This presumption is based on the fact that both of the elongated del2 transcripts (A17 and A19) had alternative AAUAAA hexonucleotides within 10–30 nt upstream of their cleavage sites, as well as GU-rich DSEs <30 nt downstream (fig. 3). Two mutations in the pAS of the human β-globin gene also have been shown to cause similar read-through transcripts in vivo, which elongated beyond the normal termination site and were presumably cleaved at downstream alternative poly(A) sites (Rund et al. 1992).

Studies have demonstrated that the efficiency of a cleavage/polyadenylation site is defined by the stability of its interaction with the cleavage/polyadenylation complex, particularly the cleavage polyadenylation–specifying factor, with the AAUAAA hexonucleotide, and the cleavage-stimulating factor, with the DSE (Wahle and Ruegsegger 1999). The AAUAAA pAS is highly conserved among mammals, whereas AUUAAA is the most common variant signal (Wahle and Ruegsegger 1999). The upstream alternative pAS, AAUACA, is rare but is used in several mammalian genes (Beaudoing et al. 2000), including the human genes spermidine synthase (Myöhänen et al. 1991) and type I keratin (Marchuk et al. 1985) and in the α-subunit of the photoreceptor cGMP phosphodiesterase (Pittler et al. 1990). In contrast, the mutated pAS of del2, AUUAAG, was not identified in a pAS motif analysis of ~5,600 ESTs, and it was suggested that either its efficiency is low or it may even have a deleterious effect (Beaudoing et al. 2000). It is probable that the del2 mutation, by altering the last nucleotide of the AUUAAA sequence to AUUAAG, caused instability of the pre-mRNA–cleavage/polyadenylation complex and led to the relative strengthening of upstream and multiple downstream pASs, thereby resulting in a wider variety of transcripts, compared with del4.

Another mechanism involved in the generation of multiple mutant transcripts is aberrant cleavage-site selection. Interestingly, even the mutant transcripts that shared the same pAS used multiple adjacent cleavage sites. For instance, transcripts A8–A15 and B5–B15 most

Table 1

In vitro α-GalA Activity and Subcellular Localization of Expressed Mutant Constructs in COS-7 Cells

Mutant Construct	Transcript Type	Length (nt)	In-Frame Thymidine	Termination Codon[a]	α-GalA Activity (% WT)[b]	Subcellular Localization
1277del2:						
A2	I	−15	+	+	116	Lysosome
A7	I	−1	+	+	91	Lysosome
A16	II	+181	−	+	0	ER
A19	II	+706	−	+	0	ER
A3	III	−14	−	−	17	Lysosome
A6	III	−5	−	−	15	Lysosome
A8	III	0	−	−	18	Lysosome
A12	III	+17	−	−	0	Undetectable
A14	III	+19	−	−	0	Undetectable
1284del4:						
B6	I	+5	+	+	105	Lysosome
B15	I	+23	+	+	12	ER
B16	II	+335	−	+	0	ER
B1	III	−36	−	−	25	Lysosome
B3	III	−9	−	−	61	Lysosome
B4	III	−5	−	−	26	Lysosome
B7	III	+6	−	−	46	Lysosome
B10	III	+7	−	−	5	Lysosome
B14	III	+22	−	−	0	Undetectable

[a] All type III transcripts have artificial termination codons (fig. 1*A*), which permit expression in vitro (see "Discussion" section for details).

[b] WT = wild type.

likely used the AUUAAG (del2) or wild-type AUUAAA signal. However, among these transcripts, only A8, which accounted for 8% of all del2 transcripts, was cleaved/polyadenylated at the same site as the wild type (fig. 2*A*). Aberrant cleavage-site selection may have resulted from alteration of the relative positions of the AAUAAA motif and the DSE, since the distance between these two elements is believed to be an important cleavage site determinant (Chen et al. 1995). Of note, both the del2 and del4 mutations occur at STRs, presumably through backward slipped strand mispairing. The del2 mutation occurs in the last two of four consecutive adenosines, whereas the del4 occurs in the second set of the repeated nucleotides ACTT (fig. 2). Although the del2 mutation altered the pAS, del4 altered neither the pAS nor the sequence surrounding the termination codon, since it occurred within a repeated sequence. Thus, the del4 mutation provides further evidence that distance, more than sequence context, is a crucial determinant of cleavage-site selection. Although there is no strict requirement for the precise sequences at which cleavage occurs, previous studies have shown that there is a preference of A > U > C ≫ G (Chen et al. 1995). Analysis of the mutant del2 and del4 transcripts, excluding those that may have emerged through false priming (i.e., del2 A16, A18, and A20 and del4 B1 and B16), shows that ~70% of del2 transcripts (A1, A3, A4, A6, A8, A15, A17, and A19) were polyadenylated at sites adjacent to As, indicating that this nucleotide preference may have contributed to cleavage-site selection. In contrast, among the del4 transcripts, only ~30% (B2, B3, B10, B13, and B14) were polyadenylated at sites adjacent to As. It should be noted that these transcripts were probably cleaved at the A nucleotide in vivo; thus, the first adenosine of their poly(A) tails was derived from their transcripts. Therefore, we propose that the del2 mutation caused activation of several alternative pASs, as well as aberrant cleavage, whereas the del4 mutation mainly caused aberrant cleavage and resulted in multiple mutant transcripts.

The fact that these mutant transcripts were stable was significant. Recently, a novel mRNA degradation pathway in which transcripts lacking termination codons are degraded by the cytoplasmic exosome was demonstrated in yeast and mammalian cells (Frischmeyer et al. 2002; van Hoof et al. 2002). This pathway was referred to as "nonstop transcript decay." The majority of del2 and del4 transcripts lacked termination codons at their 3′ terminals and, hence, were candidate substrates for this degradation pathway. However, the results of the northern blot analysis suggested that nonstop decay did not occur in the del2 and del4 lymphoblasts. Alternatively, nonstop decay may not have occurred because of the unique 3′ end structure of the human α-GalA gene. In addition, nonstop decay, a translation-dependent pathway, would be impeded if

aberrant 3′-end formation of the mutant transcripts caused inefficient translation in vivo.

To investigate whether the mutant cDNAs were translatable and to determine their protein expression patterns, representative mutant constructs were transfected into COS-7 cells. These studies demonstrated that the mutant cDNAs were translated, although the vast majority did not synthesize functional proteins (table 1). The elongated type II transcripts with downstream termination codons produced misfolded proteins that were retained in the ER. Although nonstop constructs with short (~15 nt) poly(A) tracts made proteins with partial activity, the actual lengths of poly(A) tails of in vivo type III transcripts were assumed to be normal (generally ~200 adenosines), because the mutant transcripts were transported to the cytoplasm (fig. 5) (Eckner et al. 1991; Huang and Carmichael 1996; Hilleren et al. 2001). Therefore, two sets of type III transcripts that use the same cleavage site but have longer (~60 nt) poly(A) tracts were expressed. Although the poly(A)$_{60}$ transcripts were stable in COS-7 cells (data not shown), they did not result in immunologically detectable proteins. Presumably, the strong positive charge of multiple lysines encoded by the poly(A) tail has trapped the protein in the ER membrane, thereby preventing release of the nascent protein into the ER lumen and triggering rapid degradation. Another possibility is that the mutant protein with the longer poly(A) tract may have resulted in a misfolded, highly unstable undimerized polypeptide that was not recognized by the anti–α-GalA antibodies. Hence, it is likely that the mutant proteins encoded by the nonstop transcripts in vivo are highly unstable and rapidly degraded.

Table 2

Effect of Poly(A) Tract Length on Protein Expression Pattern

	Poly(A)$_{15}$		Poly(A)$_{60}$	
Mutant Construct	Activity (% WT)[a]	Localization	Activity (% WT)[a]	Localization
A6	15	Lysosome	0	Undetectable
B7	46	Lysosome	0	Undetectable

[a] WT = wild type.

Thus, only the type I-L transcripts, which had in-frame terminal thymidines and were relatively short, synthesized functional proteins with enzymatic activity comparable to wild type. In the present studies, such transcripts accounted for only 6% of the cDNAs in both patients. The actual α-GalA activity in the lymphoblasts of the two patients was ≤2.0% of the normal mean, suggesting that these functional transcripts were even less abundant in vivo. In addition, since α-GalA is a homodimeric protein, it is possible that the few functional transcripts may have encoded polypeptides that dimerized with nonfunctional enzyme subunits, rendering the heterodimer noncatalytic or unstable.

In summary, we have identified two novel frameshift mutations at the 3′ terminus of the α-GalA gene in unrelated men with classical Fabry disease. Both mutations resulted in multiple, stable transcripts with various 3′ lengths that were classified in three major types. Representative mutant constructs were expressed, demon-

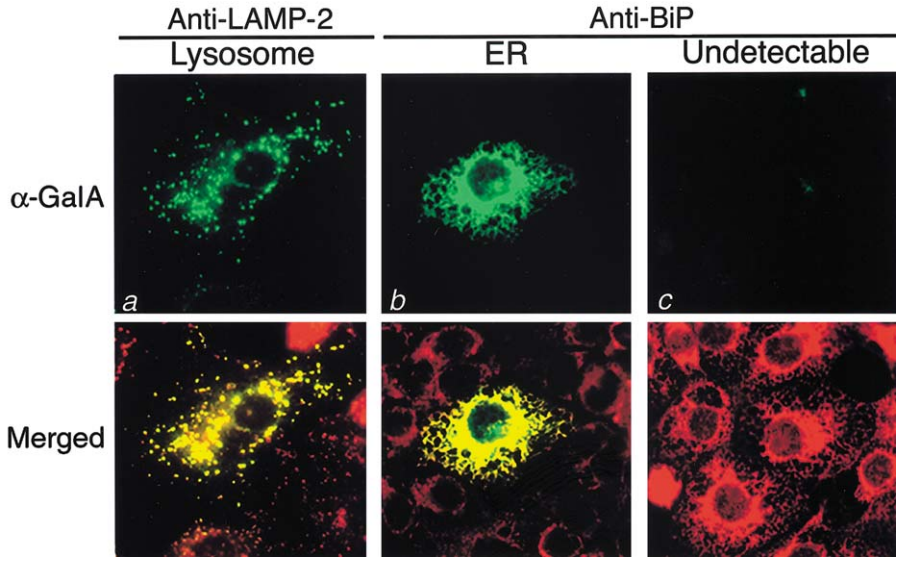

Figure 6 Subcellular localization of mutant proteins. Representative mutant constructs were double stained with rabbit anti–human α-GalA antibodies and either rabbit anti–human LAMP-2 (as the lysosome marker) antibodies (*a*) or rabbit anti–human BiP (the ER marker) antibodies (*b* and *c*). Representative staining patterns of lysosomal (*a*), ER (*b*), and undetectable (*c*) proteins are shown.

strating that the vast majority of mutant polypeptides were nonfunctional. These mutations reveal a novel molecular mechanism causing classical Fabry disease, and they emphasize the molecular genetic heterogeneity of this disorder. The present studies provide an increased understanding of 3′-end formation and nonstop transcript decay in higher eukaryotes.

Acknowledgments

We thank Drs. Harry Dietz (Johns Hopkins University) and David Bishop (Mount Sinai School of Medicine), for helpful discussions and comments on the manuscript, and Drs. E. Pettersson and M. Lipson, for referral of the patients. M.Y. is a recipient of the NORD/Roscoe Brady Lysosomal Storage Disease Fellowship from the National Organization of Rare Disorders. This work was supported in part by National Institutes of Health grants R37 DK34045 (Merit Award), 5 M01 RR00071 (for the Mount Sinai General Clinical Research Center Program), and 5 P30 HD28822 (for the Mount Sinai Child Health Research Center) and by a research grant from the Genzyme Corporation.

Electronic-Database Information

The URL for data presented herein is as follows:

Online Mendelian Inheritance of Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for Fabry disease)

References

Anderson MA, Gusella JF (1984) Use of cyclosporin A in establishing Epstein-Barr virus transformed human lymphoblastoid cell lines. In Vitro 20:856–858

Beaudoing E, Freier S, Wyatt JR, Claverie JM, Gautheret D (2000) Patterns of variant polyadenylation signal usage in human genes. Genome Res 10:1001–1010

Bennett CL, Brunkow ME, Ramsdell F, O'Briant KC, Zhu Q, Fuleihan RL, Shigeoka AO, Ochs HD, Chance PF (2001) A rare polyadenylation signal mutation of the FOXP3 gene (AAUAAA→AAUGAA) leads to the IPEX syndrome. Immunogenetics 53:435–439

Bishop DF, Calhoun DH, Bernstein HS, Hantzopoulos P, Quinn M, Desnick RJ (1986) Human α-galactosidase A: nucleotide sequence of a cDNA clone encoding the mature enzyme. Proc Natl Acad Sci USA 83:4859–4863

Bishop DF, Kornreich R, Desnick RJ (1988) Structural organization of the human α-galactosidase A gene: further evidence for the absence of a 3′ untranslated region. Proc Natl Acad Sci USA 85:3903–3907

Chen F, MacDonald CC, Wilusz J (1995) Cleavage site determinants in the mammalian polyadenylation signal. Nucleic Acids Res 23:2614–2620

Desnick RJ, Allen KY, Desnick SJ, Raman MK, Bernlohr RW, Krivit W (1973) Fabry's disease: enzymatic diagnosis of hemizygotes and heterozygotes α-galactosidase activities in plasma, serum, urine, and leukocytes. J Lab Clin Med 81: 157–171

Desnick RJ, Ioannou YA, Eng CM (2001) α-Galactosidase A deficiency: Fabry disease. In: Scriver CR, Beaudet AL, Sly WS, Valle D, Kinzler KE, Vogelstein B (eds) The metabolic and molecular basis of inherited disease. McGraw-Hill, New York, pp 3733–3774

Eckner R, Ellmeier W, Birnstiel ML (1991) Mature mRNA 3′ end formation stimulates RNA export from the nucleus. EMBO J 10:3513–3522

Frischmeyer PA, van Hoof A, O'Donnell K, Guerrerio AL, Parker R, Dietz HC (2002) An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. Science 295:2258–2261

Harteveld CL, Losekoot M, Haak H, Heister GA, Giordano PC, Bernini LF (1994) A novel polyadenylation signal mutation in the α-2-globin gene causing α-thalassaemia. Br J Haematol 87:139–143

Hilleren P, McCarthy T, Rosbash M, Parker R, Jensen TH (2001) Quality control of mRNA 3′-end processing is linked to the nuclear exosome. Nature 413:538–542

Hsu AP, Tsai EJ, Anderson SM, Fischer RE, Malech H, Buckley RH, Puck JM (2000) Unusual X-linked SCID phenotype due to mutation of the poly-A addition signal of IL2RG. Am J Hum Genet 67:A206

Huang Y, Carmichael GC (1996) Role of polyadenylation in nucleocytoplasmic transport of mRNA. Mol Cell Biol 16: 1534–1542

Ioannou YA, Zeidner KM, Grace ME, Desnick RJ (1998) Human α-galactosidase A: glycosylation site 3 is essential for enzyme solubility. Biochem J 332:789–797

Kornreich R, Desnick RJ, Bishop DF (1989) Nucleotide sequence of the human α-galactosidase A gene. Nucleic Acids Res 17:3301–3302

Marchuk D, McCrohon S, Fuchs E (1985) Complete sequence of a gene encoding a human type I keratin: sequence homologous to enhancer elements in the regulatory region of the gene. Proc Natl Acad Sci USA 82:1609–1613

Myöhänen S, Kauppinen L, Wahlfors J, Alhonen L, Janne J (1991) Human spermadine synthase gene: structure and chromosomal localization. DNA Cell Biol 10:467–474

Nakao S, Takenaka T, Maeda M, Kodama C, Tanaka A, Tahara M, Yoshida A, Kuriyama M, Hayashibe H, Sakuraba H, Tanaka H (1995) An atypical variant of Fabry's disease in men with left ventricular hypertrophy. N Engl J Med 333: 288–293

Orkin SH, Cheng TC, Antonarakis SE, Kazazian HH Jr (1985) Thalassemia due to a mutation in the cleavage-polyadenylation signal of the human β-globin gene. EMBO J 4:453–456

Pittler SJ, Baehr W, Wasmuth JJ, McConnel DG, Champagne MS, vanTuinen P, Ledbetter D, Davis RL (1990) Molecular characterization of human and bovine rod photoreceptor cGMP phosphodiesterase α-subunit and chromosomal localization of the human gene. Genomics 6:272–283

Preiss T, Hentze MW (1998) Dual function of the messenger RNA cap structure in poly(A)-tail-promoted translation in yeast. Nature 392:516–520

Rund D, Dowling C, Najjar K, Rachmilewitz EA, Kazazian HH Jr, Oppenheim A (1992) Two mutations in the β-globin polyadenylylation signal reveal extended transcripts and new RNA polyadenylylation sites. Proc Natl Acad Sci USA 89:4324–4328

Sachs AB, Sarnow P, Hentze MW (1997) Starting at the beginning, middle, and end: translation initiation in eukaryotes. Cell 89:831–838

Shabbeer J, Yasuda M, Luca E, Desnick RJ (2002) Fabry disease: 45 novel mutations in the α-galactosidase A gene causing the classical phenotype. Mol Genet Metab 76:23–30

van Hoof A, Frischmeyer PA, Dietz HC, Parker R (2002) Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. Science 295:2262–2264

von Scheidt W, Eng CM, Fitzmaurice TF, Erdmann E, Hubner G, Olsen EG, Christomanou H, Kandolf R, Bishop DF, Desnick RJ (1991) An atypical variant of Fabry's disease with manifestations confined to the myocardium. N Engl J Med 324:395–399

Wahle E, Ruegsegger U (1999) 3′-End processing of pre-mRNA in eukaryotes. FEMS Microbiol Rev 23:277–295

Zhao J, Hyman L, Moore C (1999) Formation of mRNA 3′ ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol Mol Biol Rev 63:405–445