Structural organization of the human α-galactosidase A gene: Further evidence for the absence of a 3' untranslated region

(Fabry disease/intron-exon junctions/full-length cDNA/promoter/lysosomal hydrolase)

DAVID F. BISHOP, RUTH KORNREICH, AND ROBERT J. DESNICK*

Division of Medical Genetics, Mount Sinai School of Medicine, New York, NY 10029

Communicated by Ernest Beutler, December 28, 1987

ABSTRACT Human α-galactosidase A (α-d-galactoside galactohydrolase; EC 3.2.1.22) is a lysosomal hydrolase encoded by a gene localized to the chromosomal region Xq22. The deficient activity of this enzyme results in Fabry disease, an X chromosome-linked recessive disorder that leads to premature death in affected males. For studies of the structure and function of α-galactosidase A and for characterization of the genetic lesions in families with Fabry disease, the full-length cDNA was isolated, sequenced, and used to screen human genomic libraries. The 1393-base-pair full-length cDNA had a 60-nucleotide 5' untranslated region and encoded a precursor peptide of 429 amino acids including a signal peptide of 31 residues. Three overlapping A clones spanning 32 kilobases were identified that contained the entire ≈12-kilobase chromosomal gene as well as ≈9 and ≈11 kilobases of 5' and 3' flanking sequence, respectively. The gene had seven exons. The genomic exonic and full-length cDNA sequences were identical. All intron-exon splice junctions conformed to the GT/AT consensus sequence. The 5' flanking region of this lysosomal housekeeping gene contained Sp1 and CCAAT box promoter elements as well as sequences corresponding to the activator protein 1 (AP1), octanucleotide ("OCTA"), and "core" enhancer elements. There was an upstream "HTF" island (Hpa II tiny fragments) followed by four direct repeats of the "chiorion box" enhancer. The unique lack of a 3' untranslated sequence in the α-galactosidase A cDNA was confirmed by sequencing additional cDNA clones and the genomic 3' region.

Human α-galactosidase A (α-d-galactoside galactohydrolase; EC 3.2.1.22; α-Gal A) is an X chromosome-linked lysosomal exoglycosidase that cleaves the terminal α-d-galactosyl residues from glycolipids and glycoproteins (1). The human enzyme has been purified to homogeneity from various sources and its physical and kinetic properties have been determined (2–6). The subunits of this homodimeric enzyme are synthesized as ≈50-kDa precursor glycopeptides (7). Following cleavage of the signal peptide, and carbohydrate modifications in the Golgi and lysosomes, the mature enzyme has a native molecular mass of 101 kDa (6).

The deficient activity of α-Gal A results in Fabry disease, an X chromosome-linked recessive disorder characterized by the lysosomal accumulation of the enzyme's primary substrates, globotriasylceramide, galabioasylceramide, and blood group B substance (1). In affected males, the progressive deposition of these substrates in lysosomes of vascular endothelial and smooth muscle cells leads to early demise due to occlusive disease of the heart, kidney, and brain. Classically affected males have no detectable α-Gal A activity and immunologic studies have revealed both the presence and absence of enzyme protein (1, 7–9). Atypical variants with residual α-Gal A activity and immunologically detectable enzyme protein have milder or no clinical manifestations (1, 7, 10–14). Thus, it is likely that the mutations at the α-Gal A locus are heterogeneous, involving gene rearrangements, mRNA splicing and stability defects, and point mutations that alter the expression, function, and/or stability of the enzyme protein.

Recently, we reported the cloning and nucleotide sequence of a cDNA (λAG18) encoding the entire mature lysosomal form of human α-Gal A (15, 16). The λAG18 cDNA did not contain the entire 5' sequence and encoded only five residues of the α-Gal A signal peptide. It was notable that λAG18 and a subsequently cloned α-Gal A cDNA (16) did not have 3' untranslated regions. In these cDNAs, the poly(A) tract immediately followed the termination codon. By using the λAG18 cDNA insert as probe, in situ hybridization studies revealed only one locus for α-Gal A at chromosomal band Xq22 (17). In addition, recent molecular studies of the genetic lesions in unrelated males with Fabry disease have revealed partial gene deletions as well as alterations in mRNA size and quantity (18). In this communication, we report the cloning and the 5' sequence of a full-length cDNA, its use for the isolation of the entire human α-Gal A chromosomal gene, and the characterization of the gene's intron-exon organization. In addition, 5' regulatory elements and the 3' flanking sequence of this lysosomal housekeeping gene are described.

MATERIALS AND METHODS

Materials. Restriction endonucleases, the Klenow fragment of DNA polymerase I, M13 cloning vectors, and universal sequencing primers were purchased from New England Biolabs; T4 DNA ligase was from IBI (New Haven, CT); α-cyanoethyl disopropyl phosphorodiamidites and ancillary DNA synthesis reagents were obtained from Biosearch (San Rafael, CA); agarose was from FMC (Rockland, ME); nitrocellulose filters (type HATF) were purchased from Millipore and Zetabind nylon transfer membranes were from AMF Cuno (Meriden, CT); [α-32P]- and [γ-32P]dNTPs (3000 Ci/mmoll; 1 Ci = 37 GBq) and dATP[α-33S] (1000 Ci/mmol) were obtained from Amersham. A human fibroblast cDNA library in the pCD vector (19) was kindly provided by H. Okayama (National Institutes of Health). A human lung Agt11 cDNA library (HL1004, lot 2007) was obtained from Clone tech (Palo Alto, CA). The 49,XXXXX bacteriophage library (20), designated A4X, was kindly provided by W. I. Wood (Genentech, San Francisco). A 49,XXXXX lymphoblast cell line (GM 6061A) was from the Human Genetic Mutant Cell Repository (Camden, NJ). The EMBL3 vector, in vitro packaging extracts, and Escherichia coli strain K802 were kindly provided by J. Gordon (Mount Sinai School of Medicine).

Abbreviations: α-Gal A, α-galactosidase A; nt, nucleotide(s).
*To whom reprint requests should be addressed: Division of Medical Genetics, Mount Sinai School of Medicine, 100th Street and Fifth Avenue, New York, NY 10029.
†The sequence reported in this paper is being deposited in the EMBL/GenBank data base (IntiGenetics, Mountain View, CA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03249).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
Isolation of Full-Length α-Gal A cDNA Clones. The unamplified pcD human fibroblast cDNA library was plated at a density of 5 × 10^4 colonies per 137-mm HATF filter on x-broth agar plates (19). After growth for 10 hr at 37°C, replicas were regrown and transferred to chloramphenicol plates for an additional 12 hr. Colony hybridization was carried out according to the method of Hanahan and Meselson (21). For use as hybridization probe, the 1.2-kilobase (kb) EcoRI insert of pAG18 (from λAG18; ref. 16) was purified by electroelution from a 0.8% agarose gel and nick-translated to a specific activity of 5 × 10⁶ cpm/μg. In addition, a Agt11 human lung cDNA library was screened by plaque hybridization (22) with the nick-translated pAG18 cDNA insert. The inserts from the positive clones were subcloned directly into M13mp18 and M13mp19 (23). All DNA sequencing reactions were carried out by primer extension in both orientations (16, (17-mers) constructed using either restriction enzymes, EcoRI and PvuII to fragment. 49,XXXXX library, were carried out by primer extension in both orientations (16, 24) using either the M13 universal primer or synthetic oligonucleotides (17-mers) constructed to α-Gal A gene sequences with a Biosearch 8700 DNA synthesizer.

Construction and Screening of X-Chromosome-Enriched Genomic Libraries. Genomic DNA was isolated (25) from the 49,XXXXX human lymphoblast line, partially digested with MboI, and fractionated in a 0.9% agarose gel. Purified targeted DNA (13–22 kb) was ligated to the λ replacement vector, EMBL3 (26), which had been digested to completion with BamHI and EcoRI to prevent reassociation to the middle stuffer fragment. The ligated DNA was packed with extracts prepared by the method of Ish-Horowicz and Burke (27). Approximately 700,000 plaques from the unamplified 49,XXXXX library, designated A5X, and 2 × 10^6 recombinants from the λX library (20) were screened (22) at a density of 10,000 plaques per 150-mm plate. Filters were hybridized with the 32P-labeled, nick-translated insert from pAG18 as described (28).

Characterization of Genomic Clones. Phage DNA was isolated from purified positive plaques (29), digested with various restriction endonucleases, separated by agarose gel electrophoresis (16), and transferred to nylon membranes (30). Sac I- and PvuII-digested DNAs were hybridized with the 32P-labeled, nick-translated α-Gal A M13 deletion subclones (16) to identify and orient the location of exonic fragments. Finer mapping was accomplished with double digests of the genomic inserts and isolated Sac I restriction fragments. Selected genomic fragments containing exonic sequences and 5' and 3' flanking genomic fragments were subcloned into M13 vectors and sequenced by the dideoxynucleotide chain-termination method in both orientations as described above. Additional restriction mapping was performed to position the intron-exon boundaries.

RESULTS AND DISCUSSION

Isolation of a Full-Length cDNA. pAG18 was used to screen the human fibroblast pcD cDNA library (31) and the human lung cDNA library. Following purification, hybridization with 5' and 3' λAG18 M13 subclones (16), and restriction analyses, only 1 of the 15 positive clones, pcDAG126, was putatively full-length. The 1437-nucleotide (nt) Pst I/BamHI insert was subcloned into M13mp18 and M13mp19 and sequenced (Fig. 1). The pcDAG126 insert was 136 nt longer than that of λAG18 and contained 60 nt of 5' untranslated sequence, the initiation codon, and the entire open reading frame, which encoded a 31-amino acid signal peptide and the 398 residues of the mature enzyme. The 60-nt 5' untranslated sequence contained one out-of-frame ATG at −22, which did not have the conserved purine at −25 or the other consensus nucleotides for initiation codons (32) (Fig. 1). The 31-amino acid signal peptide was consistent with that predicted from maturation studies, indicating that human α-gal A was synthesized as an ~50-kDa precursor glycoprotein that was proteolytically cleaved to the mature ~46-kDa lysosomal glycoprotein (7). As shown in Fig. 2, the predicted signal peptide was typical of such sequences (33–35) and included a basic amino acid in the first five residues (Arg-4) followed by a central hydrophobic core of at least nine residues (Leu-8).

![Fig. 1](image-url)

The 5' nucleotide sequence of the α-Gal A gene and its 5' flanking region. The 1478-nt sequence derived from restriction fragments of genomic clones λ-W2 and λ-B5 (see Fig. 3). The sequence includes 1179 nt of 5' flanking DNA, 192 nt of coding sequence, and 107 nt of intron 1. The sequence was determined in both orientations. The sequence from −60 to +192 was identical to that in the full-length cDNA clone, pcDAG126. The signal peptide cleavage site is indicated by an upward arrow. Nucleotide A of the initiation codon, ATG, is designated +1. Sequences resembling TATA or ATA and CCAAT motifs are boxed. Sp1 binding sites (Sp1) are underlined with wavy arrows; octanucleotide ("OCTA") enhancer elements (OCTA), an activator protein 1 (AP-1) enhancer sequence (AP1), cFos enhancer element, and direct repeats (DR1, DR2) are underlined with straight lines. A (G+C)-rich region ranges from −660 to +1.
to Val-22), an α-helix breaker such as proline or glycine −4 to −8 from the cleavage site (Pro-27, Gly-28), a more polar C-terminal region (Asp-25 and Arg-30), and the most frequently observed C-terminal sequence, Ala-Xaa-Ala. Identification of the signal peptide cleavage site by using the weight-matrix method of von Heijne (35) unequivocally predicted cleavage after Ala-31 (score = 7.36). This site was consistent with Leu-32 being the N-terminal residue, as had been established by microsequencing the purified enzyme (15). Thus, microsomal cleavage of the signal peptide is the only N-terminal processing of human α-Gal A. An identical signal peptide sequence for α-gal A was recently reported (36), based on sequencing an α-Gal A cDNA from the pcD library (pcD-AG210) that had a 5′ untranslated region 37 nt shorter than pcDAG126. The N-terminal processing of the human α-Gal A prepeptide is similar to that of other human lysosomal enzymes (e.g., ref. 37) and differs from those that contain pre- and pro-segments requiring a second cleavage to form the mature polypeptide (e.g., refs. 38 and 39).

**Isolation and Restriction Mapping of Genomic Clones.** Bacteriophage libraries enriched for human X chromosomal DNA were screened with the nick-translation pcDAG126 and pAG18 inserts. Three overlapping clones containing the entire α-Gal A cDNA sequence were isolated, purified, and subjected to extensive restriction mapping and Southern hybridization with the radiolabeled λAG18 M13 deletion subclones (16) to locate fragments containing exonic sequences (Fig. 3). Clone λ-W2, isolated from the A4X library, was ≈12 kb and contained the 5′ end of the gene (exon 1) and an additional 9 kb of 5′ flanking sequences. Two overlapping clones, λ-B5 and λ-B18, were isolated from the unamplified λ5X human lymphoblast library. λ-B5 was ≈14 kb and contained the entire coding sequence. λ-B18 was ≈15 kb and contained 3′ coding sequences as well as about 11 kb of 3′ flanking sequence. These three clones spanned 32 kb of the X chromosome, which included the entire α-Gal A gene (~12 kb) and about 20 kb of flanking sequences.

**Intron–Exon Boundaries.** Sequencing of selected α-Gal A restriction fragments revealed the presence of seven exons in the gene (Fig. 3). There were no discrepancies between the exonic sequence in the genomic clones and that of pcDAG126 or pAG18 (16). The exons ranged in length from 92 to 291 nt (Table 1), whereas the introns ranged from 200 nt (intron 5) to 3.7 kb (intron 1). The first exon contained the 60-nt 5′ untranslated sequence as well as the sequence encoding the signal peptide and the first 33 residues of the mature enzyme. The intron–exon boundary sequences for all seven exons are shown in Table 1. All splice junctions followed the “GT/AG” rule (40) and were consistent with consensus sequences for splice junctions of RNA polymerase II-transcribed genes (41). Putative lariat branch points were identified between −23 and −28 nt from the splice junction for all six introns by similarity to the less-well-conserved consensus sequence (C/T)(N/C/T)T(A/G)A(A/C/T)(C/T)/(G/C/T) (42, 43). All three codon phases were observed at the intron–exon junctions in the α-Gal A gene (Table 1).

To date, the β-hexosaminidase α-chain gene is the only other characterized human lysosomal gene (44). In contrast to the structure of the α-Gal A gene, it was larger (~35 kb with 14 exons) and the sequences encoding the pre- (signal peptide) and the pro-segments of the α-chain precursor were in the first exon, suggesting a possible functional domain.

**Regulatory Elements.** Several possible regulatory elements were identified in the α-Gal A 5′ flanking region (Fig. 1). TATA-like sequences were −86, −93, −102, and −129 nt from the initiation codon (all distances are from the initiation codon to the most 3′ nucleotide of the indicated element). Canonical CCAAT box sequences (5′ GGYCAATCT 3′; ref. 45, 46) were located at −101 and −127 nt from the translation initiation site as well as a TATA box at −76 nt. A GC box was identified at −151 nt from the translation initiation site. Two GC box motifs were located at −88 and −151 nt from the translation initiation site. A GC box was identified at −101 nt from the translation initiation site. Two GC box motifs were located at −88 and −151 nt from the translation initiation site. A GC box was identified at −101 nt from the translation initiation site. Two GC box motifs were located at −88 and −151 nt from the translation initiation site.
1. 

Exon sequences are in uppercase letters; intron sequences are in lowercase. Introns 5 and 6 were completely sequenced and the sizes of introns 1-4 were determined by restriction enzyme analyses. Donor and acceptor consensus sequences as well as those for the lariat branch point are indicated. The putative lariat branch points are underlined in the 3' splice acceptor sequences. Codon phase 0 intron–exon junctions occur between codons; codon phase 1 and II interrupt codons after the first and second nucleotides, respectively.

45) were located at -71, -146, and -178 nt from the initiation codon on the anti-sense strand, whereas degenerate forms, CAAT and GTGCAA, occurred at positions -104 and -203 nt, respectively. The G-C box consensus sequence for the promoter-specific transcription factor Sp1, found in other housekeeping genes (46), occurred in the reverse orientation at -63 nt and in the forward orientation at -207 nt. For comparison, the ß-hexosaminidase ß-chain gene had a 5' region -253 to -283 nt from the initiation codon (the cap site is yet to be determined) that contained a "TATA"-like box (TTATTTA) and a "CCAT" box (CCATC) with an intervening (G+C)-rich region (44).

Several potential enhancer binding sites also were present in the a-Gal A gene. The conserved recognition motif (TGACTCA) of the API enhancer protein (47) was identified at -153 nt. The immunoglobulin OCTA enhancer element (ATTTCGAT; ref. 48) was present with one mismatch at -835 nt and, in the opposite orientation, at -889 nt. A perfect match with the reverse complement of the c-fos enhancer element GATGGC (49) occurred at +70 to +78. Interestingly, the "chorion box" enhancer [TCA(T/C)(G/T) (G/A)AG(C/A); refs. 50 and 51] was found in four direct repeats, each separated by ~6 nt at -274, -290, -307, and -323 nt. Future evaluation of these putative promoter/enhancer elements in the regulation of a-Gal A gene expression will require footprinting experiments and functional assays (52) with deletion/mutation mapping.

"HTF" islands (Hpa II tiny fragments) are DNA sequences of 500-2000 nt enriched in CpG dinucleotides (CpG/CpG > 1.0) in (G+C)-rich regions (>50%) that are typically found in "housekeeping" genes (53). The a-Gal A gene contains an HTF island in the region from -260 to -660 nt. This 5' region had a G + C content of 59% and a CpG/CpG dinucleotide ratio of 1.5. Extension of the region to +1 nt, to include a Sac II site that is indicative of HTF islands (54), slightly reduced the G + C content to 54% and the dinucleotide ratio to 1.4. Since HTF islands have been implicated in maintaining inactivation of X chromosome-linked genes (55, 56), analysis of the methylation patterns in this region from active vs. inactive X chromosomes will complement previous studies of X chromosome reactivation of a-Gal A expression in mouse–human somatic cell hybrids (57).

The 3' Flanking Region. A most unusual feature of the AAG18 cDNA insert was the absence of a 3' untranslated sequence; the poly(A) signal sequence was in the coding region 12 nt from the termination codon, which was followed by the poly(A) tract (16). This finding is unique among human nuclear-encoded mRNAs. In mammals, the only other example of an mRNA lacking a 3' untranslated region is the mRNA for mouse thymidylate synthase (58). It is of interest that the human thymidylate synthase transcript did have a 3' untranslated region (58). To further investigate the possibility that a-Gal A transcripts lack a 3' untranslated region, additional cDNA clones from different human cDNA libraries were isolated and their 3' sequences were determined. The possibility that the a-Gal A transcript contained a 3' A-rich sequence that could serve as a site for oligo(dT) binding in the construction of cDNA libraries was ruled out by sequencing the 3' genomic region. As shown in Table 2, five cDNA clones from three different libraries lacked 3' untranslated regions. In two other clones (pcDAG7 and pcDAG41), the TAA termination codons were followed by the short 3' untranslated sequences, AATGGT and AATGTG, respectively, whereas a previously reported cDNA (36) had the sequence AATGTG (Table 2). Although the majority of the isolated a-Gal A cDNAs lacked a 3' untranslated sequence, alternative cleavage and polyanthylination can result in short 6- or 7-nt untranslated sequences, maintaining the termination codon. It is possible that these alternative, short 3' regions represent microheterogeneity of the primary a-Gal A cleavage reaction (59).

Table 1. Nucleotide sequence of the intron–exon boundaries in the human a-Gal A gene

<table>
<thead>
<tr>
<th>Exon number and (size)</th>
<th>cDNA position of exon</th>
<th>5' Splice donor</th>
<th>Intron number and (size)</th>
<th>3' Splice acceptor</th>
<th>Codon phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>(nt)</td>
<td>(kb)</td>
<td>5' Splice donor</td>
<td></td>
<td>5' Splice acceptor</td>
<td></td>
</tr>
<tr>
<td>1 (254)</td>
<td>-60-194</td>
<td>TGCCAGA gtaaga</td>
<td>1 (3.7)</td>
<td>gaaattgattgattatggatgttcctctctccag TGAAGCAG</td>
<td>II</td>
</tr>
<tr>
<td>2 (175)</td>
<td>195-369</td>
<td>CTAATTAT gtagt</td>
<td>2 (2.0)</td>
<td>acaagttgtcttttcctctctctctctctccag TTGGCAGC</td>
<td>0</td>
</tr>
<tr>
<td>3 (178)</td>
<td>370-547</td>
<td>GGCAGAT gtaatg</td>
<td>3 (0.9)</td>
<td>tttctctgtatctccatctctctctctctctctccag TTGGTAA</td>
<td>I</td>
</tr>
<tr>
<td>4 (92)</td>
<td>548-639</td>
<td>TCAAGAAC gtaga</td>
<td>4 (1.8)</td>
<td>aaagtagaagagagatctgtatcttttcctccag CCCATTA</td>
<td>0</td>
</tr>
<tr>
<td>5 (162)</td>
<td>640-801</td>
<td>CAGATAT gtaaa</td>
<td>5 (0.2)</td>
<td>tttctctgtgtattcattcttttttcctctctccag TTATGAG</td>
<td>0</td>
</tr>
<tr>
<td>6 (198)</td>
<td>802-999</td>
<td>TTAGCAC gtaaat</td>
<td>6 (0.3)</td>
<td>gttgtagaacaacaccactttttatccttccag GGGAGAAG</td>
<td>0</td>
</tr>
<tr>
<td>7 (291)</td>
<td>1000-1290</td>
<td>TACTTTAAAattg 3'</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Consensus sequences: donor: AAG gtagg a-Gal lariat: Cnt t tG acceptor: ccccccccccccgac

Table 2. 3' Sequences of a-Gal A cDNAs and genomic clone λ-B18

<table>
<thead>
<tr>
<th>Clone</th>
<th>3' Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>cDNA</td>
<td></td>
</tr>
<tr>
<td>AAG18</td>
<td>...TAA ...</td>
</tr>
<tr>
<td>λHLG4</td>
<td>...TAA ...</td>
</tr>
<tr>
<td>pcDAG8</td>
<td>...TAA ...</td>
</tr>
<tr>
<td>pcDAG9</td>
<td>...TAA ...</td>
</tr>
<tr>
<td>pcDAG126</td>
<td>...TAA ...</td>
</tr>
<tr>
<td>pcDAG41</td>
<td>...TAA AATGGTc.</td>
</tr>
<tr>
<td>pcDAG7</td>
<td>...TAA AATGTTT</td>
</tr>
<tr>
<td>Genomic</td>
<td></td>
</tr>
<tr>
<td>λ-B18</td>
<td>...TAA AATGGTATTTTTTCCCCACT ACTCTCTCTCTGTACCTTTCTCCT...</td>
</tr>
</tbody>
</table>

Downstream consensus: YTGTTT TTTTTTTT

The a-Gal A cDNA and genomic clones were isolated and sequenced. The 51-nt genomic 3' sequence, beginning with the termination codon (TAA) and ending with CTCC, is aligned with the sequence in pcDAG7. The "GT-box" and "T-rich" element each occur twice in the genomic sequence as indicated by underlines.
Additionally, the second upstream (~28 nt) poly(A) signal, AATAAC, may be involved in the variant cleavages. Since the murine thymidylate synthase gene did not have a 3’ untranslated sequence, it was proposed that cleavage and polyadenylation was facilitated by the formation of a stem and loop structure in which the poly(A) signal completely base-paired with downstream sequences (58). When the human α-Gal A genomic sequence was similarly folded (BIOFLD program, BIONET, IntelliGenetics, Mountain View, CA), the poly(A) signal was not completely base-paired with downstream signals. In contrast to the human thymidylate gene, the 3’ flanking region of the α-Gal A gene contained sequences with similarities to the consensus downstream elements, YTGTGTYY (“GT-box”) and Tn (“T-rich”), recently shown to be involved in polyadenylation (60-62). Two pairs of these downstream elements occur in the α-Gal A gene first at +16 and +22 nt and again at +43 and +52 nt from the poly(A) signal (Table 2). It is possible that the alternative cleavage sites may be due to the presence of these additional, more 3’, elements. Whatever the mechanism responsible for these alternative 3’ sequences, it is notable that the α-Gal A transcript does not require a 3’ untranslated region for stability or translation.

In conclusion, knowledge of the genomic organization of the α-Gal A gene should facilitate studies of gene expression and the characterization of the molecular nature of the mutations causing Fabry disease.

Note Added in Proof. Quinn et al. (63) also reported a similar, but shorter, genomic 5’ sequence and a single cap site at ~60 nt. The latter result differs from our findings, which will be described elsewhere.

We thank Dr. J. Gordon for the construction of the human genomic library and Dr. C. Eng, T. Fitzmaurice, H. Giordano, and Y. Ioannou for expert assistance in cDNA cloning and sequencing. This work was supported in part by a grant (DK 34045) from the National Institutes of Health and a grant (1-578) from the March of Dimes Birth Defects Foundation. R.K. was the recipient of a predoctoral fellowship from the National Institutes of Health (5T32 HD07105).


