1990

Isolation, Characterization, and Mapping of a Human Acid β-Galactosidase cDNA

Yoshimi Yamamoto

Cynthia A. Hake

Brian M. Martin

Keith A. Kretz

Amelia J. Ahern-Rindell
University of Portland, ahernrin@up.edu

See next page for additional authors

DNA AND CELL BIOLOGY
Volume 9, Number 2, 1990
Mary Ann Liebert, Inc., Publishers
Pp. 119-127

Isolation, Characterization, and Mapping of a Human Acid β-Galactosidase cDNA

YOSHIMI YAMAMOTO,*,§ CYNTHIA A. HAKE,* BRIAN M. MARTIN,† KEITH A. KRETZ,* AMELIA J. AHERN-RINDELL,* SUSAN L. NAYLOR,‡ MICHAL MUDD,‡ and JOHN S. O'BRIEN*

ABSTRACT

A λgt11 human testicular cDNA library was screened with degenerate oligonucleotide probe mixtures based on amino acid sequence data generated from cyanogen bromide fragments and tryptic fragments of purified human β-galactosidase. Six positive clones were identified after screening 2 × 10^6 plaques. The sequences of these six clones were determined and found to be derived from two different cDNAs. The sequence of the longest of these cDNAs is nearly identical to that recently determined by Oshima *et al.* (1988). It codes for a 76-kD protein and all 11 peptides that were generated from the purified enzyme. The second clone is shorter by 393 bp in the central portion of the coding region. Analysis by Northern blotting revealed the presence of a single mRNA species of 2.45 kb in lymphoblasts and testicular tissue. It is deduced from the amino acid sequence data that proteolytic processing of the precursor form of β-galactosidase must occur by cleavage in the carboxy-terminal portion of the polypeptide perhaps around amino acid 530 at a uniquely hydrophilic sequence. Using a probe generated from the 3′ region of the cDNA, we have mapped the locus coding for human β-galactosidase to chromosome 3p21-3pter.

INTRODUCTION

β-D-GALACTOSIDASE (EC 3.2.1.23) is a lysosomal hydrolase responsible for the removal of terminal galactosyl moieties from glycoproteins and glycolipids, most notably GM1-ganglioside. The primary defect in the metabolic storage diseases GM1-gangliosidosis and Morquio B syndrome is a deficiency of this enzyme (for review, see O'Brien, 1989).

The structural gene for β-galactosidase has been localized to the short arm of human chromosome 3 (Shows *et al.*, 1979; Naylor *et al.*, 1982). In normal human fibroblasts, an 84-kD precursor is synthesized and processed *via* intermediates into a mature 64-kD monomer (Nanba *et al.*, 1988). In the lysosome, a "protective protein" interacts with β-galactosidase monomers effecting their multimerization into a high-molecular-weight aggregate of 600–700 kD (Hoogeveen *et al.*, 1983). The aggregated β-galactosidase plus protective protein and another lysosomal enzyme, α-neuraminidase (EC 3.2.1.18), form a lysosomal membrane-bound complex. Formation of this complex is necessary to protect β-galactosidase from intralysosomal degradation by proteases and to stabilize α-neuraminidase (Hoogeveen *et al.*, 1983; Verheijen *et al.*, 1985; Nanba *et al.*, 1987).

The existence of the complex has made it difficult to isolate and purify β-galactosidase. This difficulty, along with the fact that lysosomal enzymes in general are encoded by mRNAs that are present in extremely low abundance, has hampered attempts to clone the β-galactosidase gene until now.

The sequence of a full-length cDNA coding for β-galactosidase was reported recently by Oshima *et al.* (1988). We also have cloned a full length human cDNA encoding β-galactosidase whose sequence agrees almost completely with theirs. In addition, we have identified the sequence which

*Department of Neurosciences and Center for Molecular Genetics, University of California, San Diego, School of Medicine, La Jolla, CA 92093.
†Molecular Neurogenetics Unit, National Institute for Mental Health, Alcohol, Drug Abuse, Mental Health Administration, Bethesda, MD 20892.
‡Department of Cellular and Structural Biology, Health Science Center, University of Texas, San Antonio, TX 78284.
§Present address: Department of Biology, Hamamatsu University, School of Medicine, Hamamatsu, Shizuoka 431-31, Japan.

is removed by proteolytic processing by structural analysis of pure enzyme, mapped the cDNA to chromosome 3p21-3pter and determined the size of the β-galactosidase mRNA.

MATERIALS AND METHODS

Purification and amino acid sequence analysis of β-galactosidase

Human liver β-galactosidase was purified as described (Norden *et al.*, 1974; Miller *et al.*, 1976) with minor modifications (Yamamoto *et al.*, 1982). Briefly, β-galactosidase was purified by chromatography on concanavalin A-Sepharose, Sepharose-*p*-aminophenyl-β-D-thiogalactoside, Sepharose-6B, and octyl-Sepharose. The purified preparation consisted of β-galactosidase protein (64 kD) and protective protein (32 kD and 20 kD) as described previously (Yamamoto *et al.*, 1982). Homogeneous β-galactosidase protein (64 kD) was obtained by reverse-phase HPLC on a Vydac C4 column (4.6 mm × 25 cm). Before injection, the sample was heated at 100°C for 5 min in 2% $NaDodSO_4$ and 5% 2-mercaptoethanol. A linear gradient of acetonitrile (0–80%) in water containing 0.1% trifluoroacetic acid was used. Three major peaks were collected; these were found after gel electrophoresis to correspond to proteins of 20, 32, and 64 kD. The peak containing the 64-kD protein gave a single band on $NaDodSO_4$ polyacrylamide gel electrophoresis. The 64-kD protein was cleaved with cyanogen bromide (CNBr) or digested with trypsin as follows.

For CNBr cleavage, pure human β-galactosidase was dissolved in 70% formic acid and then a 100× molar excess of solid CNBr was added. After incubation overnight at room temperature, 9 volumes of water were added and the reaction products were lyophilized. CNBr cleaved peptides were separated by $NaDodSO_4$ polyacrylamide gel electrophoresis and electroblotted onto a polyvinylidene difluoride membrane (Immobilon Transfer, Millipore) as described by Matsudaira (1987). Protein was stained with Coomassie blue and well separated bands were cut out and sequenced (Matsudaira, 1987).

Probe B1

A		Thr	Asn	Phe	Ala	Tyr	Cys	Asn	Gly	Ala	
B	5'	ACN	AA$^{U}_{C}$	UU$^{U}_{C}$	GCN	UA$^{U}_{C}$	UG$^{U}_{C}$	AA$^{U}_{C}$	GGN	GC	3'
C	3'	TGN	TT$^{A}_{G}$	AA$^{A}_{G}$	CGN	AT$^{A}_{G}$	ACI	TTI	CC$^{I}_{C}$	CG	5'

Probe B2

A		Gln	His	Tyr	Gly	Phe	Val	Leu	Tyr	Arg	
B	5'	CA$^{A}_{G}$	CA$^{U}_{C}$	UA$^{U}_{C}$	GGN	UU$^{U}_{C}$	GUN	CUN	UA$^{U}_{C}$	CG	3'
C	3'	GT$^{T}_{C}$	GT$^{A}_{G}$	AT$^{A}_{G}$	CCN	AAI	CAN	GA$^{I}_{C}$	ATI	GC	5'

FIG. 1. Probe mixtures used to screen for human β-galactosidase. The two peptide sequences listed were used to synthesize degenerate oligonucleotide probe mixtures as shown. A, Amino acid sequence; B, deduced nucleotide sequence; C, complimentary oligonucleotide sequence synthesized for screening. N = A, C, G, and T. I = inosine.

Trypsin

1	Tyr	Phe	Ala	Leu	Arg								
2	Glu	Ser	Ile	Leu	Leu	Arg							
3	Met	Phe	Glu	Ile	Asp	Tyr	Ser	Arg					
4	Gln	His	Tyr	Gly	Phe	Val	Leu	Tyr	Arg				
5	Val	Ile	Ser	Gly	Ser	Ile	His	Tyr	Ser	Arg			
6	Val	Asn	Tyr	Gly	Ala	Val	Ile	Asn	Asp	Phe			
7	Ala	Tyr	Val	Ala	Val	Asp	Gly	Ile	Pro	Ser			
8	Phe	Ser	Asp	Pro	Asp	Try	Leu	Ala	Ala	Val	Lys		
9	Asn	Asn	Val	Ile	Thr	Leu	Asn	Ile	Thr	Gly	Lys		
10	Thr	Thr	Leu	Pro	Gln	Asp	Thr	Ser	Asn	Pro	Gly	Pro	Leu	Ser

Cyanogen Bromide

11	Phe	Ile	Gly	Gly	Thr	Asn	Phe	Ala	Tyr	Cys	Asn	Gly	Ala	Asn	Ser	Pro	Tyr	Ala

FIG. 2. Amino acid sequences obtained after cleavage of purified human β-galactosidase with trypsin or CNBr. The underlined residues differ from those determined from the nucleotide sequence. It is assumed that the potential glycosylation site in trypsin fragment 9 is not glycosylated because it was identified during amino acid sequencing.

Trypsin digestion was performed in 2 *M* urea, 0.1 *M* Tris-HCl pH 8.0, and 20 m*M* methylamine-HCl with 2% wt/wt trypsin (Boehringer-Mannheim, sequencing grade) β-galactosidase at 37°C overnight. Separation of tryptic peptides was carried out by reverse-phase HPLC on a Brownlee Aquapore column (RP-300, 2.1 mm × 10 cm) using a gradient of acetonitrile in 0.1% trifluoroacetic acid. Twenty-one well-separated peaks were collected and used for amino acid sequence analysis. Amino acid sequencing was accomplished using an Applied Biosystems model 470A gas-phase sequencer equipped with a model 120A on-line PTH-amino acid analyzer.

Oligonucleotide probe synthesis

Two 26-bp-long degenerate oligonucleotide probe mixtures were synthesized by the phosphoramidite method (Matteucci and Caruthers, 1981) based on the amino acid sequence of two β-galactosidase peptides (Fig. 1). The oligonucleotides were 5'-end-labeled using T4-polynucleotide kinase and [γ-^{32}P]ATP as follows: 5 μg of oligonucleotide probe were resuspended in 200 m*M* Tris-HCl pH 7.4, 10 m*M* $MgCl_2$, and 40 m*M* 2-mercaptoethanol with 250 μCi [γ-^{32}P]ATP and 10 units of T4-polynucleotide kinase in a total reaction volume of 50 μl. The reaction mixture was incubated at 37°C for 1 hr and the reaction was stopped by heating at 60°C for 5 min. Unincorporated [^{32}P]ATP was removed from the labeled sample by passage through a Nensorb column (Dupont) using the manufacturer's instructions. The specific activity obtained was 5–10 × 10^8 cpm/μg.

Library screening

A human testis cDNA library constructed in λgt11 was purchased from Clontech. After plating on *E. coli* Y1090, a total of 2 × 10^6 plaques were screened.

The phage DNA was transferred to nitrocellulose (Davis *et al.*, 1986) and fixed by heating at 80°C under vacuum for 2 hr. The filters were prehybridized at 65°C for 4 hr in a solution containing 6× SSC (sodium chloride/sodium citrate buffer), 0.05% sodium pyrophosphate, 0.5% $NaDodSO_4$, 100 μg/ml of denatured salmon sperm DNA, and 10× Denhardt's solution (Maniatis *et al.*, 1982).

Hybridization was carried out at 37°C overnight in the same solution (without $NaDodSO_4$) as above with 20% formamide and ^{32}P-labeled probe mixture (4–10 × 10^6 cpm/ml). The filters were rinsed three times with 6× SSC/0.05% sodium pyrophosphate and washed two times for 30 min each in the same solution at room temperature. The filters were then rinsed with Me_4NCl wash solution (3 *M* tetramethylammonium chloride, 50 m*M* Tris-HCl pH 8.0, 2 m*M* EDTA, 0.1% $NaDodSO_4$ at 37°C and washed two

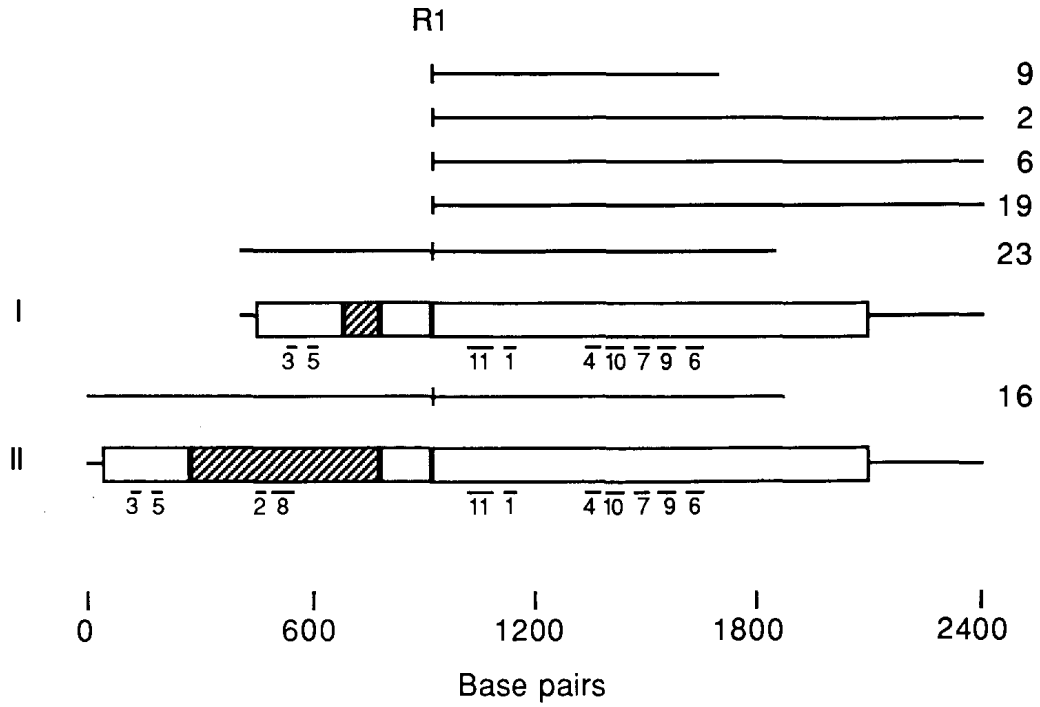
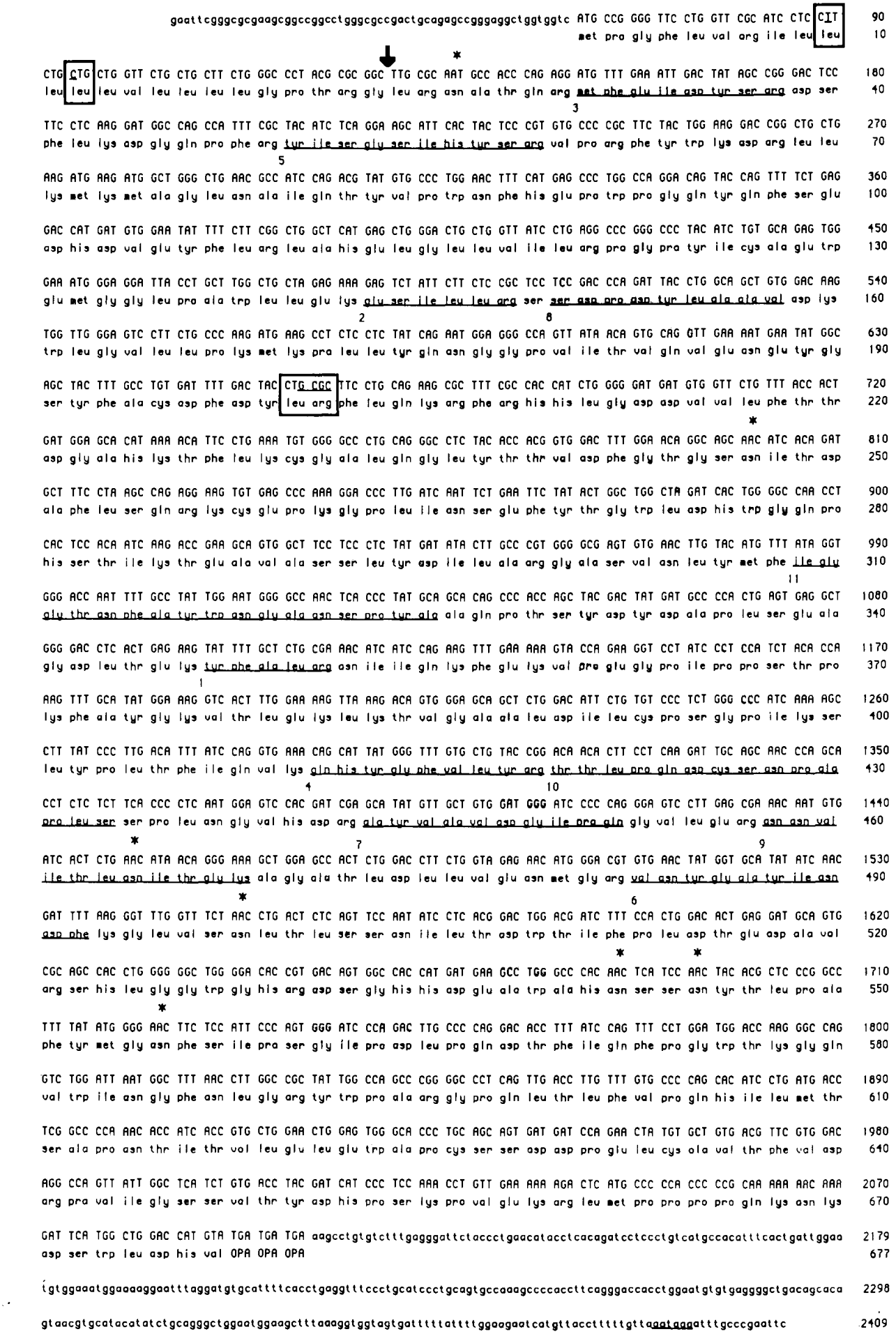


FIG. 3. Structure of human β-galactosidase cDNAs. I and II represent the two cDNA structures determined from sequencing the clones shown. The open box represents the open reading frame while the solid lines represent untranslated regions. The numbered dashes (–) under the cDNA structure indicate the location of the 11 sequenced peptides from Fig. 2. The individual clones (2, 6, 9, 16, 19, and 23) are denoted. R1 denotes the position of the internal *Eco* RI site. The cross-hatched regions denote the region unique to each structure as discussed in the text.



times for 20 min each with the Me_4NCl wash solution at 55°C. The dried filters were exposed to Kodak X-Omat film at −70°C with an intensifying screen. Six positive clones were isolated after five rounds of screening and used for nucleotide sequence analysis.

Cloning and sequencing

Inserts from the positive clones were obtained by *Eco* RI digestion of large-scale phage preparations according to Maniatis *et al.* (1982). The inserts were separated on 0.7% agarose gels and were electroeluted from the gels using an IBI Model UEA electoeluter according to the manufacturer's instructions. The inserts were ligated into pBluescript KS II⁻ (Stratagene) and unidirectional deletions were made as suggested by the manufacturer. Sequencing was performed using modified T7-DNA polymerase (Sequenase, US Biochemicals) according to instructions for double-stranded templates.

Northern blotting and hybridization

Poly(A)RNA was prepared from human lymphoblasts and human testicular tissue using the Fast-Track RNA isolation kit (Invitrogen) as recommended by the manufacturer. Approximately 1 μg of this RNA was separated on a 1% formaldehyde/agarose gel (Maniatis *et al.*, 1982) and blotted onto a nylon membrane (Schleicher & Schuell). RNA probes for hybridization were prepared using the T3 promoter of the pBluescript vector with T3-polymerase and [^{32}P]CTP. Hybridization was carried out in 50% formamide/5× SSC at 55°C overnight according to the method supplied with the vector. The filter was washed at 65°C in 2× SSC/0.1% $NaDodSO_4$ twice for 15 min each, and twice in 0.1× SSC/0.1% $NaDodSO_4$ for 15 min each. The filter was exposed to Kodak X-Omat film for 2 days at −70°C with an intensifying screen.

Chromosomal mapping

Human–rodent somatic cell hybrids were produced by the fusion of human fibroblasts or lymphocytes with mouse cell lines (Rag, LMTK⁻ or LTP) containing selectable markers (see references in Naylor *et al.*, 1983). Each hybrid line was karyotyped and assayed for isozyme markers at the time of DNA harvest (Shows *et al.*, 1982). Two series of hybrids segregated portions of chromosome 3: the TSL series was made from an individual with a reciprocal 3;17 translocation [GM2808-46,XX,t(3;17)(p21;p13)] and the XTR series from an individual with a reciprocal X;3 translocation [GM194-46,X,t(X;3)(q28;q21)] (Naylor *et al.*, 1982).

RESULTS

Purification of human β-galactosidase by column chromatography yielded a mixture of three proteins. The final separation of β-galactosidase was only accomplished after heating at 100°C for 5 min in a solution of 2% $NaDodSO_4$ and 5% 2-mercaptoethanol followed by HPLC separation. After either CNBr cleavage or trypsin digestion, amino acid sequence analysis of purified human β-galactosidase fragments yielded one CNBr fragment sequence and ten tryptic fragment sequences (Fig. 2). Many of the CNBr fragments and trypsin fragments did not yield enough protein to analyze, whereas some of the fragments were multiples upon further analysis and could not be used.

Two degenerate oligonucleotide probe mixtures were generated from appropriate regions of two of the amino acid sequences (Fig. 1). Screening of the human testicular cDNA library with the end-labeled probes yielded six positive clones, and sequencing of these clones revealed the presence of two different cDNA species (Fig. 3). Examination of the sequence of clone 23 revealed that a 485-bp region of clone 16 had been replaced by a 92-bp segment also coding for an uninterrupted open reading frame. The sequence of clone 16 was used to deduce the amino acid sequence of human β-galactosidase (Fig. 4). This sequence was nearly identical to that recently reported by Oshima *et al.* (1988).

The presence of two different cDNA sequences, both coding for uninterrupted open reading frames, led us to examine the mRNA species present. Northern blotting and hybridization using an RNA probe generated from clone 6 (Fig. 3) revealed the presence of a single RNA species of 2.45 kb in human lymphoblasts and human testicular tissue (Fig. 5).

Chromosomal localization was determined using clone 6 to screen a series of 28 hybrids. The human β-galactosidase locus maps to chromosome 3p21-3pter (Fig. 6).

FIG. 4. Nucleotide sequence of β-galactosidase cDNA and the deduced amino acid sequence. Lower-case letters indicate 5′ and 3′ untranslated sequence. Boxes denote differences in sequence from that reported by Oshima *et al.* (1988). The bold arrow (↓) demarcates the proposed signal sequence cleavage site. Asterisks (*) denote potential glycosylation sites. Numbered and underlined peptides are those determined by chemical sequencing of pure β-galactosidase (Fig. 2). The designation of amino acid 1 was made to the methionine encoded by the first ATG.

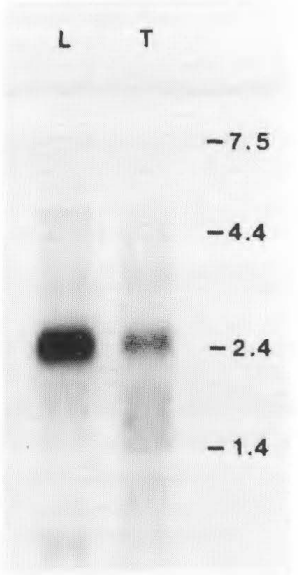


FIG. 5. Northern blot analysis of lymphoblast (L) and testicular (T) poly(A)RNA. Samples (1 μg) of poly(A)RNA were run on a 1% formaldehyde/agarose gel and blotted using labeled RNA from clone 6 as probe. Size standards are the 0.24 to 9.5-kb RNA ladder (Bethesda Research Laboratories).

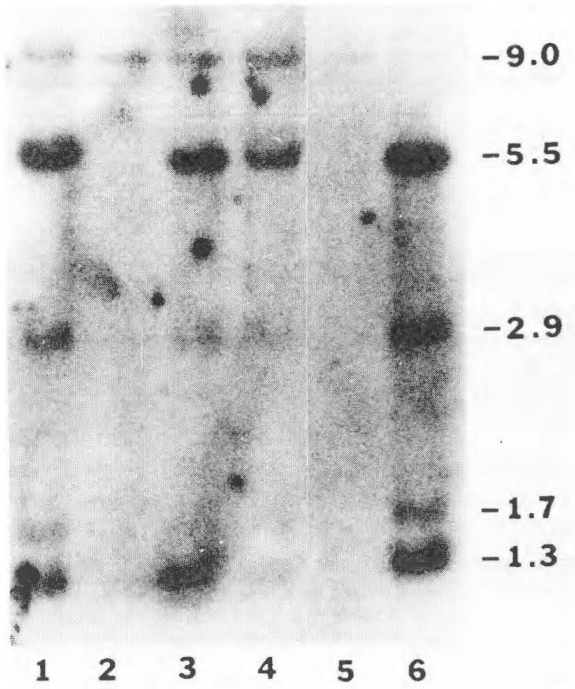


FIG. 6. Hybridization of the β-galactosidase clone 6 probe to human–mouse somatic cell hybrids. Human (lane 6) and mouse (lane 5) DNA digested with *Eco* RI yield distinguishable patterns. Somatic cell hybrids are either positive for the human fragments (lanes 1, 3, and 4) or negative (lane 2). Lane 1 is TSL-2, a hybrid containing the p21-pter region of chromosome 3. Lanes 2 and 3 have both parts of a reciprocal translocation: XTR-22 (lane 2) contains 3q21-3qter and XTR-3BSAGB (lane 3) contains 3pter-3q21. Lane 4 contains a hybrid whose only human chromosome is 3. These data place the 5.5-, 2.9-, and 1.3-kb *Eco* RI fragments hybridizing to clone 6 in the p21-pter region of chromosome 3.

DISCUSSION

Although the sequence of the cDNA for human β-galactosidase was reported recently (Oshima *et al.*, 1988), additional information on the β-galactosidase cDNA remains to be determined. Amino acid sequencing of CNBr and tryptic peptides confirms the identification of the cDNA sequence as that of human β-galactosidase (Figs. 2, 3, and 4). All 10 tryptic peptides and the single CNBr peptide are located in the amino acid sequence deduced from the nucleotide sequence of clone 16 (Figs. 3 and 4), which is identical to that reported by Oshima *et al.* (1988) except for a few base changes.

One of the nucleotide base differences we note occurs at nucleotide 89 (Fig. 4), numbered 63 in Oshima's *et al.* (1988) published sequence, where we found thymine rather than cytosine. This changes the amino acid at this position, which is within the signal peptide sequence, from proline to leucine. Another difference is at nucleotide 94, numbered 68 in Oshima's *et al.* (1988) sequence, where we found cytosine rather than thymine; no difference in amino acid sequence results because both triplets code for leucine. Beginning at nucleotide 660 there is a GC region where the order of 2 guanosine and 2 cytosine bases is unclear, most likely due to intrastrand base pairing that causes artifacts in the sequencing gel. Substituting dITP for dGTP in the sequencing reaction, we found the sequence in this region reads GCGC, resulting in the incorporation of amino acid arginine rather than alanine as noted in Oshima's *et al.* (1988) sequence. The sequence of the protein described in Fig. 4 was compared with sequences in the protein database and no significant homology was detected.

The presence of a second cDNA of slightly different structure opens the possibility of two different mRNA species coding for similar proteins. Northern blot analysis with a probe spanning a region common to both species (Fig. 3, clone 6) reveals the presence of only a single mRNA species of 2.45 kb in human lymphoblasts and human testicular tissue. On the basis of the present evidence, we believe that the larger cDNA species codes for the 2.45-kb mRNA transcript which is translated into human β-galactosidase. Because the sequence of the shorter clone 23 does not code for two of the peptides found during amino acid sequence analysis (Fig. 3, peptides 2 and 8), the mRNA from clone 16 must be responsible for the signal on the Northern blot. The significance of the shorter cDNA is unclear.

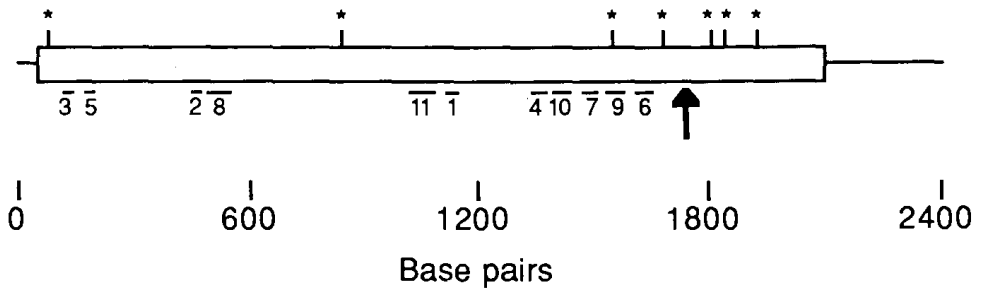


FIG. 7. Proposed structure of the β-galactosidase cDNA. The open box represents the open reading frame while the solid lines represent untranslated regions. The numbered dashes (—) under the cDNA structure indicate the location of the 11 sequenced peptides from Fig. 2. Stars (*) indicate potential glycosylation sites. The bold arrow (↑) indicates a potential protease cleavage site.

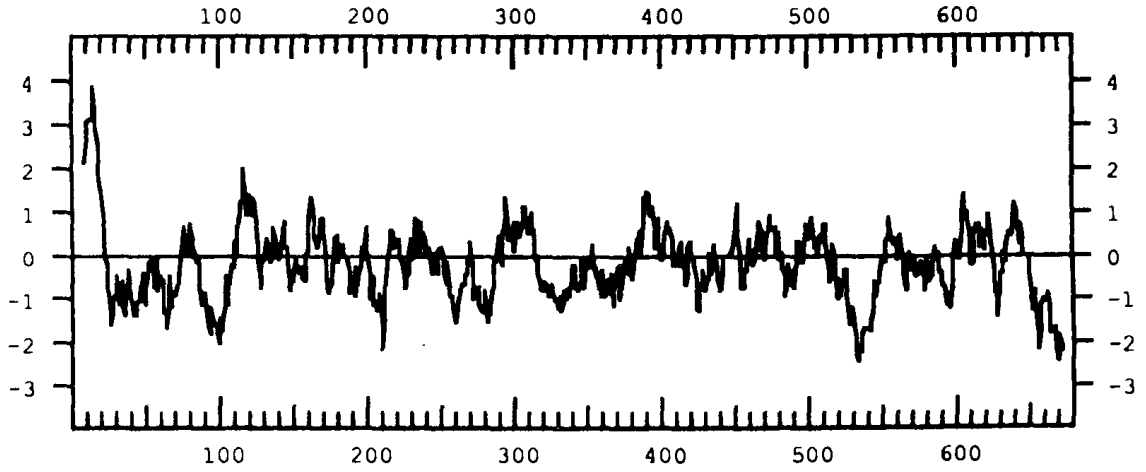


FIG. 8. Hydrophobicity plot of human β-galactosidase using the method of Kyte and Doolittle (1982). Sequence evaluated contains 677 amino acids. A potential cleavage site is present around amino acid 530 in a hydrophilic region.

Previously, Shows *et al.* (1979) used human-mouse hybrids and species-specific anti-β-galactosidase antiserum to identify the locus coding for human β-galactosidase and assigned the gene to chromosome 3. Naylor *et al.* (1982) used the same antibody to place the β-galactosidase locus in the region 3p21-3q21. However, TSL-6F, the hybrid cell line used to locate the gene previously exhibits a very weak hybridization signal with the β-galactosidase probe (clone 6). The expression of the human β-galactosidase protein in TSL-6F is probably too weak to be detected reliably. This accounts for the lack of immunoreactivity with TSL-6F in the previous study. Consequently, these new results are not in conflict with the earlier report.

Jones *et al.* (1984) have placed the β-galactosidase gene in the cen-pter region, and recently, Hertz *et al.* (1988) excluded 3p11-3p14.2. All the data are now consistent with the human β-galactosidase gene being located in the 3p21-3pter region.

It has been reported that β-galactosidase is initially synthesized as an 84-kD precursor which is processed to an intermediate form of 88 kD and then to a 64-kD mature enzyme form (Nanba *et al.*, 1988). We reported previously that the 64-kD mature enzyme contains 7.5% carbohydrate (Frost *et al.*, 1978) which gives a molecular weight of the polypeptide portion of the mature enzyme of 59 kD. The deduced amino acid sequence of β-galactosidase yields a protein of 73.5 kD after cleavage of the putative signal peptide (Oshima *et al.*, 1988), suggesting that proteolytic cleavage removes about 15 kD of polypeptide. Based on the positions of the peptides determined here from amino acid sequencing (Fig. 7), most of the proteolytic processing must occur at the carboxyl terminus of the precursor protein.

Although we found that the amino terminus of the mature enzyme was blocked, preventing sequence identification of the initial residues, one tryptic peptide was located

8 amino acid residues from the predicted signal cleavage site. No peptides were identified in the mature enzyme, that were derived from the carboxy-terminal residues (after residue 492) of the precursor sequence. When the deduced amino acid sequence of the precursor is analyzed on a Kyte-Doolittle hydrophobicity plot (Kyte and Doolittle, 1982), a large hydrophilic region is noticeable in the area of amino acid 530 (Fig. 8). If we assume that this is the region where proteolytic cleavage occurs, a protein of approximately 58 kD is generated which contains the entire region coding for the 11 peptides found by chemical sequencing (Figs. 4 and 7). Three potential glycosylation sites also are present in the region we believe is processed, which, if glycosylated, could account for some additional mass removed during processing (Figs. 4 and 7).

It has recently been suggested that the protective protein that stabilizes β-galactosidase may do so by proteolytic modification of β-galactosidase at its carboxyl terminus (Galjart *et al.*, 1988). The cDNA encoding the protective protein precursor has been isolated and sequenced (Galjart *et al.*, 1988). Extensive homology between the protective protein and the yeast proteases carboxypeptidase Y and KEX1 was found, including conserved stretches that contain the three active-site residues (Galjart *et al.*, 1988), suggesting that the former is also a serine protease. The carboxy cleavage region close to amino acid residue 530 in β-galactosidase may be the site where protective protein cleaves β-galactosidase to stabilize it and prevent its intralysosomal degradation.

ACKNOWLEDGMENTS

This work was funded by National Institutes of Health grants NS 08682-19 and HD 18983-4 to J.S. O'Brien and by the Mather Charitable Foundation, the Council for Tobacco Research, and the National Cancer Institute CA 44764 to S.L. Naylor. This project was supported by the Cooperative Human Tissue Network, which is funded by the National Cancer Institute.

REFERENCES

DAVIS, L.G., DIBNER, M.D., and BATTEY, J.F., eds. (1986). Screening a plated library with radiolabeled probes. In *Basic Methods in Molecular Biology*. (Elsevier, Amsterdam) pp. 185–189.

FROST, R.G., HOLMES, E.W., NORDEN, A.G.W., and O'BRIEN, J.S. (1978). Characterization of purified human liver acid β-D-galactosidase A_3 and A_4. Biochem. J. **175,** 181–188.

GALJART, N.J., GILLEMANS, N., HARRIS, A., VAN DER HORST, G.T.J., VERHEIJEN, F.W., GALJAARD, H., and D'AZZO, A. (1988). Expression of cDNA encoding the human "protective protein" associated with lysosomal β-galactosidase and neuraminidase: Homology to yeast proteases. Cell **54,** 755–764.

HERTZ, J.M., COERDT, W., HAHNEMANN, N., and SCHWARTZ, M. (1988). Interstitial deletion of the short arm of chromosome 3: Fetal pathology and exclusion of the gene for beta-galactosidase-1 (GLB-1) from 3(p11-p14.2). Hum. Genet. **79,** 389–391.

HOOGEVEEN, A.T., VERHEIJEN, F.W., and GALJAARD, H. (1983). The relation between human and lysosomal β-D-galactosidase and its protective protein. J. Biol. Chem. **258,** 12143–12146.

JONES, C., MILLER, Y.E., PALMER, D., MORSE, H., KIRBY, M., and PATTERSON, D. (1984). Regional mapping of human chromosome 3. Cytogenet. Cell Genet. **37,** 500.

KYTE, J., and DOOLITTLE, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157,** 105–132.

MANIATIS, T., FRITSCH, E.R., and SAMBROOK, J. (1982). In *Molecular Cloning: A Laboratory Manual.* (Cold Spring Harbor Laboratory, Cold Spring Hawrbor, New York).

MATSUDAIRA, P. (1987). Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262,** 10035–10038.

MATTEUCCI, M.D., and CARUTHERS, M.H. (1981). Synthesis of deoxyoligonucleotides on a polymer support. J. Am. Chem. Soc. **103,** 3185–3189.

MILLER, A.L., FROST, R.G., and O'BRIEN, J.S. (1976). Purification of human liver acid β-D-galactosidases using affinity chromatography. Anal. Biochem. **74,** 537–545.

NANBA, E., TSUJI, A., OMURA, K., and SUZUKI, Y. (1987). Galactosialidosis: A direct evidence that a 46-kilodalton protein restores deficient enzyme activities in fibroblasts. Biochem. Biophys. Res. Commun. **144,** 138–142.

NANBA, E., TSUJI, A., OMURA, K., and SUZUKI, Y. (1988). GM1-Gangliosidosis: Abnormalities in biosynthesis and early processing of β-D-galactosidase in fibroblasts. Biochem. Biophys. Res. Commun. **152,** 794–800.

NAYLOR, S.L., ELLIOTT, R.W., BROWN, J.A., and SHOWS, T.B. (1982). Mapping of aminoacylase-1 and β-galactosidase-A to homologous regions of human chromosome 3 and mouse chromosome 9 suggests location of additional genes. Am. J. Hum. Genet. **34,** 235–244.

NAYLOR, S.L., SAKAGUCHI, A.Y., SZOKA, P., HENDY, G.N., KRONENBERG, H.M., RICH, A., and SHOWS, T.B. (1983). Human parathyroid hormone gene (PTH) is on the short arm of human chromosome 11. Somat. Cell Genet. **9,** 609–616.

NORDEN, A.G.W., TENNANT, L.L., and O'BRIEN, J.S. (1974). GM1 Ganglioside β-galactosidase A: Purification and studies of the enzyme from human liver. J. Biol. Chem. **249,** 7969–7976.

O'BRIEN, J.S. (1989). Gangliosidosis and Mucolipidosis VI. In *The Metabolic Basis of Inherited Disease*, 5th ed. C. Scriver, A. Beaudet, W. Sly, and D. Valle, eds. (McGraw-Hill, New York) pp. 945–969.

OSHIMA, A., TSUJI, A., NAGAO, Y., SAKURABA, H., and SUZUKI, Y. (1988). Cloning, sequencing, and expression of cDNA for human β-galactosidase. Biochem. Biophys. Res. Commun. **157,** 238–244.

SHOWS, T.B., SAKAGUCHI, A.Y., and NAYLOR, S.L. (1982). Mapping the human genome, cloned genes, DNA polymorphisms, and inherited disease. Adv. Human Genet. **12,** 341–452.

SHOWS, T.B., SCRAFFORD-WOLFF, L.R., BROWN, J.A., and MEISLER, M. (1979). GM1-Gangliosidosis: Chromosome 3 assignment of the β-galactosidase-A gene (βGAL_A). Somat. Cell Genet. **5,** 147–158.

VERHEIJEN, F.W., PALMERI, S., HOOGEVEEN, A.T., and GALJAARD, H. (1985). Human placental neuraminidase: Activation, stabilization and association with β-D-galactosidase and its 'protective' protein. Eur. J. Biochem. **149,** 315–321.

YAMAMOTO, Y., FUJIE, M., and NISHIMURA, K. (1982). The interrelation between high- and low-molecular-weight forms of GM1-β-galactosidase purified from porcine spleen. J. Biochem. **92,** 13-21.

Address reprint requests to:
Dr. John S. O'Brien
Department of Neurosciences, M-034J
University of California, San Diego
La Jolla, CA 92093

Received for publication November 14, 1989.

NOTE ADDED IN PROOF

While this paper was in press, H. Morreau *et al.* (1989) (J. Biol. Chem. **264,** 20655–20663) reported the isolation and sequencing of two cDNAs encoding human β-galactosidase. They identified the same short cDNA as we discovered (the two sequences are identical) as well as the longer cDNA, and demonstrated that the shorter cDNA is produced by alternative splicing of β-galactosidase mRNA. Small amounts of mRNA corresponding to the shorter cDNA were detected in some human tissues.