This is to certify that the

dissertation entitled

ISOLATION AND CHARACTERIZATION OF BOVINE β-MANNOSIDASE cDNA

presented by

Hong Chen

has been accepted towards fulfillment
of the requirements for

Ph.D degree in Genetics

Major professor

Date June 27, 1994

ISOLATION AND CHARACTERIZATION OF BOVINE β-MANNOSIDASE cDNA

By

Hong Chen

A DISSERTATION

Submitted to
Michigan State university
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Genetics Interdepartmental Doctoral Program

1994

Karen Friderici

ABSTRACT

ISOLATION AND CHARACTERIZATION OF BOVINE β-MANNOSIDASE cDNA

By

Hong Chen

Lysosomal β-mannosidase is an acidic glycosidase involved in the degradation of N-linked glycoproteins. The deficiency of β-mannosidase activity results in an autosomal recessive inherited disorder, β-mannosidosis. This lysosomal storage disease has been found to cause a severe and fatal neurovisceral storage disorder in goats and cattle. The human counterpart is milder and exhibits extreme clinical heterogeneity.

The β-mannosidase cDNA had not been previously cloned from any species. Cloning and characterization of the normal β-mannosidase gene is essential in order to better understand the genetic defects underlying the β-mannosidoses, the regulation of expression of normal β-mannosidase, and the nature of the enzyme. In this dissertation, a variety of methods have been attempted to isolate a mammalian β-mannosidase cDNA. The availability of β-mannosidase peptide sequencing finally led to the isolation and characterization of a full-length bovine β-mannosidase cDNA. This cDNA contains a 3852-bp insert, comprising a 74-bp 5′ non-coding region, a 2637-bp coding

region encoding 879 amino acids, a 1141-bp 3′ non-coding region, and a 13-bp poly (A) tail. A 17-residue signal peptide sequence and possible polyadenylation signal sequences were identified. The deduced amino acid sequence was colinear with all peptide sequences determined by protein microsequencing. Northern analysis demonstrated a 4.2 kb single transcript in both normal and affected animals and in various tissues. The mRNA level was decreased in tissues from goats and cattle affected with β-mannosidosis, especially in the thyroid gland which contains the highest expression of β-mannosidase. The gene encoding β-mannosidosis was localized on human chromosome 4 by Southern analysis of rodent/human somatic cell hybrids.

This dissertation is dedicated to my parents, who have given me so much love and support in my whole life. Also to my dear husband, Wumin, and our lovely son Allan.

ACKNOWLEDGMENTS

I sincerely thank Dr. Rachel Fisher for accepting me as her graduate student and for her advice and support. Her generous financial support through the five-year period of my Ph.D studies made it possible for me to explore new research areas and work with a group of friendly people that I will never forget.

I thank Dr. Margaret Jones for being a member of my guidance committee. Her assistance, valuable suggestions, dedication to this project, and time spent reviewing my dissertation are deeply appreciated. I am also grateful for her financial support in the last year of my Ph.D studies and her friendship towards my family.

I thank Dr. Emanuel Hackel for his guidance and willingness to be the chairperson of my guidance committee. I thank, as well, Dr. Steven Triezenberg for his valuable discussions and suggestions.

I am greatly indebted to Dr. Karen Friderici. She was my advisor, my mentor, and my friend. She led me to master the molecular biology technology which will help my future career. I thank her for excellent academic advice, technical assistance, understanding, and encouragement

throughout the development of this dissertation. I appreciate her sharing the frustrations and joys of this project with me.

I thank Jeff Leipprandt (my co-worker) for his important contribution to this dissertation during the last crucial year. My appreciation extends to Dr. Kevin Cavanagh and Christine Traviss for their contributions, particularly, related to the work of amino acid sequencing analysis.

My thanks also go to Nancy Truscott, Jack Truscott, Dr. Susan Lootens, Dr. Kathy Lovell, Dr. Keiji Marushige, Dr. Bryce Sopher, Dr. Yasuko Marushige, Dr. P. Storto, Dr. Xiaotan Qiao, Ralph Common, Barbara Hoefler, Helen Wells, members of The AFP laboratory and the Department of Pathology. I was fortunate to have had these friendships and associations.

Studying in the Genetics Program and working in the Department of Pediatrics and Human Development, and particularly in the Department of Pathology at Michigan State University, will always remain memorable.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AP	alkaline phosphate
bp	base pair
cDNA	complementary DNA
CNBr	cyanogen bromide
Con A	concanavalin A
cpm	counts per minute
DMEM	Dulbecco's modified Eagle media
dNTP	deoxyribonucleoside triphosphate
DTT	dithiothreitol
EDTA	disodium ethylenediaminetetra-acetate
exoIII	exonuclease III
G_{m2}	GalNAcβ1→4(NeuAcα2→3)Galβ1→4Glcβ1→ceramide
GAPDH	glyceraldehyde-3-phosphate-dehydrogenase
HPLC	high performance liquid chromatography
Ig	immunoglobulin
IPTG	isopropyl-thiogalactopyranoside
MAbs	monoclonal antibodies
M-MLV	Moloney murine leukemia virus
MOPAC	mixed oligonucleotide primed amplification of cDNA
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
RACE	rapid amplification of cDNA ends
SDS	sodium dodecyl sulfate
TBS	tris-buffered saline
TMAC	tetramethylammonium chloride
TNT	Tris-NaCl-Tween 20
UAP	universal amplification primer

INTRODUCTION

INTRODUCTION

Lysosomal storage diseases are a group of inherited or induced disorders caused by deficient activity of one or several lysosomal enzymes (e.g. I-cell disease and multiple sulfatase deficiency) (Neufeld, 1991; von Figura *et al.*, 1984). The concept of lysosomal storage disease was first introduced in the context of α-glucosidase deficiency, known as the Pompe disease (Hers, 1965; 1973), on the basis of classical studies on the biochemistry of lysosomes. So far, three dozen lysosomal storage diseases have been discovered (Neufeld, 1991). The classification of lysosomal storage diseases is typically based on the main chemical type of accumulated material or the function of the defective proteins and enzyme involved (Watts and Gibbs, 1986; Neufeld, 1991). Currently, lysosomal storage diseases are divided into six major categories (Neufeld, 1991) comprising: (1) disorders of sphingolipid degradation; (2) disorders of glycoprotein degradation; (3) disorders of glycosaminoglycan degradation; (4) disorders of one single enzyme deficiency; (5) disorders of lysosomal enzyme biosynthesis; and (6) disorders of lysosomal membrane transport. Other classification methods, e.g. according to

the principles that cause the deficiency of the specific lysosomal enzyme activities, were also introduced (von Figura *et al.*, 1984).

Within each of the human lysosomal storage diseases, considerable clinical, biochemical, and molecular heterogeneity has been observed. The pathogenetic mechanisms for lysosomal storage diseases are still unclear. The storage of undegraded macromolecules in lysosomes enlarges and probably perturbs the physiology of the cell, which presumably contributes to the various clinical manifestations.

The impairment of a specific lysosomal enzymatic function is attributed to gene mutations that may (1) decrease either the amount or the rate of enzyme synthesis; (2) produce a catalytically inactive enzyme which either fails to bind its substrate or fails to catalyze reactions after binding to the substrate; (3) impair the transport of a newly synthesized enzyme into lysosomes; (4) increase the rate of enzyme degradation; or (5) decrease the concentration of a protecting or an activating factor (Watts and Gibbs, 1986). In addition, in some lysosomal storage diseases, i.e. mucopolysaccharidoses, the accumulated substrates interact with an enzyme outside the main pathways affected by the genetic deficiencies and cause a "secondary lysosomal disease" (von Figura *et al.*, 1984)

β-Mannosidosis, a recently defined inherited autosomal

recessive disorder (Jones and Laine, 1981; Jones and Dawson, 1981), is one of the lysosomal storage diseases. It is caused by decreased activity of β-mannosidase which is involved in glycoprotein catabolism. β-Mannosidosis was first described in Nubian goats and has more recently also been found in humans and Salers cattle (Hartley *et al.*, 1973; Healy *et al.*, 1981; Jones and Laine, 1981; Jones and Dawson, 1981; Bryan *et al.*, 1990; Jolly *et al.*, 1990; Abbitt *et al.*, 1991; Wenger *et al.*, 1986; Cooper *et al.*, 1986; 1991; Dorland *et al.*, 1988; Kleijer *et al.*, 1990; Wijburg *et al.*, 1991; Poenaru *et al.*, 1992; Levade *et al.*, 1991; 1994). The disease in goats and cattle has been extensively characterized in this laboratory (Jones and Laine, 1981; Jones and Dawson, 1981; Matsuura, Laine and Jones, 1981; Jones *et al.*, 1982; 1983; 1984; 1986; 1992; Masturra and Jones, 1985; Jones and Abbitt, 1993; Lovell and Jones, 1983; 1985; Lovell and Boyer, 1987; Lovell, 1990; Lovell *et al.*, 1991; 1994; Dahl *et al.*, 1986; Kumar *et al.*, 1986; Fisher *et al.*, 1987). β-Mannosidase from goats and cattle has also been purified and characterized (Cavanagh *et al.*, 1982; 1985; 1992; Dunstan *et al.*, 1983; Frei *et al.*, 1988; Sopher *et al.*, 1992; 1993). However, nothing is known about the molecular defect of this disease. The gene encoding β-mannosidase has not been cloned from any species, although an attempt was made to clone guinea pig β-mannosidase (McCabe and Dawson, 1990; Sopher, 1992a). The aim of this

project was to isolate β-mannosidase cDNA. The cloning of the β-mannosidase gene will assist with the characterization of the structure, function, and expression of the gene product. Identification of mutations may then be possible, and sequentially, gene therapy.

CHAPTER ONE

LITERATURE REVIEW

CHAPTER ONE

LITERATURE REVIEW

1.1 MOLECULAR ANALYSIS OF NORMAL LYSOSOMAL GENES

Cloning and characterization of normal lysosomal genes is a first step toward understanding the structure and function of lysosomal enzymes and toward the analysis of mutations underlying lysosomal storage diseases. Studies of these normal lysosomal genes and transcripts revealed some common characteristics.

1.1.1 The size and sequence features

Nearly 20 complementary DNAs encoding lysosomal enzymes have been cloned and characterized (Table 1.1) during the past ten years. The sizes of full-length cDNAs vary from 1.4 kb to 3.6 kb. The consensus sequence (GCCA/GCCATGG) (Kozak, 1986) for the translation initiation site was present in most of the lysosomal genes. It has been well demonstrated that within the consensus sequence, the purine at -3 position was most highly conserved (Kozak, 1986). This purine is also conserved in all the lysosomal enzymes except arylsulfatase A (Stein *et al.*, 1989). Other portions

Table 1.1 cDNA Clones of Normal Lysosomal Enzyme Genes

Genes	Disorders	cDNA Clones	mRNA	References
α-N-Acetyl-galactosaminidase	Schindler disease	2.2 kb** 3.6 kb	2.2 kb 3.6 kb	Wang *et al.*, 1990 Tsuji *et al.*, 1989 Yamauchi *et al.*, 1990
Acid phosphatase		2.1 kb	2.3 kb	Pohlmann *et al.*, 1988
Acid shingomyelinase	Niemann-Pick disease	2.3 kb**	2.5 kb	Quintern *et al.*, 1989 Schuchman *et al.*, 1991
Aspartyl-glucosaminidase	Aspartyl-glucosaminuria	2.1 kb	2.1 kb 3.4 kb	Ikonen et al., 1991
Arylsufatase A	Metachromatic leukodystrophy	1.6 kb	2.0 kb 3.9 kb 6.0 kb 4.8 kb	Stein *et al.*, 1990 Kreysing *et al.*, 1990
Arylsulfatase B	Maroteaux-Lamy syndrome	2.8 kb	1.8 kb 2.5 kb 4.8 kb	Schuchman *et al.*, 1990 Peters *et al.*, 1990
α-Fucosidase	Fucosidosis	2.0 kb 1.5 kb*	2.3 kb	Occhiodoro *et al.*, 1989 Kretz *et al.*, 1992 Fisher *et al.*, 1989
α-Galactosidase	Fabry disease	1.4 kb	1.4 kb	Bishop *et al.*, 1986; 1988
α-Glucosidase	Pompe disease	3.6 kb	3,6 kb	Hoefsloot *et al.*, 1988
α-L-iduronidase	Hurler and Scheie syndrome	2.2 kb*	2.2 kb* 2.8 kb*	Stoltzfus *et al.*, 1992*
		2.1 kb	2.3 kb	Scott *et al.*, 1991

Table 1.1 (cont'd.)

Genes	Disorders	cDNA Clones	mRNA	References
Galacto-cerebrosidase	Krabbe disease	?	?	Wenger *et al.*, 1993
β-Galactosidase	GM1 gangliosidosis	2.4 kb**	2.4 kb	Yamamoto *et al.*, 1990
	Morquio B		2.0 kb	Morreau *et al.*, 1989
		2.4 kb*	2.4 kb*	Oshima *et al.*, 1988
				Nanba *et al.*, 1990*
Glucocerebrosidase	Gaucher disease	2.4 kb	2.2 kb	Reiner *et al.*,1987; 1988
			2.6 kb	
			6.0 kb	
Glucosamine-6-sulfatase	Sanfilippo D	2.3 kb		Robertson *et al.*, 1992
β-Glucuronidase	Sly syndrome	2.4 kb*	2.6 kb*	Pail *et al.*, 1988*
				Watson *et al.*, 1985
		2.1 kb**		Oshima *et al.*, 1987
				Miller *et al.*, 1990
β-Hexosaminidase α-subunit	Tay-Sachs disease	1.9 kb**	2.0 kb	Korneluk *et al.*, 1986
			2.6 kb	Myerowitz *et al.*, 1985
β-Hexosaminidase β-subunit	Sandohoff disease	1.7 kb	2.2 kb	O'Dowd *et al.*, 1985
		1.7 kb*	2.0 kb*	Bapat *et al.*, 1988*
		1.7 kb*	1.8 kb*	Graham *et al.*, 1988*
Iduronate sulfatase	Hunter syndrome	2.3 kb	1.4 kb	Wilson *et al.*, 1990
			2.1 kb	
			5.4 kb	
			5.7 kb	

* Clones were isolated from non-human sources. **Clones with different structures have been isolated. Gene encoding lysosomal enzymes which did not cause a known disease e.g. proteases are not listed in this table.

of the consensus sequence are more variable. Like the secretory proteins, lysosomal enzymes contain a short signal peptide at the N-terminus, which typically contains 15-30 amino acids and has the following characteristic compositions (von Heijne, 1983; 1986; 1990): (a) a positively charged amino acid within the first five residues; (b) a hydrophobic core (7-15 residues); and (c) a more polar C-terminal region (3-7 residues). The signal peptide directs the nascent protein across the membrane and into the lumen of endoplasmic reticulum by interacting with a signal recognition particle (Kornfeld, 1987; Lewin, 1990). It is usually cleaved immediately after its transfer into the endoplasmic reticulum and before the synthesis of the peptide is completed. The typical signal sequence has been observed in all the lysosomal enzyme genes cloned.

All lysosomal enzymes studied so far are synthesized as polypeptides that undergoes limited proteolytic processing which may involve removal of a signal sequence, an N-terminal sequence, a C-terminal sequence, and/or internal cleavage (Kornfeld, 1987; Holtzman, 1989). Apart from the signal sequence cleavage, cleavage at the C-terminus seems to be the most common proteolytic processing event associated with the maturation of lysosomal precursors (Yamamoto *et al.*, 1990; Erickson and Blobel, 1983; Gottschalk *et al.*, 1988; Quinn *et al.*, 1987). To date, arylsulfatase A enzyme is the only lysosomal enzyme for

which post-translational proteolytic processing appears to be restricted to cleavage of the signal peptide only (Stein *et al.*, 1989). This was demonstrated by three lines of evidence: (1) the predicted molecular mass after removal of the signal peptide sequence was close to the size of deglycosylated arylsufatase A on SDS-PAGE; (2) an arylsulfatase A peptide corresponded to the predicted C-terminal sequences and lacked only the three residues located just before the stop codon (The three residues were thought to be cleaved by the staphylococcal V8 proteinase during the step of proteinase digestion in amino acid sequencing (Stein *et al.*, 1989)); and (3) the N-terminal sequence was found to follow immediately after the signal peptide sequence. Proteolytic processing is not essential for activation as it was found that the precursor forms of many lysosomal enzymes except proteases were active (Holtzman, 1989). It has been speculated that the precursor forms may be important for folding, stabilizing, or sorting the proteins in their early stages.

Most cDNAs encoding lysosomal enzymes have a 5′ untranslated region containing less than 600 bp compared to a 3′ untranslated region which is usually more than 500 bp long. cDNA containing 3′ untranslated regions as long as 2 kb in length has been documented for the human N-acetylgalactosidase gene (Wang *et al.*, 1990). The 3′ untranslated region is believed to be important to mRNA

turnover in eukaryotic cells (Ross *et al.*, 1988). However, this may not be true in all cases. The termination codon of human α-galactosidase transcript is followed immediately by a poly (A) tail without a 3′ untranslated region according to the analyses of both cDNA and genomic clones, which indicates that the 3′ untranslated region is not absolutely critical for its stability (Bishop *et al.*, 1986; 1988).

Like most eukaryotic mRNAs, all full length cDNAs of lysosomal enzyme genes possess a poly A tail at their 3′ termini. The function of the poly A tail is still not fully understood. It may affect mRNA stability and translation (Bernstein *et al.*, 1989; Proudfoot, 1991; Jackson *et al.*, 1990). The addition of a poly A tail occurs in the nucleus, and involves cleavage of the primary transcript with subsequent addition of a poly A tail to the newly formed 3′ end (Wickens, 1990). Polyadenylation requires the sequence AAUAAA. This hexanucleotide is usually located at 15-35 bp upstream of the poly A addition site of mRNAs and is highly conserved. Mutation analysis indicated that any base change in AAUAAA affects the accuracy and efficiency of cleavage and polyadenylation (Sheets *et al.*, 1990). However, polyadenylation signal sequences other than the consensus AAUAAA such as CUUAAA, AAUGAA, AAUAAC, or AUUAAA have been found in a few lysosomal enzyme genes (Stoltzfus *et al.*, 1992; Pohlmann *et al.*, 1988; Stein *et al.*, 1989; Proia *et al.*, 1986).

1.1.2. Sequence homologies

There is little, if any, homology among the cloned lysosomal genes at either the DNA or amino acid sequence levels. However, it has been observed that there are striking sequence similarities between lysosomal enzymes and non-lysosomal enzymes with similar catalytic functions. Sequence homology has been observed between lysosomal acid phosphatase and prostatic acid phosphatase (Pohlmann *et al.*, 1988); between α-glucosidase and intestinal sucrase-isomaltose complex (Hoefsloot *et al.*, 1988); and between lysosomal arylsulfatases A and B, arylsulfatase in the sea urchin, and steroid sulfatase (Stein *et al.*, 1989; Schuchman *et al.*, 1990; Sasaki *et al;* 1988). There is strong sequence conservation in lysosomal enzymes from different species including lower eukaryotes and even prokaryotes (Peters *et al.*, 1990; Schuchman *et al.*, 1990; Stein *et al.*, 1989; Wilson *et al.*, 1990; Wang *et al.*, 1990; Stoltzfus *et al.*, 1992). As summarized by Neufeld (1991), there are families of β-hexosaminidases, β-glucuronidases, α-galactosidases, α-glucosidases, phosphatases, sulfatases, α-L-iduronidases and other enzymes. Genes in each family presumably originated from the same ancestral gene by duplication and their divergence may reflect their different specificities regarding locations and substrates.

1.1.3. Alternative transcripts

In most eukaryotic genes, coding regions are interrupted by intervening sequences, known as introns. Introns have to be spliced out from the primary transcript to produce a mature mRNA. Splicing out of introns occurs by the formation of a lariat intermediate in which the left end of the intron is joined to a site near the right end of the intron (Lewin, 1990). Sequences at the left and right splice junctions appear to be conserved. The location of the branch point (i.e. the site near the right end of the intron) seems to play a role in the constitutive splicing (Smith *et al.*, 1989). Alternative splicing by various mechanisms is an efficient way to generate protein isoforms and may play important roles in gene regulation (Smith *et al.*, 1989). Transcripts that may be generated by alternative splicing have been reported for several lysosomal enzyme genes (Morreau *et al.*, 1989; Oshima *et al.*, 1987; Schuchman *et al.*, 1991; Wang *et al.*, 1990; Ikonen *et al.*, 1991; Scott *et al.*, 1991).

Two cDNAs with different lengths for β-galactosidase were isolated by Morreau *et al.* (1989). The short one lacked two stretches of coding sequences (212 bp and 181 bp) in the 5′ region. The reading frame was shifted by missing the 212 bp, but it was restored after exclusion of another 181 bp. The study of a genomic DNA sequence spanning the two missing areas demonstrated that the 212 bp sequence was

encoded by two separate exons of 151 and 61 bp. The 181 bp sequence was encoded in another exon. Furthermore, an exon specifying a 95 bp sequence between the two missing stretches was found. These results confirmed that the shorter β-galactosidase cDNA resulted from alternative splicing of its precursor mRNA by skipping two adjacent exons that encode a total of 212 bp and another exon specifying the 181 bp region. The short cDNA was only represented in anti-β-galactosidase antibody selected mRNA. No β-galactosidase activity was expressed by the corresponding protein (Morreau *et al.*, 1989).

Two cDNAs of the human α-N-acetylgalactosaminidase gene have been reported (Yamauchi *et al.*, 1990). The transcript encoding a cDNA which contained a 70-bp insertion in the coding region was present only in brain and in a small amount. The possibility of alternative splicing in this gene was studied by another group (Wang *et al.*, 1990). The region flanking the 70 bp insertion was amplified by PCR. cDNA transcribed from mRNAs in various tissues, cloned α-N-acetylgalactosaminidase cDNAs, and genomic clones were used as DNA templates in the PCR. No alternative transcripts in various tissues were observed. A different PCR product in one of cDNA clones was found. However, the PCR product contained an in-frame 45 bp deletion instead of the 70 bp insertion reported by Yamauchi *et al.* (1990). Strikingly, both the 70 bp insertion and the 45 bp deletion occurred at

the same intron. The investigators proposed that a unique sequence and/or secondary structure in this intron or surrounding region may impair the fidelity of the constitutive splicing.

Studies of the β-glucuronidase gene demonstrated that there were two types of mRNA corresponding to the two types of β-glucuronidase cDNA clones isolated from human placenta (Oshima *et al.*, 1987). Only the protein specified by the longer mRNA showed β-glucuronidase activity. Cloning and characterization of the genomic β-glucuronidase gene illustrated that there was an exon corresponding exactly to the 153-bp deletion in the shorter cDNA. This observation supported the previous hypothesis that the short transcript was generated from alternative splicing by exclusion of that exon (Miller *et al.*, 1990).

More recently, multiple transcripts generated from alternative splicing have been demonstrated for the human acid sphingomyelinase gene (Quintern *et al.*, 1989; Schuchman *et al.*, 1991). Three types of cDNA were isolated. Type 1 represented the majority of the acid sphingomyelinase clones identified and expressed lysosomal activity of sphingomyelinase in *COS* cells. It contained an insert of 1879 bp and possessed a unique 172 bp sequence encoding 57 amino acids. Type 2 contained an insert of 1382 bp and an unrelated 40 bp in-frame sequence that substituted for the 172 bp in-frame sequence presented in type 1. Type 3

contained a truncated open reading frame and lacked both the 172 bp and 40 bp sequences. The PCR results of a genomic sequence proved that the 172 bp was encoded by an exon, whereas the 40 bp sequence was of intronic origin. Furthermore, a weak donor splice site adjacent to the 172 bp exon was observed. The authors thought that the occurrence of type 2 cDNA may be due to the involvement of this cryptic 5′ donor splice site, which excluded the entire 172-bp exon and left 40 bp of intron sequence. The type 3 resulted from exon skipping, which excluded the 172-bp exon as well as the whole intronic sequence flanking the exon.

The use of a cryptic acceptor splice site within an exon may also be responsible for an alternative canine α-L-iduronidase transcript (Stoltzfus *et al.*, 1992). In addition to alternative splicing, multiple transcripts can arise from differential polyadenylation. This has been observed in iduronate-2-sulfatase (Wilson *et al.*, 1990), α-N-acetylgalactosaminidase (Wang *et al.*, 1990), and β-hexosaminidase (Proia *et al.*, 1987). The physiological functions of these alternative transcripts remain unclear. The proteins resulting from alternatively spliced transcripts do not have the same catalytic activity as the corresponding hydrolases (Oshima *et al.*, 1987; Morreau *et al.*, 1989).

1.1.4. Promoters

A promoter is a region of DNA involved in binding RNA polymerase to initiate transcription. "Housekeeping" genes produce proteins that have a wide tissue distribution and provide the essential functions needed for many types of cells (Dynan, 1986). The promoters for housekeeping genes typically have the following characteristic features: (1) no TATA box; (2) a high G/C content; and (3) the presence of CpG islands.

To date, about a dozen lysosomal enzyme genes have been isolated and characterized (Table 1.2). All except the β-glucocerebrosidase gene have promoters that are characteristic of housekeeping genes. The promoters of the human acid phosphatase gene (Geier *et al.*, 1989), the human arylsulfatase A gene (Kreysing *et al.*, 1990), the β-subunit of β-hexosaminidase (Neote *et al.*, 1988), the α-glucosidase gene (Martiniuk *et al.*, 1990; Hoefsloot *et al.*, 1990), the β-galactosidase (Suzuki *et al.*, 1991; Morreau *et al.*, 1991), and the α-L-iduronidase gene (Moskowitz *et al.*, 1992) all lack a TATA box and have a high G/C content with possible sp1 binding sites, i.e. the GGCGGG motif. Multiple initiation sites of transcription were observed in promoters of the acid phosphatase gene, the arylsulfatase A gene, and the mouse β-galactosidase gene. This is typical of promoters lacking a TATA box. Some lysosomal enzyme genes whose promoters are G/C rich do contain a TATA box. These

Table 1.2 Putative Promoter Regions of Genes Encoding Lysosomal Enzymes *

Genes	Transcription start site(s)	Percentage of GC	TATA	CAAT	Sp1	Others		References
α-N-Acetyl-galactosaminidase	-347	56	none	-678 -644	-364 -410 -596	IR	-354 -410	Wang *et al.*, 1991
Acid phosphatase	--6 to -23	59	none	-150 -380 -450	-60 -326			Geier *et al.*, 1989
Acid sphingomyelinase	nd	63	-863 -894	-704 -867	-150 -256 -267 -285 -715	Ap1 NK1	-604 -461 -583	Schuchman *et al.*, 1992
Arylsulfatase A	-367 to -387		none	none	-10			Kreysing *et al.*, 1990
α-L-Fucosidase	nd		none	none	-33 -70 -89 -131 -164			Kretz *et al.*, 1992
α-Galactosidase A	-60	59	-86 -93 -102 -129	-71 -104 -146 -178 -203	-63 -207	AP1 OCTA	-153 -835 -889	Quinn *et al.*, 1987
α-Glucosidase	-220	80	none	-262 -107 -180	-29	AP2	-316 -287	Hoefsloot *et al.*, 1990

Table 1.2 (cont'd)*

Genes	Transcription start site(s)	Percentage of GC	TATA	CAAT	Sp1	Others		References
β-Galactosidase	-53 to -56	71	none	-309 -132 -153 -163 -556	-80	AP2	-166	Morreau *et al.* 1991
Glucocerebrosidase	--657		-680 -691	-754 -760	none			Horowitz *et al.*, 1989 Reiner *et al.*, 1988
β-Glucuronidase	-30 -126	72	none	none	-248 -64	AP2	-164	Shipley *et al.* 1991
β-Hexosaminidase	nd		none	-381	-59 -123 -169 -338	AP1	-99 -107	Neote *et al.*, 1988
Iduronidase	-177		none	none	-150 -171 -185 -214 -232 -297 -350			Moskowitz *et al.*, 1992

* This table is expanded from Schuchman's table (Schuchman *et al.*, 1992)

include the α-subunit of the β-hexosaminidase (Proia *et al.*, 1987), α-galactosidase (Bishop *et al.*, 1988), acid sphingomyelinase (Schuchman *et al.*, 1992), and murine β-glucuronidase (D'Amore *et al.*, 1988) genes. The human glucocerebrosidase gene is not a typical housekeeping gene. Its promoter contains a TATA box and CAAT-like box, and lacks the sp1 binding site (Reiner *et al.*, 1988a; Grabowski *et al.*, 1990). TATA and CAAT-like boxes are common in regulated genes. Reiner *et al.* (1988b) reported that the glucocerebrosidase gene had different levels of expression in a variety of cell types and the level of mRNA was well correlated to the level of chloramphenicol acetyltransferase (CAT) activity directed by the promoter in different cells.

1.1.5 Pseudogenes

Pseudogenes are defined by stable genomic sequences that are related to those of functional genes, but they can not be translated into functional proteins. There are two types of pseudogenes. The first type includes those that retain the introns found in their functional counterparts. They are typically derived by duplications and mutations of ancestral active genes. The second type are so called "processed" pseudogenes. They contain no intervening sequences and resemble the RNA transcripts of their active counterparts. Processed pseudogenes presumably resulted from integration of their mRNAs or cDNA copies into the

genome. They may be located anywhere in the genome. Reiner *et al.* (1988) reported two different genomic clones for glucocerebrosidase gene that had different lengths. These two clones were later found to have 96% identity at the nucleotide level and were located on the same chromosome (Horowitz *et al.*, 1989). However, the promoter of the shorter clone showed very low or negligible activity when coupled to a bacterial gene. So far, no protein product has been found to be produced by the shorter gene. This pseudogene lost the open reading frame due to a large deletion in several introns and exons. Presumably, most pseudogenes do not have transcriptional and translational activities. However, some exceptions were reported in which pseudogenes had transcripts or protein products (Sorge *et al.*, 1990). The glucocerebrosidase pseudogene was found to be transcribed at a similar level to its active gene counterpart (Sorge *et al.*, 1990). Extra precaution is needed in molecular diagnosis by PCR due to the presence of high levels of transcripts of pseudogenes.

Besides the glucocerebrosidase pseudogene, a pseudogene for human α-L-fucosidase gene was discovered by O'Brien's group (Kretz *et al.*, 1992). The α-L-fucosidase pseudogene has 80% identity with the active gene but is located on a different chromosome than the functional gene. The α-L-fucosidase pseudogene contains no introns and the sequence diverges at the beginning and the end of the transcript of

the functional gene. These phenomena are considered to be part of the general structural characteristics of the processed pseuogenes (Vanin, 1985). However, the α-L-fucosidase pseudogene lacks a poly A tract at the 3′ end and the usual direct repeats flanking a pseudogene sequence.

More recently, multiple pseuogenes for human β-glucuronidase gene were identified by PCR when mutation analysis was carried out in a patient with mucopolysaccharidosis (Shipley *et al.*, 1991; Vervoort *et al.*, 1993). These pseudogenes shared high sequence homologies with the functional counterparts, but the normal reading frame was disrupted in most cases. Some of the β-glucuronidase pseudogenes containing uninterrupted open reading frames may, in fact, represent gene families with closely related functions. Amplification of genomic DNA from a panel of human/rodent somatic cell hybrid lines demonstrated that the multiple unprocessed pseudogenes for human β-glucuronidase were located on six different chromosomes. Three types of β-glucuronidase pseudogenes were identified with respect to the exon 11 region. Each of these pseudogenes was found to be located in two different chromosomes: type 1 was located in chromosome 5 and 6, type 2 in chromosome 22 and 22, and type 3 in chromosome 7 and Y.

A conflict arose regarding the chromosomal location of the gene encoding the G_{m2} activator protein. In 1985, Burg

et al. (1985) mapped the gene for human G_{m2} activator protein to chromosome 5 by expressing the protein in human /mouse somatic cell hybrids. However, in 1991, Kleyn *et al.* (1991) using PCR analysis reported that the gene for the G_{m2} activator protein was not located on chromosome 5. This discrepancy was elegantly explained by the discovery of a processed pseudogene coding for the G_{m2} activator protein (Xie *et al.*, 1992). A 7300-bp intron between nucleotides G_{243} and G_{244} of the cDNA was present in the functional gene but missing in the pseudogene. The primers chosen by Kleyn *et al.* (1991) happened to flank the intron region. Therefore, amplification of genomic DNAs by these primers could only generate products originating from the pseudogene but not from the functional gene due to the presence of a large sized intron. By using primers specific to the functional gene or the pseudogene, Xie *et al.* (1992) were able to place the functional gene on chromosome 5 and the pseudogene on chromosome 3. These studies illustrate that caution is required in mutation analysis of genomic DNA. However, it is still possible to use genomic DNA for mutation analysis even if a pseudogene is present.

1.2 STRATEGIES FOR THE MOLECULAR CLONING OF LYSOSOMAL ENZYME GENES

In order to better understand the fundamental abnormalities of lysosomal storage diseases at the gene level, it is necessary to isolate and characterize the normal structural gene which codes for the specific lysosomal enzyme or enzyme protector. Not only will the molecular cloning of the normal lysosomal gene facilitate the understanding of the structure, function, and expression of the gene products, but it will also lead to the possibility of molecular diagnosis and gene therapy. Lysosomal enzymes are encoded by mRNAs that are present at low abundance (O'Brien *et al.*, 1984). Therefore, cloning of the genes for these enzymes has proven to be difficult. Nevertheless, thanks to the rapid progress in molecular cloning techniques, 16 lysosomal enzyme genes whose deficiencies lead to known lysosomal storage diseases have been isolated (Table 1.1). The isolation of cloned cDNA involves several major steps: enrichment of mRNA, synthesis of cDNA, insertion into a cloning vector, and identification of a clone. This review is to discuss the various techniques involved in the final step of molecular cloning.

1.2.1 Identification of a cDNA clone by hybrid-selected translation

β-Glucuronidase, one of the first lysosomal enzymes to be cloned, was isolated by the technique of hybrid-select translation (Hieber, 1982). First, a cDNA library was screened using a population of mRNAs as probes. The probes were either total mRNA fractions or β-glucuronidase-enriched mRNA. DNA was isolated from positive clones, bound to nitrocellulose filters and hybridized with the population of mRNAs. After washing, the bound mRNAs were eluted from the filters. The selected mRNAs were used to direct cell-free protein synthesis in a rabbit reticulocyte lysate system. Positive clones were identified by looking for bands of total *in vitro* translated protein which had the same mobility as the immunoprecipitated and translated proteins on SDS-polyacrylamide gel (PAGE). The same technique has been used by other researchers for verifing clones coding for α-glucosidase (Konings, 1984) and the α-chain of β-hexosaminidase (Myerowitz *et al.*, 1984).

Hybrid-selected translation is a valuable technique both as a screening tool and as an independent means of confirming the authenticity of a clone. However, this method is time- and labor-consuming. In addition, it may be inefficient for large mRNA because of the possible degradation of mRNA and inefficient transcription (Horwich *et al.*, 1984). Today, this technique is mainly used to

verify positive clones isolated by other methods.

1.2.2 Nucleic acid hybridization

When a cDNA probe is available, it is often used to screen cDNA libraries from the same species to isolate the full-length cDNA, or it is used to screen cDNA libraries from other species. High or relatively high stringency hybridization conditions can usually be applied when homologous or partially homologous cDNA probes are used. Therefore, screening with cDNA probes seldom has the problem of false positive clones. cDNA probes synthesized from a population of mRNAs can also be used to screen clones. However, for rare genes, it is necessary to enrich the mRNA before the preparation of cDNA probes. β-Glucuronidase (Nishimura *et al.*, 1986), α-glucosidase (Konings, 1984), and the α-chain of β-hexosaminidase (Myerowitz *et al.*, 1984) were successfully isolated by differential hybridization with cDNA probes prepared by reverse transcription of both immunoselected and depleted polysomal mRNA.

1.*2.3 Identification of genes by antibody probes*

The technique of screening cDNA libraries with antibodies was introduced after the development of the phage λgt11 as an expression vector (Young *et al.*, 1983). In this system, double stranded cDNA is inserted into a restriction site in the *E. coli* lacZ (β-galactosidase) gene carried by

λgt11, λZAPII, or other expression vectors. If the foreign DNA is inserted in the correct orientation and reading frame, a fusion protein will be produced with β-galactosidase at its amino terminal and the foreign protein at the carboxyl terminus. To isolate cDNA clones by immunoscreening, *E. coli* cells transformed with recombinant phage from a cDNA expression library are plated and induced to make fusion proteins when phage plaques are just visible. The plates are overlaid with nitrocellulose filters to absorb the fusion proteins. Immunoreactions are carried out by sequential incubation of filters with a primary antibody specific for the enzyme being cloned and a secondary antibody against the primary one. Finally, positive clones are identified by either chromogenic immunodetection or auto-radiography depending on the secondary antibody. By screening a human hepatoma cDNA library constructed in a λgt11 expression vector, O'Brien's group was able to isolate partial cDNAs encoding lysosomal enzymes α-fucosidase, galactosidase, the α-chain of β-hexosaminidase (O'Brien *et al.*, 1984; de Wet *et al.*, 1984). This technique was also used by other groups to identify genes encoding lysosomal enzymes such as acid phosphatase, α-glucosidase, α-galactosidase A, α-fucosidase, and β-glucocerebrosidase (Pohlmann *et al.*, 1988; Hoefsloot *et al.*, 1988; Martiniuk *et al.*, 1986; Calhoun *et al.*, 1985., Sorge *et al.*, 1985., Ginns *et al.*, 1984; Fisher *et al.*, 1989; Fukushima *et al.*, 1985).

There are several problems which may affect the success of immunoscreening. First, most human and rabbit antisera usually contain significant amounts of antibodies against the *E. coli* or phage proteins (Snyder *et al.*, 1987). These nonspecific antibodies can be removed by pseudoscreening, i.e., by preincubating the diluted antisera with filters coated by lysed *E. coli* for several times or, more efficiently, by passing the polyclonal antisera through an affinity column to which a large amount of lysed *E. coli* proteins has been coupled (Sambrook *et al.*, 1989). Secondly, false positive clones may be detected by contaminating antibodies generated from impure proteins. The use of affinity purified antisera may overcome this problem. However, any protein possessing a common epitope to the protein of interest may still be isolated. Monoclonal antibodies usually generate less background than polyclonal antibodies. However, screening with polyclonal antibodies has a higher possibility of isolating a gene since it recognizes multiple epitopes. All in all, the ideal probes are pools of monoclonal antibodies directed against different epitopes of the protein of interest. Although antibody screening is frequently used, the success of this approach relies not only on the quality of antibody used, but also the abundance of the protein expressed. Generation of monoclonal or polyclonal antisera is labor-intensive and time-consuming. More plaques or colonies may

need to be screened in the isolation of a gene encoded in very low abundance compared with other methods, since only one out of six recombinant clones, in theory, will produce fusion proteins and react with antibody. The most difficult and tedious part of isolating genes by antibody screening is to verify the immunoreactive clones. A cDNA selected by antibody probes is by no means a true positive. Failure to isolate genes by antibody screening was documented (Moreman, 1989) and is not uncommon.

1.2.4 Identification of cDNA clones by oligonucleotides

Using degenerate oligonucleotides to screen cDNA libraries is currently one of the most commonly employed techniques for cloning genes. To apply this technique, a portion of the amino acid sequence of the protein is required. Based on the known peptide sequence information, either a set of short degenerate oligonucleotides or a unique long guessmer can be synthesized chemically.

Short degenerate oligonucleotides are usually 17 to 20 nucleotides long and include all possible codon choices for each amino acid. Only one sequence in the mixture will perfectly match the coding region in the cDNA. The disadvantages of using short degenerate oligonucleotides are: (1) An amino acid peptide with low degeneracy is necessary in order to synthesize an oligonucleotide mixture with low complexity. The more complicated the

oligonucleotide mixture is, the more false positives will be isolated. (2) The sequence information of amino acid peptides should be correct. A single mismatch of nucleotides will substantially affect the hybridization of the probe.

The long guessmer is typically 30 to 100 nucleotides long with a unique sequence. The best guesses are made for each amino acid based on the typical codon usage. The guessmer overcomes the disadvantages of short degenerate oligonucleotides because the long length in the guessmer compensates the possible effect of nucleotide mismatches. Thus the amino acid sequence need not be absolutely correct and amino acids with high codon degeneracy (e.g. arginine, leucine, and serine) need not be avoided when designing a guessmer. However, at least ten amino acid residues are needed for synthesizing a guessmer.

Screening with oligonucleotides is done by either plaque or colony hybridization. It appears that plaque hybridization with short mixed oligonucleotides was seldom used because there is less DNA per plaque than per bacterial colony. For short oligonucleotides, however, the recombinant plaques may be amplified *in situ* before hybridization. By amplification of plaques *in situ*, there are more copies of the phage DNA per plaque fixed on the filters. Therefore, the intensity of hybridization signals is increased substantially. Generally, it is not necessary

to prepare duplicate filters after *in situ* amplification. However, the phage need to be plated at much lower density than in the unamplified method (Wozney, 1990).

In some cases, tetramethylammonium chloride (TMAC) has been used as the hybridization solvent. In TMAC, the melting temperature of an oligonucleotide is independent of the oligonucleotide sequence, i.e. base composition, and is a function of probe length (Wood *et al.*, 1985; Jacobs *et al.*, 1988). Therefore, by hybridization in TMAC, the preferential hybridization of G/C rich sequences may be abolished.

Most of the lysosomal enzyme genes isolated by oligonucleotides were screened by plaque hybridization with unique oligonucleotide guessmers. They include arylsulfatase A (Stein *et al.*, 1989), arylsulfatase B (Peters *et al.*, 1990), α-L-iduronidase (Scott *et al.*, 1991; Stoltzfus *et al.*, 1992), iduronate-2-sulfatase (Wilson *et al.*, 1990), cathepsin B (Segundo *et al.*, 1985), N-acetylgalactosaminidase (Tsuji *et al.*, 1989), aspartylglucosaminidase (Ikonen *et al.*, 1991); and α-mannosidase (Schatzle *et al.*, 1992). Colony hybridization with a short oligonucleotide mixture was used in the cloning of α-N-acetylgalactosaminidase (Wang *et al.*, 1990), the β-subunit of β-hexosaminidase (O'Dowd *et al.*, 1985), the α-subunit of β-hexosaminidase (Korneluk *et al.*,1986), co-β-glucosidase (Rorman *et al.*, 1989), and sphingomyelinase

(Quintern *et al.*, 1989). Glucosamine-6-sulfatase was the only lysosomal enzyme gene isolated by plaque hybridization with short mixed oligonucleotides (Robertson *et al.*, 1988).

1.2.5 Isolation of genes by polymerase chain reaction (PCR)

False positive clones are often encountered when screening with degenerate oligonucleotides, while screening with guessmers require a decrease in the stringency of hybridization due to codon uncertainty. Recently, molecular cloning of genes using PCR techniques has been introduced, which diminishes the disadvantages of these conventional methods described above. A mixed oligonucleotide primed amplification of cDNA (MOPAC) method was first introduced by Lee *et al* (1988). Unlike conventional PCR, the amplification reaction in this system is primed by degenerate oligonucleotides specified by peptide sequences rather than by two unique oligonucleotides. The template is either cDNA synthesized by reverse transcription or DNA prepared from cDNA libraries. By the MOPAC technique, specific PCR products can be generated, which can then be used as unique probes for screening DNA libraries. The advantage of screening cDNA libraries using longer, unique nucleic acid probes instead of mixed oligonucleotides or guessmers is obvious, i.e. hybridization can be carried out in high stringency conditions so that fewer false positive clones are produced. However, a relatively long piece of

amino acid sequence, or the information regarding relative orientation of two different peptide sequences is generally required so that a sense and an antisense mixed oligonucleotide based on the amino acid sequences can be designed. Multiple PCR products are expected due to the use of mixed oligonucleotides.

Mixed oligonucleotides with low degeneracy are recommended in order to increase the specificity of PCR. However, priming by mixed oligonucleotides with as high as 8200 fold degeneracy is possible (Ottolie *et al.*, 1991). The authenticity of PCR products can be confirmed by the following methods: (1) Southern hybridization if an internal mixed oligonucleotide that lies between two primers used in the PCR reaction is available; (2) the size selection if the length between the two PCR primers is known; and (3) sequencing (Lee *et al.*, 1988; 1991; Moremen, 1989; Schuchman *et al.*, 1990). Other ways of identifying the complex PCR products are possible (Strub *et al.*, 1989). In addition to the generation of a unique DNA fragment as a probe, the MOPAC method can also be used for direct production of a portion of the DNA using either two gene specific oligonucleotides (Moreman, 1989) or one gene specific oligonucleotide primer and a vector primer (Gonzalez and Chan, 1993; Ikonen *et al.*, 1991).

The MOPAC technique has been used successfully in the isolation of a number of genes including the gene encoding

lysosomal arylsulfatase B enzyme and the gene encoding the endoplasmic reticulum-targeting protein of β-glucuronidase (Schuchman *et al.*, 1990; Ovnic *et al.*, 1991; Camirand *et al.*, 1991; Bischoff *et al.*, 1990; Moremen, 1989). More recently, a lysosomal gene encoding galactocerebrosidase enzyme was isolated by screening with a PCR product generated by the MOPAC technique (Wenger *et al.*, 1993). This approach is especially useful for cloning members of gene families.

Conventional screening of cDNA libraries is time- and labor-consuming, especially when the gene is expressed at a very low level, necessitating the screening of a large number of clones. This problem can be overcome by using the MOPAC technique for diagnostic screening, i.e. a group of minilysates, each containing a certain number of phage plaques from a library, is first prepared, followed by amplification of DNA with gene specific primers. Library subsets represent by positive minilysates are then screened (Moremen, 1989).

Recently, a PCR method, called rapid amplification of cDNA ends (RACE) (also known as anchored- or one-sided PCR) has been developed (Frohman *et al.*, 1988; 1993; Ochara *et al.*, 1989). This method has been used for rapid cloning of the 3′ or 5′ end of cDNA (Frohman *et al.*, 1988; Casella *et al.*, 1989). By conventional screening, it may take weeks or months to screen a cDNA library and to isolate and analyze

candidate cDNAs in order to clone a full-length cDNA. Using the RACE method, a gene specific primer based on a portion of internal DNA sequence is required. In both 5′ and 3′ RACE systems, mRNAs are first copied by reverse transcription, and the first strand cDNAs are then amplified. The amplification is accomplished by two primers: a gene specific oligonucleotide and a synthetic homopolymer tail (poly-T for the 3′ RACE; poly-T or G for the 5′ RACE) tagged by restriction sites. For the 5′ RACE system, a terminal deoxynucleotidyl transferase (TdT) tailing reaction is required before amplification. The RACE technique provides a rapid and efficient way not only for cloning full-length cDNAs, but also for obtaining information regarding alternative splicing, alternative polyadenylation, or alternative promoters. Kits for RACE are commercially available.

1.3 β-MANNOSIDOSIS

β-Mannosidosis is an inherited glycoprotein storage disorder. It was first described in Nubian goats (Jones *et al.*, 1981a; Hartley and Blakeman, 1973; Healey *et al.*, 1981). Affected goats all have the following phenotypic features: inability to stand, dome-shaped skulls, carpal contractures, pastern joint hyperextension, narrowed

palpebral fissures, deafness, and intention tremor (Jones *et al.*, 1982, 1983; Kumar *et al.*, 1986). Pathologically, the affected goat displayed widespread cytoplasmic vacuolation in a variety of cell types of the central nervous system and viscera as well as central nervous system axonal spheroids and myelin deficits (Jones *et al.*, 1983; Lovell *et al.*, 1983). Goats with β-mannosidosis were found to have a deficiency of the enzyme β-mannosidase in plasma and various tissues such as kidney, brain, liver, and skin fibroblasts (Jones and Dawson, 1981; Jones *et al.*, 1984). The main storage products associated with the enzyme deficiency are a disaccharide (Manβ1-4GlcNAc) and a trisaccharide (Manβ1-4GlcNAcβ1-4GlcNAc) (Jones and Laine, 1981; Matsuura *et al.*, 1981; Jones *et al.*, 1983; 1984). Minor accumulations of tetrasaccharide and pentasaccharide were also observed (Matsuura and Jones, 1985). Studies of accumulated oligosaccharides suggested that the di- and tri-saccharides were generated independently (Hancock *et al.*, 1986). Two catabolic pathways involving the activities of an endo-β-N-acetylglucosaminidase and an aspartylglucosaminidase in goat are partially responsible for the heterogeneity in storage and excreted material (Hancock and Dawson, 1987).

In 1990, β-mannosidosis was identified in Salers calves in Australia, New Zealand, and the U.S.A. (Jolly *et al.*, 1990; Abbitt *et al.*, 1991; 1990; Bryan *et al.*, 1990; Orr, 1990). The clinical, pathological, and biochemical features

of affected calves are very similar to these of affected goats and include: facial dysmorphism, inability to stand, intention tremors, profound central nervous system dysmyelination, extensive tissue cytoplasmic vacuolation, di- and tri-saccharide accumulation in tissues and urine, and marked deficiency of β-mannosidase activity in plasma and various tissues. However, affected cattle had less hearing defect, but more profound enlargement of the kidney and thyroid (Bryan *et al.*, 1990; Abbitt *et al.*, 1991). The latter difference is believed to be due to the longer gestation period of the bovine species and thus more accumulation of the undegraded or uncatabolized oligosaccharides (Abbitt *et al.*, 1991; Jones and Abbitt, 1993). β-Mannosidosis in the caprine and bovine species has a rapidly fatal course. Affected animals die in the neonatal period if intensive care is not administered.

In 1986, Cooper *et al.* and Wenger *et al.* described the first cases of human β-mannosidase deficiency. Cooper *et al.* (1986) described 44-year-old and 19-year-old brothers. Wenger *et al* (1986) reported a four-year-old boy. In both cases, the patients suffered from mental retardation and hearing loss but no other neurological symptoms. Coarse facial features were observed in Wenger's patient. There was no detectable β-mannosidase activity in leukocytes, fibroblasts, and plasma in these patients. The four-year-old boy also had heparin sulfamidase deficiency. Excess

disaccharide was observed in these patients' urine. So far, eleven patients from eight families have been diagnosed with β-mannosidosis (Cooper *et al.*, 1986; 1991; Wenger *et al.*, 1986; Dorland *et al.*, 1988; Kleijer *et al.*, 1990; Wijburg *et al.*, 1991; Poenaru *et al.*, 1992; Levade *et al.*, 1991; 1994). There is great clinical heterogeneity among these patients. The most characteristic symptoms are mental retardation and deafness. The course of β-mannosidosis in human is generally much milder than that in either ruminant form, although three patients died before age 20. More recently, Levade *et al.* (1994) reported a 14-year-old patient with complete deficiency of β-mannosidase activity in plasma and leukocytes, but no mental retardation or hearing loss were revealed. Moreover, there was no detectable disaccharide in the patient's urine. The major storage product in other patients is disaccharide rather than the trisaccharide which is the major accumulating oligosaccharide in ruminants. Recently, a new oligosaccharide complex (sialyl-α(2→6)-mannosyl-β(1→4)-N-acetylglucosamine), resulting from an α2-6-sialylation of the accumulated disaccharide product, was isolated from the urine of a patient with β-mannosidosis (Hokke *et al.*, 1990; van Pelt *et al.*, 1990). The more complex oligosaccharides similar to those accumulated in affected animals have not been found in human cases. The striking biochemical differences between ruminant and human β-mannosidosis (i.e. accumulation of trisaccharide versus

disaccharide) are probably due to the different catabolic pathways for glycoproteins (Hancock *et al.*, 1986; Jones, 1992). In ruminants, removal of the reducing-end GlcNAc from oligosaccharides that have been cleaved by an amidohydrolase is catalyzed by β-hexosaminidase, while in human and rats, lysosomal chitobiase is responsible for this activity (Aronson and Kuranda, 1989; Hancock *et al.*, 1986).

Dysmyelination in the central nervous system is a consistent feature among affected β-mannosidosis animals. The degree of myelin deficits in regions of the central nervous system varies and it correlates with the time of myelination and with more severe myelin paucity in regions that develop myelin at a later stage (Lovell *et al.*, 1983; 1987; Patterson *et al.*, 1991). The dysmyelination appears to be associated with a defect in oligodendrocytes (Boyer *et al.*, 1990; Lovell *et al.*, 1987). The observation of profound cytoplasmic vacuolation in thyroid and decreased thyroid hormone in affected animals raises the possibility that hypothyroidism may play a role in the central nervous system dysmyelination (Boyer *et al.*, 1990b; Lovell *et al.*, 1991). However, the factors responsible for the severity of dysmyelination in various regions of the central nervous system are still unclear.

β-Mannosidosis is identified by its characteristic clinical manifestations and urinary excretion of di- and tri-saccharides. The final diagnosis relies on the

deficiency of β-mannosidase activity in plasma, lymphocytes, or tissues. Intermediate levels of β-mannosidase activity are generally observed in heterozygotes (Jones *et al.*, 1981; Cavanagh *et al.*, 1992; Cooper *et al.*, 1987), however, age matched controls should be used to detect carriers. The determination of plasma β-mannosidase activity within the same herd can provide useful information for β-mannosidase carrier detection (Cavanagh *et al.*, 1992; Taylor *et al.*, 1993). Further studies are needed to discover a more efficient method for carrier detection. Preliminary studies suggested that prenatal diagnosis of β-mannosidosis by oligosaccharide detection in allantoic or amniotic fluid was possible (Jones *et al.*, 1984; Dahl *et al.*, 1986). The high expression of β-mannosidase activity in amniotic cells and in chorionic villi make it feasible to do prenatal diagnosis by amniocentesis or chorionic villi sampling (Jones *et al.*, 1984; Poenaru *et al.*, 1992).

β-Mannosidosis is inherited as a simple autosomal recessive genetic defect. The disease has been found in both sexes. Obligate carriers generally show an intermediate level of β-mannosidase activity. To study the possibility of a defect in a common factor in a patient with a combined deficiency of β-mannosidase and heparin sulphate sulfamidase (Wenger *et al.*, 1988), complementation studies were carried out by Hu *et al.* (1991) using cultured skin fibroblasts from the patient with combined deficiency, other

human β-mannosidosis cases from five different families, and one caprine β-mannosidosis case. No complementation was observed between the combined deficiency cells and other human β-mannosidase cells, between the combined deficiency case and caprine one, and between these patient cell lines. The results indicated that an allelic mutation in the same gene is responsible for these human as well as caprine β-mannosidosis cases. The molecular basis for the combined deficiency case is still unclear.

1.4 β-MANNOSIDASE

The degradation of N-linked glycoproteins involves three mechanisms: (1) the sequential removal of the sugar monomers from the non-reducing end of carbohydrate side chains by exo-glycosidases; (2) cleaving the chitobiose link between the two N-acetylglucosamine residues located in the "core" portion by endo-β-N-acetylglucosaminidase; and (3) the hydrolysis of the bond between N-acetylglucosamine and asparagine by aspartylglycosaminidase (Watts and Gibbs, 1986). Species differences in this pathway have been reported by Aronson *et al.* (1989).

β-Mannosidase (EC 3.2.1.25) is one of the exo-glycosidases involved in the catabolism of glycoproteins. It cleaves the β-linked mannosyl residue at the final step of the degradation of N-linked oligosaccharides. β-

Mannosidase is widely distributed in fungi, insects, plants, and animals. Its activity has been detected in plasma and in various tissues such as brain, liver, kidney, thyroid, spleen, urine, skin fibroblast, granulocytes, and lymphocytes (Cavanagh *et al.*, 1982; 1992; Bernard *et al.*, 1986; Dunstan *et al.*, 1983; Pearce *et al.*, 1987; Cooper *et al.*, 1987; Colin *et al.*, 1987; Jones *et al.*, 1984). β-Mannosidase activity is highest in thyroid, with decreasing levels in kidney, liver, muscle, and brain (Pearce *et al.*, 1987; Lovell *et al.*, 1994). The enzyme activity can be influenced by age, sex, reproductive status, and other factors (Cooper *et al.*, 1987; Dunstan *et al.*, 1983). It was found that the β-mannosidase activity in goats as well as in humans was significantly decreased during the period of sexual maturation and leveled off in adulthood.

Lysosomal enzymes are glycoproteins with acidic pH optima. Neutral counterparts of α-mannosidase, β-mannosidase, β-glucosidase, β-galactosidase, and sphingomyelinase have been discovered. Except sphingomyelinase, all neutral forms are present exclusively in liver (Chatterjee *et al.*, 1989). The discovery of considerable residual activity of β-mannosidase in the liver of affected kids (Jones and Dawson, 1981; Cavanagh *et al.*, 1982) led to Dawson's investigation of the neutral form of β-mannosidase in goat liver tissue (Dawson, 1982). By concanavalin A (Con A)-sepharose 4B chromatography, the

unbound and bound forms of β-mannosidase were separated. The latter represents typical lysosomal enzyme with acidic pH optimum (5.0-5.5), while the former one showed a broad and high pH optimum (6.0-8.0). The study of liver β-mannosidase activity of affected animals demonstrated that it consisted exclusively of the unbound forms. Further studies of the lysosomal and non-lysosomal forms of β-mannosidase in goats by Cavanagh *et al.* (1985) demonstrated that these two forms have different molecular weights, isoelectric points, and substrate specificities and reacted differently towards inhibitors. Pearce *et al.* (1987) reported that the activity of the non-lysosomal form increased progressively in liver during the second half of gestation. These reports clearly illustrated that the lysosomal and non-lysosomal forms are genetically, structurally, and functionally distinct from each other. The substantial residual activity found in the affected animal liver was due to the activity of the non-lysosomal form which does not hydrolyse trisaccharide. The biological function of liver specific β-mannosidase is still unknown. The neutral form of β-mannosidase has not been found in humans (Noeske *et al.*, 1983; Iwasaki *et al.*, 1989).

Apart from the non-lysosomal form, multiple isoforms of lysosomal β-mannosidase have been demonstrated in goats and humans by different groups (Pearce *et al.*, 1987; Frei *et al.*, 1988; Percheron *et al.*, 1992). The difference among

these isoforms largely reflects the degree of sialylation.

β-Mannosidase has been purified to various degrees from several mammalian sources including guinea pig liver, human placenta, and bovine and goat kidney (Sukeno *et al.*, 1972; Noeske *et al.*, 1983; Kyosaka *et al.*, 1985; Frei *et al.*, 1988; Iwasaki *et al.*, 1989; Sopher *et al.*, 1992; 1993). The first mammalian β-mannosidase purified to homogeneity was from guinea pig liver. The molecular size of the purified β-mannosidase protein was found to be 97-110 kDa (Kyosaka *et al.*, 1985; McCabe *et al.*, 1990). The protein was thought to be monomeric since the molecular weight estimated by gel filtration and SDS-PAGE were very close. However, additional studies suggested that it consisted of at least 3 subunits (McCabe and Dawson, 1991). McCabe and Dawson prepared a specific polyclonal antibody using Kyosaka's purified pig liver β-mannosidase protein. Western analysis of a fresh liver homogenate using this polyclonal antibody revealed a major band at 150 kDa. However, when frozen proteins were used, the immunoreaction pattern changed: the 150 kDa band disappeared, but two bands at 120 kDa and 20 kDa were observed. These were replaced finally by three bands at 97, 37, and 20 kDa after more cycles of freezing and thawing of proteins. Fresh proteins from a crude lipid extraction showed the 97, 37, and 20 kDa immunoreactive bands, too, which led the authors believe that the three subunits of guinea pig β-mannosidase were stabilized by a

lipid environment. The 120 and 150 kDa proteins were not observed in guinea pig kidney. The reason for this heterogeneity was not addressed. The functional and structural relationship between the two smaller subunits and the 97 kDa subunit was not clear.

A five-step purification of β-mannosidase from human placenta yielded a 10,000-fold purification (Iwasaki *et al.*, 1989). However, SDS-PAGE still revealed multiple bands with molecular masses from 98 to 57 kDa. The molecular weight calculated by gel filtration was 110 kDa, which suggested that the human β-mannosidase may consist of multiple subunits.

More recently, specific monoclonal and polyclonal antibodies against caprine and bovine β-mannosidases were produced in our laboratory. β-Mannosidase can now be highly purified (12,000 to 16,000 fold) from bovine and goat kidney tissues by a five-step purification procedure including an immunoaffinity chromatography step using a monoclonal antibody (Sopher *et al.*, 1992; 1993). Upon separation by SDS-PAGE, the purified goat protein revealed three bands (80, 90, and 100 kDa) by silver staining and Western analysis. The two major bands (90 and 100 kDa) were clearly associated with β-mannosidase activity. They both reacted with a monoclonal antibody specific against β-mannosidase and exhibited similar peptide patterns under limited proteolysis. Neither the 90 nor the 100 kDa peptide were

detectable in caprine β-mannosidosis kidney as indicated by silver staining and Western analysis of affinity purified protein. The 80 kDa peptide which copurified with the β-mannosidase peptides represents a small amount of the total amount of this protein in the Con A bound fraction. This peptide was clearly not related with the catalytic activity of the β-mannosidase and was also found to be present in β-mannosidosis goats by Western analysis. The purification of bovine β-mannosidase produced a very similar result. Three bands with the size of 84 kDa, 100 kDa, and 110 kDa were observed by silver staining and Western analysis. Glycosylation studies indicated that the size difference of the peptides between the two species was due to the different amounts of carbohydrate side chains. Removal of N-linked carbohydrate decreased the sizes of the major peptides to 86 kDa and 91 kDa in both species. Besides the size difference, another difference observed between cattle and goat was that the affected goats lack any 90 and 100 kDa but contain the 80 kDa protein as documented by silver staining and Western analysis. In contrast, the 84 kDa peptide found in normal bovine tissues and cattle with β-mannosidosis reacted with anti-β-mannosidase antiserum and did not react with anti-80 kDa antiserum. It was proposed that the 84 kDa peptide found in variable amounts in bovine kidney probably resulted from degradation of β-mannosidase peptides. The reason for this difference and the functional

relationship between the 80 kDa peptide and β-mannosidase are still unclear. Like Kyosaka's guinea pig (Kyosaka *et al*; 1985), the bovine and goat β-mannosidases appear to be monomeric.

The structural gene encoding β-mannosidase had not been previously cloned from any species. In 1987, Lundin (1987) localized a gene controlling mouse β-mannosidase activity in the distal part of mouse chromosome 3. However, whether it is a β-mannosidase structural or regulatory gene is unclear. In the same year, Fisher *et al*. (1987) tentatively localized the gene responsible for human β-mannosidase to human chromosome 4 whose long arm is syntenically related to the distal part of mouse chromosome 3. A putative cDNA was cloned from guinea pig using a specific anti-β-mannosidase polyclonal antibody (McCabe and Dawson, 1991). However, chromosome mapping and linkage studies suggested that this clone did not represent the β-mannosidase gene (Sopher, 1992a).

This investigation was designed to isolate and characterize the cDNA encoding β-mannosidase. It is an important step towards the achievement of the goal of development of an animal model system for gene therapy of a lysosomal storage disease characterized by neurodegeneration and early onset.

The isolation of β-mannosidase cDNA was attempted by several approaches. First, a human cDNA library was

screened with putative goat cDNA clones in order to clone a human β-mannosidase cDNA and to verify the goat cDNA clones. Secondly, goat and bovine polyclonal antibodies were used to screen goat and bovine cDNA libraries after the generation of polyclonal antibodies and the failure of the isolation human β-mannosidase cDNA with goat clones. Finally, purified bovine β-mannosidase was microsequenced and peptide sequences were obtained. Bovine cDNA libraries were then screened with degenerate oligonucleotides. By this method a cDNA encoding bovine β-mannosidase was finally isolated and characterized.

CHAPTER TWO

ISOLATION AND CHARACTERIZATION OF HUMAN cDNA HOMOLOGOUS TO GOAT CLONES

CHAPTER TWO

ISOLATION AND CHARACTERIZATION OF HUMAN cDNA HOMOLOGOUS TO GOAT CLONES

2.1 INTRODUCTION

Since the discovery of β-mannosidosis in Nubian goats (Jones and Laine, 1981; Jones and Dawson, 1981), extensive studies regarding β-mannosidosis and β-mannosidase have been carried out in this laboratory (Matsuura *et al.*, 1981; Matsuura and Jones, 1985; Jones *et al.*, 1982; 1983; 1984; 1986; 1992; Jones, 1989; Jones and Abbitt, 1993; Lovell and Jones, 1983; 1985; Lovell and Boyer, 1987; Lovell, 1990; Lovell *et al.*, 1991; 1994; Dahl *et al.*, 1986; Kumar *et al.*, 1986; Fisher *et al.*, 1987; Cavanagh *et al.*, 1982; 1985; 1992; Dunstan *et al.*, 1983; Frei *et al.*, 1988; Sopher *et al.*, 1992; 1993). The partial purification and characterization of goat β-mannosidase protein in 1990 led to the preparation of monoclonal antibodies (MAbs) (Sopher *et al.*, 1992). Three MAbs (43F10S, 44A6, and 44D9) which were capable of reacting with both native and denatured goat β-mannosidase proteins were produced (Sopher *et al.*, 1992). The production of MAbs not only resulted in the generation

of an immunoaffinity column and high purity β-mannosidase protein (Sopher *et al.*, 1992; 1993), but also initiated the molecular cloning of the β-mannosidase gene. Two goat thyroid cDNA libraries constructed in the λgt11 vector in this lab were screened by a combination of two of the three MAbs (44D9 and 43F10S) in an attempt to find a caprine β-mannosidase clone. A total of 4 positive clones were identified. Preliminary studies suggested that two of the clones (p5m8 and p5m11) might be good candidates for β-mannosidase (Sopher, 1992a). However, these two clones were not long enough to encode the whole β-mannosidase protein if they were authentic clones for β-mannosidase.

Cloning and characterization of human β-mannosidase cDNA were the original goal of this project. A high quality human cDNA library was chosen for isolation of a full-length clone. Two goat cDNA clones were used to probe a human placenta cDNA library in order to isolate human β-mannosidase gene. A full-length clone could be transferred to a expression system for verification of β-mannosidase activity. Thus, it was postulated that screening of a high quality human cDNA library might enable us to isolate human β-mannosidase cDNA and verify the goat clones.

2.2 MATERIALS AND METHODS

2.2.1 Materials

A human placenta cDNA library cloned in λZAPII vector and XL1-Blue *E. coli* cells were obtained from Stratagene (La Jolla, CA). pSVL vector and DEAE-dextran were purchased from Pharmacia LKB Biotechnology Inc. (Piscataway, NJ). Sequenase Version 2.0 T7 DNA polymerase, M13 reverse and -21 primers, exonuclease III, and mung bean nuclease were purchased from United States Biochemical Corp. (Cleveland, OH). T4 ligase, random primed DNA labeling kit, and restriction enzymes were from Boehringer Mannheim Biochemicals (Indianapolis, IN). Nitrocellulose paper BA85 (0.45 μm) was from Schleicher & Schuell (Keene, NH). [α-^{32}P]dCTP (3000 Ci/mmol; 111 TBq/mmol) was from Amersham Life Science (Arlington Heights, IL). ^{35}Sequetide (1500 Ci/mmol; 55.5 TBq/mmol) was from NEN/Du Pont (Wilmington, DE). DH5α competent cells and Dulbecco's Modified Eagle Media (DMEM) were purchased from Gibco BRL (Gaithersburg, MD). Chloroquine, polyethylene glycol, 4-methylumbelliferyl β-D-mannopyranoside, and 4-methylumbelliferyl α-D-mannopyranoside were purchased from Sigma (St. Louis, MO).

2.2.2 Library screening

Approximately 1 × 10^6 phage plaque-forming units of the human placenta cDNA library were plated at a density of

5×10^4 phage plaque-forming units per 150 mm plate and screened with two goat cDNA probes. Duplicate nitrocellulose filters were prehybridized (2 hours) and hybridized (16-20 hours) in a solution containing 0.4 M NaCl/0.01 M Pipes (pH 6.5)/50% formamide/0.5% SDS/100 µg/ml denatured herring sperm DNA at 42°C. The cDNA probes were labeled using the random primed DNA labeling method according to the manufacturer's instructions. Approximately 1×10^6 counts per minute (cpm) of radioactive DNA probe per milliliter were used during hybridization. The hybridized filters were washed in a step-wise fashion with the final wash in 0.2 × SSC/0.1% SDS at 65°C. Positive recombinant clones were isolated by plaque purification and phage DNA was prepared by the standard procedures (Sambrook *et al.*, 1989).

2.2.3 Subcloning

Positive clones were subcloned into pBluescript plasmid vector by *in vivo* excision following the manufacturer's instructions. Plasmid DNA was prepared by a boiling method according to the Stratagene instructions. Clone #1p5m11 was subcloned into pSVL vector essentially as described in Current Protocols in Molecular Biology (Ausubel *et al.*, 1989). Briefly, 4 µg of plasmid DNA from clone #1p5m11 were digested by *Acc*I enzyme, then treated with Klenow enzyme to generate blunt ends. Phosphorylated *Xho*I

linker was ligated to the plasmid DNA followed by digestion with *Xho*I enzyme. pSVL vector DNA was digested by *Xho*I and dephosphorylated by calf intestine alkaline phosphatase. Both the insert of clone #1p5m11 and vector DNA were then gel purified. Finally, the plasmid DNA of clone #1p5m11 was ligated with pSVL vector. One fourth of the ligation was used to transform subcloning efficiency DH5α competent cells following the manufacturer's procedures. Subclones were identified by *in situ* hybridization with #1p5m11 insert as a probe by the standard procedures (Sambrook *et al.*, 1989).

2.2.4 Nucleotide Sequencing

Sequencing was performed using M13 reverse and -21 primers on the intact plasmid. To obtain internal sequence, exoIII/mung bean nuclease deletion was chosen to prepare deletion clones. Large amounts of #1p5m11 DNA were first purified by precipitating with polyethylene glycol essentially according to the method described in Molecular Cloning (Sambrook *et al.*, 1989). Approximately 20 μg of plasmid DNA was subjected to double restriction enzyme digestions. For preparation of the 5′ deletion, restriction enzymes *Cla*I and *Kpn*I were selected. For the 3′ deletion, *BamH*I and *Sac*I were used. The DNA was first digested by one restriction enzyme, then extracted once with phenol/chloroform (1:1) and once with chloroform. DNA was reprecipitated by adding 1/10 volume of 3 M NaOAC, pH 5.6

and two volumes of ethanol. The resuspension was then subjected to the second restriction enzyme digestion. After the completion of double enzyme digestions, exoIII/mung bean nuclease deletions were carried out at 37°C following the Stratagene instructions. DNA that had been treated by exoIII for different times and thus fragmented into different lengths was ligated and transformed into XLI-Blue or DH5α competent cells according to the manufacturers' instructions. DNA from overlapping clones was prepared and sequenced by the dideoxy chain-termination method using T7 sequenase DNA polymerase, ^{35}S-labelled dATP, and M13 universal primers. DNA sequence data and homology searches were analyzed using the GCG program, version 7 by the Genetics Computer Group DNA Sequencing Analysis Software, Madison, Wisconsin.

2.2.5 In Vitro Expression

COS-7 cells (gift from Dr. D. Dewitt) were grown in DMEM containing 10% fetal bovine serum (FBS) in 60 mm culture dishes. After the cells were about 75% confluent, 10 μg of plasmid DNA from recombinant clones pSVL-15, pSVL-32, and pSVL were transfected into the cells by a DEAE-dextran method (Oshima *et al.*, 1988). After 10 hours of incubation with DEAE-dextran, the cells were washed with DMEM/10% fetal bovine serum and treated for 3 hours with 100 μM chloroquine in 2 ml of DMEM/10% FBS. Cells were

harvested after 48 to 72 hours of incubation. Cell lysates and supernatant were assayed for β-mannosidase and α-mannosidase activities with 4-methylumbilliferyl substrates as described previously (Jones *et al.*, 1984).

2.3 RESULTS AND DISCUSSION

2.3.1 Cloning the human homologues of goat clones

By screening 1 × 10^6 recombinants from the human placenta λZAPII cDNA library, twelve positive clones were identified by a cDNA probe from the goat cDNA clone p5m8, and five clones by a cDNA probe from the goat cDNA clone p5m11. Six clones from the p5m8 group and three from the p5m11 group were excised as pBluescript plasmids. Restriction enzyme digestion analysis suggested that the two groups of clones had different types of restriction maps (Fig. 2.1). This result was consistent with previous Southern hybridization data (Friderici, unpublished data) that indicated there was little homology between the two goat clones p5m8 and p5m11. Except for clone 15p5m8 which showed a weak signal in Southern hybridization analysis (data not shown), identical or overlapping restriction maps within each group were observed (Fig. 2.1).

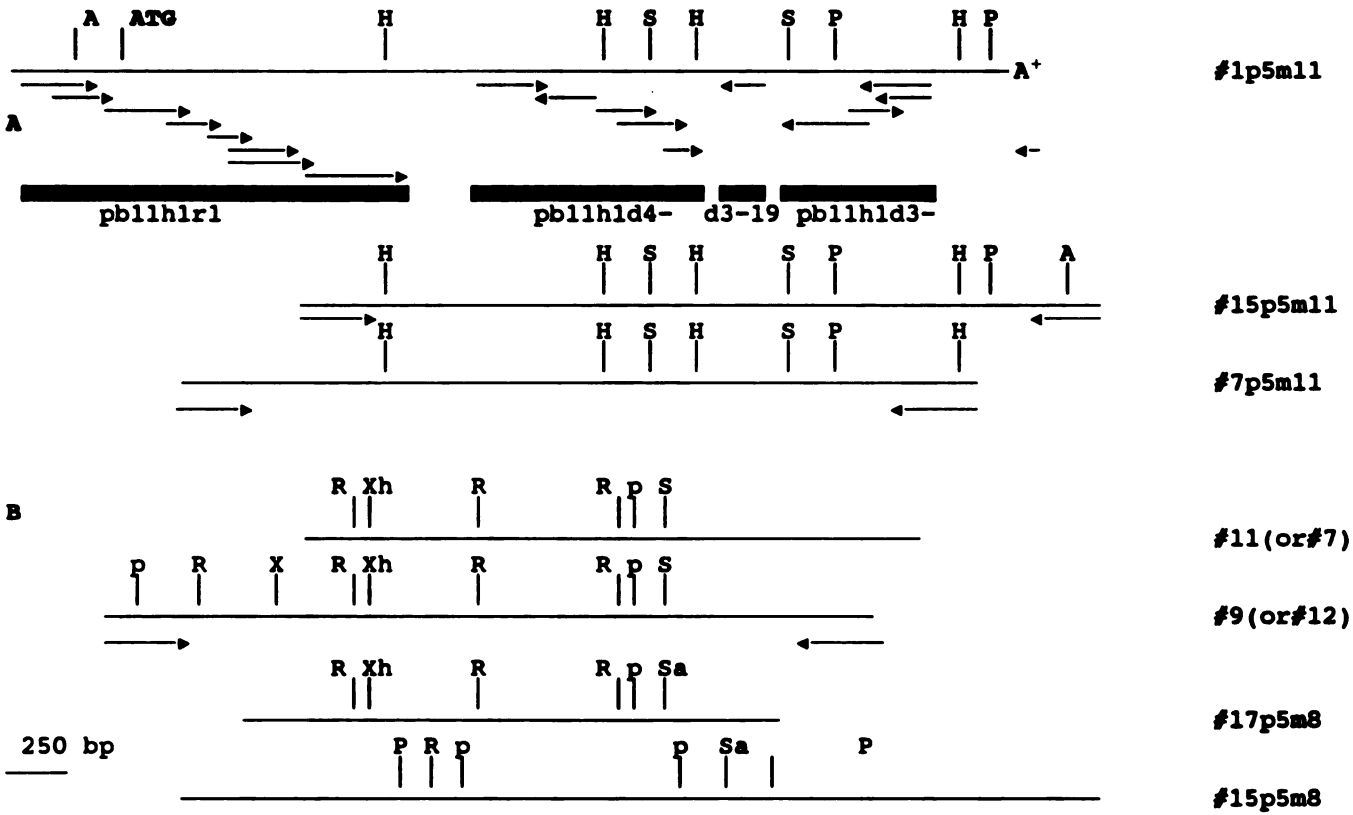


Fig. 2.1 The restriction map and sequencing strategy of human cDNA clones. Panel A: three clones isolated by a goat cDNA p5m11. Panel B: Clones isolated by a goat cDNA p5m8. Clone #15p5m8 showed a weak signal in a Southern hybridization with p5m8 probe. ATG, the putative translational initiation codon. Arrows represent sequence directions. Restriction enzymes: H, *Hind*III; S, *Stu*I; Sa, *Sac*I; P, *Pvu*II; A, *Acc*I; R, *Rsa*I; Xh, XhoI; X, *Xba*I; p, *Pst*I. Solid bars represent consensus sequences in Fig. 2.2.

2.3.2 Sequence of human p5m11 clones

Clone #1 from the p5m11 group, which contains a 3.6 kb insert was subjected to exoIII/mung bean nuclease deletions and sequenced from both directions by the dideoxy chain-termination method. A partial DNA sequence (3070 bp) (Fig. 2.2) revealed a putative initiation codon, a poly (A) tail, and three stretches of open reading frames that were interrupted by three short gaps of unsequenced DNA. The sequence flanking the initiation codon (AGCATGG) was in good agreement with the consensus sequence (A/GCCATGG) for a eukaryotic initiation codon by Kozak (1986). No typical signal peptide sequence was found. Sequencing searches did not reveal any significant sequence similarity with proteins in the database. However, sequence similarity was found between the goat clone p5m11 and the human clone #1p5m11. Interestingly, the deduced amino acid sequences from the goat and human clone had repetitive sequences consisting of 5 residues with a proline residue at the beginning (Fig 2.2 underlines). The molecular mass of human β-mannosidase peptide was 98 kDa (Iwasaki *et al.*, 1989). Clone #1p5m11 was long enough to encode a 98 kDa protein. Sequencing both ends of clones #7p5m11 and #15p5m11 with M13 reverse and -21 primers confirmed their homologies with clone #1p5m11. Partial sequencing (both ends) of clone #9 from the p5m8 group did not reveal any open reading frames or possible translational initiation codon.

Fig. 2.2 Partial nucleotide sequence of human #1p5m11 clone. Three fragments were displayed in a 5′ → 3′ direction. Nucleotides and amino acids are numbered noncontinously between each fragment. Gaps are marked with ///////. The repetitive proline sequences in fragment pb11h1d4- are marked with underlines (the human clone) and double underlines (the goat clone). The putative initiation codon is marked by *.**

```
          .         .         .         .         .         .
  1 ggctgcaggaattcctgaacttgtgcaaataactttattaccataaacctatgaatactc
          .         .         .         .         .         .
 61 atgaatagtttcccaattctggggcactcagatagagagcaaaagcaaatgtttcaattt
          .         .         .         .         .         .
121 ttgtttacaaaagtatactttaccaattgctgaagaaaaaaagttcataaatctggagaa
          .         .         .         .         .         .
181 taaaacattccaagaatcagcacattttccaataaaaaattatgaaaacattatcctttt
          .         .         .         .         .         .
241 gattatttagtccaataacattgagtttttttcttctaattcatctcttgttttatcagg
          .         .         .         .         .         .
301 tgtgtgtggtttcagcgcagcatggctgtggtcatccgtttgcaaggtctcccaattgtg
                     ***
  1                  M  A  V  V  I  R  L  Q  G  L  P  I  V
          .         .         .         .         .         .
361 gcggggaccatggacattcggcacttcttctctggattgaccattcctgatggggcgtg
 14 A  G  T  M  D  I  R  H  F  F  S  G  L  T  I  P  D  G  G  V
          .         .         .         .         .         .
421 catattgtaggggtgaactgggtgaggctttcatcgttttgccactgatgaagatgca
 34 H  I  V  G  G  E  L  G  E  A  F  I  V  F  A  T  D  E  D  A
          .         .         .         .         .         .
481 aggcttggtatgatgcgcacaggtggtacaattaaagggtcaaaagtaacactattgttg
 54 R  L  G  M  M  R  T  G  G  T  I  K  G  S  K  V  T  L  L  L
          .         .         .         .         .         .
541 agtagtaagacggaaatgcagaatatgattgaactgagtcgtaggcgttttgaaactgcc
 74 S  S  K  T  E  M  Q  N  M  I  E  L  S  R  R  R  F  E  T  A
          .         .         .         .         .         .
601 aacttagatataccaccagcaaatgccagtagatcaggaccaccacctagctcaggaatg
 94 N  L  D  I  P  P  A  N  A  S  R  S  G  P  P  P  S  S  G  M
          .         .         .         .         .         .
661 agtagcagggtaaacttncccacaacagtatccaactttaataatccatcacccagtgta
114 S  S  R  V  N  X  P  T  T  V  S  N  F  N  N  P  S  P  S  V
          .         .         .         .         .         .
721 gttactgccaccacttctgttcatgaaagcaacaaaaacatacagacattttccacagcc
134 V  T  A  T  T  S  V  H  E  S  N  K  N  I  Q  T  F  S  T  A
          .         .         .         .         .         .
781 agcgtaggaacagctcctccaaatatgggggcttcctttgggagcccaacgtttagctca
154 S  V  G  T  A  P  P  N  M  G  A  S  F  G  S  P  T  F  S  S
```

Fig. 2.2

```
                .         .         .         .         .         .
 841  actgttccaagcacagcctctccaatgaacacagtcccgccgccaccaattcctccaatt
 174  T  V  P  S  T  A  S  P  M  N  T  V  P  P  P  P  I  P  P  I
                .         .         .         .         .         .
 901  ccagcgatgccatctctgccaccaatgccatccattcccccaattccagttcctcctcca
 194  P  A  M  P  S  L  P  P  M  P  S  I  P  P  I  P  V  P  P  P
                .         .         .         .         .         .
 961  gtacctacattgcctcctgtncctcctgtnccccgattnccccagttccttctgtgcca
 214  V  P  T  L  P  P  V  P  P  V  P  P  I  X  P  V  P  S  V  P
                .         .         .         .         .         .
1021  cccatgaccccactgccacccatgtcgggcatgccgcccttgaatccgccacctgtggca
 234  P  M  T  P  L  P  P  M  S  G  M  P  P  L  N  P  P  P  V  A
                .         .         .         .         .         .
1081  cctctacctgctggaatgaatggctctggagcacctatgaatttgaacaataatctgaat
 254  P  L  P  A  G  M  N  G  S  G  A  P  M  N  L  N  N  N  L  N
                .         .         .         .         .         .
1141  cctatgtttcttggtccgttgaatcctgttaaccctatccagatgaactctcagagcagt
 274  P  M  F  L  G  P  L  N  P  V  N  P  I  Q  M  N  S  Q  S  S
                .         .         .         .         .         .
1201  gtgaagccactccccatcaaccctgatgatctgtatgtcagtgtgcatggaatgcccttt
 294  V  K  P  L  P  I  N  P  D  D  L  Y  V  S  V  H  G  M  P  F
                .         .         .         .         .         .
1261  tctgcaacggaaaatgatgtcagagatttttttcatgggctccgtgttgatgcagtgcat
 314  S  A  T  E  N  D  V  R  D  F  F  H  G  L  R  V  D  A  V  H
                .         .         .         .         .         .
1321  ttgttgaaagatcatgtaggtcgaaataatgggaatggattggttaagtttctctcccct
 334  L  L  K  D  H  V  G  R  N  N  G  N  G  L  V  K  F  L  S  P
                .         .         .         .         .         .
1381  caagatacatttgaagctttgaaacgaaacagaatgctgatgattcaacgctatgtggaa
 354  Q  D  T  F  E  A  L  K  R  N  R  M  L  M  I  Q  R  Y  V  E
                .         .         .         .         .         .
1441  gttagccctgccacagaaagacagtgggtagctgctggaggccatatcacttttaagcaa
 374  V  S  P  A  T  E  R  Q  W  V  A  A  G  G  H  I  T  F  K  Q
                .         .         .         .         .         .
1501  aatatgggaccttctggacaaactcatcccctcctcagacacttccaggtcaaatcgcca
 394  N  M  G  P  S  G  Q  T  H  P  L  L  R  H  F  Q  V  K  S  P
                .         .
1561  gtggcagaaagatcaggtcaaga          (pb11h1r1)
 414  V  A  E  R  S  G  Q
```

Fig. 2.2

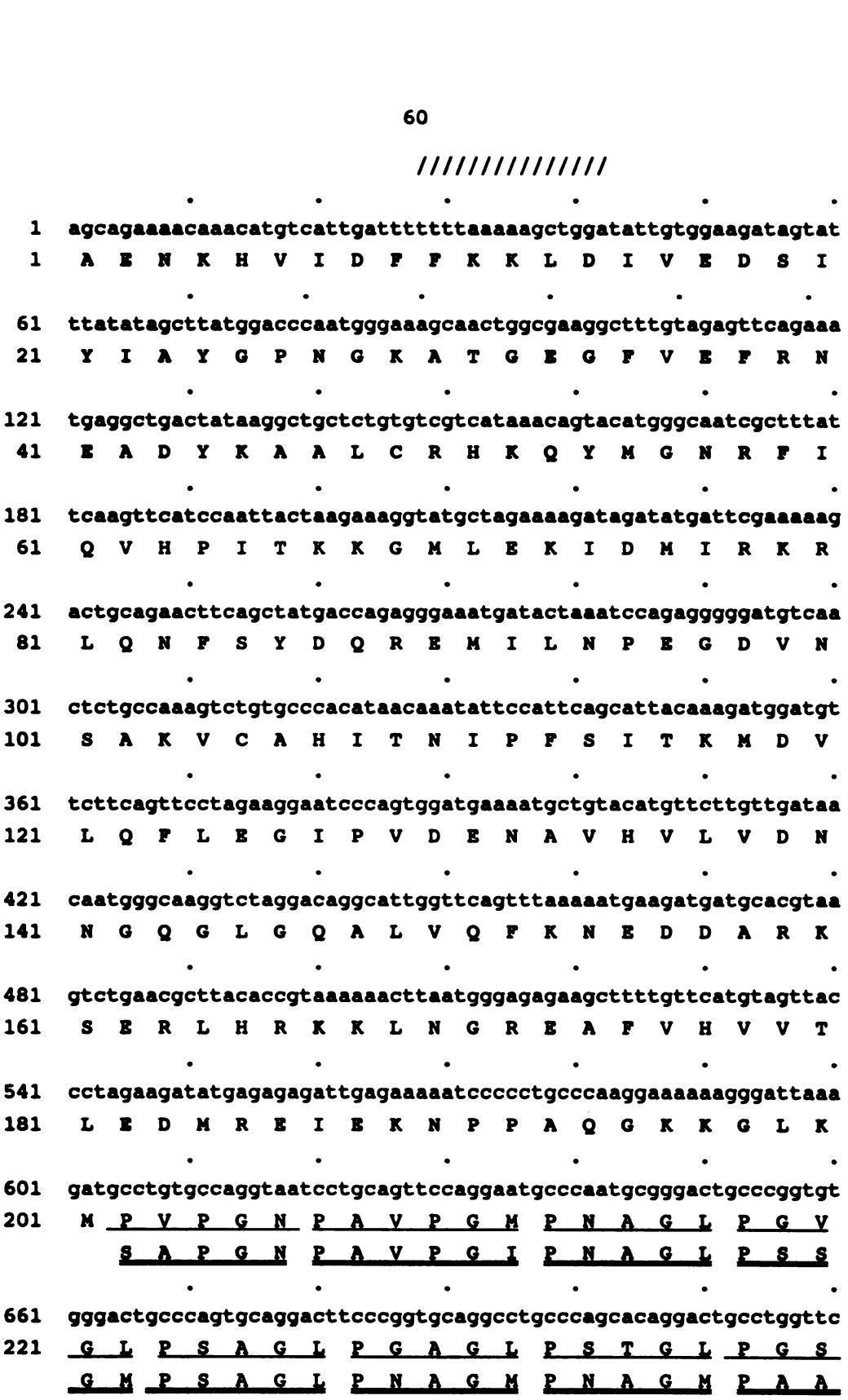

60
1 agcagaaaacaaacatgtcattgatttttttaaaaagctggatattgtggaagatagtat
1 A E N K H V I D F F K K L D I V E D S I
61 ttatatagcttatggacccaatgggaaagcaactggcgaaggctttgtagagttcagaaa
21 Y I A Y G P N G K A T G E G F V E F R N
121 tgaggctgactataaggctgctctgtgtcgtcataaacagtacatgggcaatcgctttat
41 E A D Y K A A L C R H K Q Y M G N R F I
181 tcaagttcatccaattactaagaaaggtatgctagaaaagatagatatgattcgaaaaag
61 Q V H P I T K K G M L E K I D M I R K R
241 actgcagaacttcagctatgaccagagggaaatgatactaaatccagaggggatgtcaa
81 L Q N F S Y D Q R E M I L N P E G D V N
301 ctctgccaaagtctgtgcccacataacaaatattccattcagcattacaaagatggatgt
101 S A K V C A H I T N I P F S I T K M D V
361 tcttcagttcctagaaggaatcccagtggatgaaaatgctgtacatgttcttgttgataa
121 L Q F L E G I P V D E N A V H V L V D N
421 caatgggcaaggtctaggacaggcattggttcagtttaaaaatgaagatgatgcacgtaa
141 N G Q G L G Q A L V Q F K N E D D A R K
481 gtctgaacgcttacaccgtaaaaaacttaatgggagagaagcttttgttcatgtagttac
161 S E R L H R K K L N G R E A F V H V V T
541 cctagaagatatgagagagattgagaaaaatccccctgcccaaggaaaaaagggattaaa
181 L E D M R E I E K N P P A Q G K K G L K
601 gatgcctgtgccaggtaatcctgcagttccaggaatgcccaatgcgggactgcccggtgt
201 M P V P G N P A V P G M P N A G L P G V
S A P G N P A V P G I P N A G L P S S
661 gggactgcccagtgcaggacttcccggtgcaggcctgcccagcacaggactgcctggttc
221 G L P S A G L P G A G L P S T G L P G S
G M P S A G L P N A G M P N A G M P A A

Fig. 2.2

```
                .         .         .         .         .         .
721  agcaataaccagtgcaggactgcctggtgcgggaatgcccagtgcaggaatacctngtgc
241  A  I  T  S  A  G  L  P  G  A  G  M  P  S  A  G  I  P  X  A
     G  M  P  N  A  G  M  P  A  A  G  M  P  N  A  G  I  P  S  T
                .
781  aggaggtgaagagcat
261  G  G  E  E  H                (pb11h1d4-)
     G  M

                          ///////////////

                .         .         .         .         .         .
  1  gcggggcctttggtgatgctaggcctggtatgccttcagttggaaacagtggtttgcctg
  1    G  A  F  G  D  A  R  P  G  M  P  S  V  G  N  S  G  L  P  G
                .         .         .         .         .         .
 61  gtctagnactggatgttccgggttttggaggtggaccaaacaatttaagtgggccatcgg
 21    L  X  L  D  V  P  G  F  G  G  G  P  N  N  L  S  G  P  S  G
                .         .         .         .         .         .
121  gatttggaggggccctcagaattttggaaatggccctggtagcttaggcggtcccccgg
 41    F  G  G  G  P  Q  N  F  G  N  G  P  G  S  L  G  G  P  P  G
                .
181  ggtttggaagtcccggc  197          (d3-19)
 61    F  G  S  P  G    65
                          ///////////////

                .         .         .         .         .         .
  1  taaagggaacaaaagctggagccatggtggcctttgagtctcgggatgaagccacagctg
                .         .         .         .         .         .
 61  ctgtcattgacttaaatgacaggcctataggttcaagaaaagtaaaacttgtattagggt
                .         .         .         .         .         .
121  agccattcacatcatttttatngggtagatcttcatattgctgtgattaatgcatccag
                .         .         .         .         .         .
181  attgttttcctagtatttccaggttagaacctgtggattgtttcaattgcatatagcttg
                .         .         .         .         .         .
241  gtttccataacatagagcattggttgactgtttacagaagactcactcaccaggatgggc
                .         .         .         .         .         .
301  attgctgtatgttacagtaaagctatctggagagaacacatgggtgattttggcatacca
                .         .         .         .         .         .
361  ttagagaaaccatttgtaaaactcaaatgaccacataaagcttatcaaggagtctagatt

421  ggtttttg  428          (pb11h1d3-)
```

Fig. 2.2

2.3.3 *Expression of human p5m11*

In order to determine whether clone #1p5m11 was β-mannosidase, the insert was excised by *EcoRI* digestion, gel purified, and subcloned into an eukaryotic expression vector pSVL. A recombinant clone (pSVL-15) which contains an insert with the correct size and orientation and a clone (pSVL-32) containing an insert with the right size but reverse orientation were identified by in situ hybridization and restriction enzyme analysis. Transfection of clones pSVL-15 and pSVL-32 into *COS*-7 cells led to no elevation of the β-mannosidase activity in the cells and supernatant compared to the transfection of pSVL only. Northern analysis indicated the level of RNA extracted from cells that were transfected by pSVL-15 after 72 hours was increased (data not shown). Positive results of *in vitro* expression would certainly confirm the authenticity of clones which were analyzed, however, negative results could not rule out the authenticity of clones. The successful expression of an enzyme activity *in vitro* not only relies on a full-length cDNA, but may also be influenced by others factors, such as a cofactor needed for expression of the activity or for stabilizing. While the verification of the human cDNA clones was underway, characterization of the goat clones p5m8 and p5m11 in parallel experiments (Sopher and Friderici, unpublished data) provided more evidence that these goat clones were not the β-mannosidase gene. The

evidence for this included: (1) when genomic DNA from several human-rodent cell hybrids was hybridized with clones p5m8 and p5m11, none of these clones were located on human chromosome 4 as was expected by previous work (Fisher et al. 1987; Lundin, 1987); (2) the fusion proteins produced by the two goat clones could not be recognized by any of the specific β-mannosidase polyclonal antisera that became available later on. Therefore, it was clear that the human clones identified by p5m8 or p5m11 did not correspond to the β-mannosidase gene.

2.4 SUMMARY

A human placenta λZAPII cDNA library was screened by two goat cDNA probes, p5m8 and p5m11, separately. Of 10^6 plaques screened, twelve positive clones were hybridized with p5m8 probe, and five with p5m11 probe. Six clones in the p5m8 group and three in the p5m11 group were excised as pBluescript plasmids and subjected to further analyses by restriction mapping and DNA sequencing. The possibility that these clones might be candidates for the β-mannosidase gene was ruled out by *in vitro* expression in *COS* cells and more importantly by the further analyses of the two goat cDNA clones (p5m8 and p5m11). The two goat cDNA clones were originally thought to be β-mannosidase but subsequently it was found that they more likely represented proteins sharing

the same or similar epitopes with the β-mannosidase protein rather than the β-mannosidase itself.

CHAPTER THREE

ISOLATION AND CHARACTERIZATION OF BOVINE β-MANNOSIDASE cDNA CLONES

CHAPTER THREE

ISOLATION AND CHARACTERIZATION OF BOVINE β-MANNOSIDASE cDNA CLONES

3.1 SCREENING WITH POLYCLONAL ANTIBODIES

3.1.1 Introduction

Immunoscreening is one of the common strategies used for the isolation of cDNA clones, especially when no amino acid sequences are available. Since screening goat libraries with MAbs did not yield β-mannosidase clones, attention was redirected toward identifying the β-mannosidase using polyclonal antiserum to purified enzyme.

Polyclonal antisera specifically against β-mannosidase were generated as the consequence of the purification of β-mannosidase protein (Sopher *et al.*, 1992; 1993). Antiserum to caprine β-mannosidase was prepared by immunizing rabbits with crushed gel slices containing the 90-kDa β-mannosidase peptide. A polyclonal antiserum, Z28, referred to as anti-80/90/100 in a previous paper (Sopher *et al.*, 1992) was produced. This antiserum could detect as little as 1 ng of purified β-mannosidase protein by dot blot analysis. It reacted mainly to the 90- and 100-kDa peptides, but also had

cross-immunoreaction to an 80 kDa peptide which was copurified with β-mannosidase as judged by Western analysis. Due to limited availability of goat kidneys, large scale enzyme purification was redirected to use bovine kidneys as an enzyme source (Sopher *et al.*, 1993). Studies of purified bovine β-mannosidase revealed three peptides. To generate very specific bovine polyclonal antibodies, the purified bovine protein was deglycosylated and the major band was gel purified and injected into rabbits. As a result, two anti-β-mannosidase antisera were generated from the deglycosylated bovine peptide by Sopher *et al.* (1993). The bovine specific antisera, Z59 and Z69 (referred to as the anti β-mannosidase in the published paper) were more specific than the goat antiserum Z28, with little cross-immunoreaction to the 80 kDa peptide in goat kidney. The bovine antisera could recognize β-mannosidase peptides from goat, and furthermore, goat antisera could react with bovine β-mannosidase peptides.

Polyclonal antibodies generally recognize multiple epitopes, therefore more peptides can be identified. The project was directed to bovine and caprine β-mannosidase at that time. In order to isolate the mammalian β-mannosidase gene, goat polyclonal antiserum Z28 and bovine antisera Z59 and Z69 were used to screen cDNA expression libraries in the following studies.

3.1.2 Materials and methods

3.1.2.1 Titer of polyclonal antisera

The titer and specificity of each polyclonal antiserum were determined by either dot blot test or Western analysis. Partially purified bovine protein consisting mainly of the 100 kDa with small amounts of the 110 kDa and 84 kDa β-mannosidase peptides was used as an antigen. In dot blot tests, serial dilutions of the antigen (50 ng to 100 pg) were spotted onto strips of nitrocellulose filter (Bio-Rad Laboratory, Melvill, NY). After air drying, the filter strips were blocked in a solution containing 10 mM Tris.Cl, pH 7.5/150 mM NaCl/0.05% Tween 20 (Bio-Rad) (TNT) and 5% dry milk for 1 hour at room temperature followed by incubation with serial dilutions (1:250 to 1:1000) of polyclonal antisera, individually, for 2 hours. The filters were then washed in TNT solution three times, each for 5 min, with the final wash in a solution containing 20 mM Tris-Cl, pH 7.5/150 mM NaCl (TBS) for 5 min. After washing, the filters were incubated with alkaline phosphatase (AP)-conjugated goat anti-rabbit antibody (Bio-Rad) for 1 hour followed by the washing-step described before. Finally, color was developed by reacting with 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium (BCIP/NBT) (Promega, Madison, WI) solution.

3.1.2.2 Purification of IgG fractions and affinity *purification of polyclonal antibodies*

The IgG fraction was purified by using a protein A column (1 ml) (Pierce, Rockford, IL) according to the procedures described by Harlow *et al.* (1988). Briefly, 1-2 ml of crude antiserum was adjusted to pH 8.0 by adding 1/10 volume of 1 M Tris.Cl, pH 8.0 and then applied to a protein A column which was preequilibrated with 0.1 M Tris, pH 8.0. After sequential washing with 10 volumes of 100 mM and 10 mM Tris.Cl, pH 8.0, the IgG fraction was eluted with 0.1 mM glycine pH 3.0. Approximately 1 ml elution fractions were collected into microcentrifuge tubes containing about 100 µl of 1 M Tris.Cl, pH 8.0. IgG-containing fractions were pooled together and bovine serum albumin (BSA) was added up to 1%.

To purify small amounts of epitope-selected anti-β-mannosidase antibodies, a total of 1 µg protein was fractionated on 7.5% SDS-PAGE (10 ng/well) and transferred onto PVDF membrane (Millipore, Bedford, MA) after electrophoresis. The filter was blocked and incubated with 1:500 diluted antiserum Z28 essentially as described above. The strips between 100 and 110 kDa were cut and the bound antibodies were eluted by incubating with 400 µl of 0.2 M glycine pH 2.6 for 10 min at room temperature. The elution was immediately adjusted to pH 8.0 by adding tris base solution. The activity of the elution was assayed by

Western analysis with 10 ng of TSK-butyl purified bovine protein.

3.1.2.3 Removal of anti-*E. coli* antibody

Before being used in library screening, diluted polyclonal antisera or IgG fractions were preincubated with nitrocellulose filters (Schleicher & Schuell, Keene, NH) coated with XL1-blue *E. coli* (Stratagene, La Jolla, CA) cell lysate until no substantial background was present. In some cases, the anti-*E. coli* antibodies were removed from polyclonal antisera by affinity chromatography (Sambrook *et al.*, 1989). A large amount of lysate of XL1-blue cells was prepared first according to the procedures described by (Sambrook *et al.*, 1989). The cell lysate was chilled to 0°C and bound to cyanogen bromide (CNBr)-activated sepharose 4B (Pharmacia LKB Biotechnology Inc., Piscataway, NJ) according to the manufacture's instructions. After equilibration with TBS solution containing 0.2% sodium azide, the slurry of CNBr-activated sepharose 4B coupled with *E. coli* lysate was mixed with IgG fractions or antiserum for more than 12 hours at room temperature. The slurry was then loaded into a column and eluted with TBS solution. Fractions containing antibody were pooled, diluted, and used to screen libraries.

3.1.2.4 Screening cDNA libraries

Approximately 5 × 10^4 phage plaque-forming units from goat kidney λZAPII (Clontech, Palo Alto, CA) or bovine liver λZAPII cDNA library (Stratagene) per 150-mm petri dish were plated. After about 3.5 hours incubation at 42°C, nitrocellulose filters (Schleicher & Schuell) presoaked in 10 mM β-D-isopropyl-thiogalactopyranoside (IPTG) (Boehringer Mannheim Biochemicals, Indianapolis, IN) were overlaid on the plates, and plaques were allowed to grow for another 3.5 hours. Duplicate filters were applied at the first round of screening. The immunoreaction procedures were essentially as described above in the dot blot test. Polyclonal antisera were diluted 300-500 fold in TNT solution containing 5% dry milk.

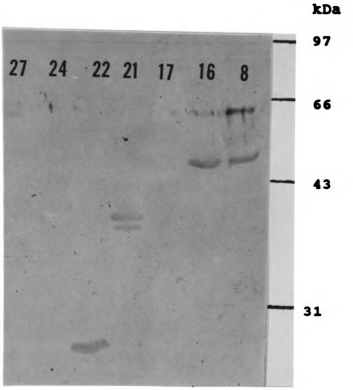
3.1.2.5 Identification of expressed fusion proteins

Clones identified by antibodies were plaque purified by several rounds of rescreening at lower density and excised into pBluescript plasmid (Stratagene) according to the manufacturer's instructions. To express fusion proteins, 10 ml of cell cultures inoculated by colonies from positive clones were grown at 37°C overnight. An aliquot of 50 μl of the overnight cell culture was added into 5 ml of LB medium containing ampicillin (100 μg/μl). After two hours of incubation at 37°C, fusion proteins were induced by adding IPTG solution up to a final concentration of 1 mM and

incubation of the cultures at 37°C for another two to four hours. An aliquot of 1 ml culture was removed after two and four hours of induction, respectively, and collected by centrifugation at 12,000 g for 1 min. The pellet was then resuspended in 100 μl of 1 × SDS loading buffer (2% SDS/10% glycerol/32.5 mM Tris.Cl (pH 6.8)/0.1% bromophenol blue/5% β-mercaptoethanol), boiled for 3 min, and centrifuged at 12,000 g for 1 min. Finally, 30 μl of suspension was loaded on 10% SDS-PAGE. After electrophoresis, the proteins were transferred to the Hybond-N membrane. The blot was stained with Ponceau S and subjected to Western analysis essentially according to Current Protocols in Molecular Biology (Ausubel *et al.*, 1987).

3.1.3 Results

Approximately 1 × 10^6 plaques from a goat kidney λZAPII cDNA library were first screened with the antiserum Z28. Nine positive clones were identified and plaque purified. Restriction enzyme analysis and Southern cross-hybridization suggested that clones G♯26Z28 and G♯27Z28 were virtually identical. Clones G♯8Z28 and G♯11Z28 also appeared to be identical and had sequence homologies with clone G♯16Z28. The remaining clones each appeared to represent a different gene. Of the nine clones, five produced fusion proteins (Fig. 3.1). Clones G♯8Z28, G♯11Z28 (not shown on Fig. 3.1), and G♯16Z28 all gave rise to a 50 kDa fusion protein, clone



3.1 Western Analysis of fusion proteins with antiserum Z28. Seven goat cDNA clones isolated by antiserum Z28 were analyzed. Clone #26 and #11 were identical clones of #27 and #8 respectively and were not included in this figure. Fusion proteins were induced by 1 mM IPTG for 2 hours. The protein lysates were fractionated by 10% SDS-PAGE. The gel was run at 25 mA for 2 hours, then 40 mA for 4 hours. The number on top of the figure represents each clone. Low range SDS-PAGE standard from Bio-Rad was used as a molecular weight marker.

G#22Z28 produced a 25 kDa fusion protein, while clone G#21Z28 generated a 38 kDa and a 39 kDa protein. These fusion proteins were all recognized by antiserum Z28 as judged by Western analysis (Fig. 3.1). Fusion proteins produced by clones G#8Z28, G#11Z28, G#16Z28, and G#22Z28 were later subjected to Western analysis to check cross-immunoreaction with two anti-β-mannosidase antisera (Z59 and Z69) and antiserum Z30. The latter antiserum reacted specifically with the copurified caprine 80 kDa peptide. Cross-immunoreaction to all three antisera in various degrees was observed in these four putative clones (table 3.1). The cross-immunoreaction to antiserum Z59 by the

Table 3.1 Western analysis of putative clones

Clones	fusion protein	polyclonal antisera Z28	Z59	Z69	Z30
G#8Z28	50 kDa	+	+	+	+?
G#16Z28	50 kDa	+	+	+	+?
G#17Z28	-	+	-		
G#21Z28	38, 38 kDa	+	-		
G#22Z28	25 kDa	+	-	+?	+
G#24Z28	-?	+	-		-
G#27Z28	-	+	-		

Based on the results of plaque immunoreactions and fusion protein analyses.

fusion proteins from clones G#8Z28, G#11Z28, and G#16Z28 was also demonstrated by immunoscreening of plaques. However, it appeared these fusion proteins could be recognized

neither by monoclonal antibodies 44D9 and 43F10 nor by affinity purified antibody against bovine 100 kDa and 110 kDa peptides. This suggested that these putative clones were not likely to be β-mannosidase.

After specific anti-β-mannosidase antisera Z59 and Z69 became available, the goat kidney cDNA library was rescreened using IgG fractions purified from polyclonal antibody Z69. In addition, Z59 antibody was used to screen a bovine liver λZAPII cDNA library. Upon screening of approximately 1×10^6 phage plaques from each library, only one positive clone (B#9Z59) was identified by antiserum Z59 after plaque purification, and 19 positive clones were selected by Z69-IgG. Clone B#9Z59 did not show any cross-immunoreaction with antiserum Z28 at all. Two out of the 19 clones identified by Z69-IgG had weak cross-immunoreaction to antiserum Z59. When these two clones (G#9Z69 and G#12Z69) were probed with antisera Z30 and Z28, it appeared that both clones had cross-immunoreaction to antiserum Z30. However, only clone G#12Z69 had weak immunoreaction with antiserum Z28. No obvious fusion proteins were produced by these clones as judged by a Ponceau S staining. Nevertheless, unexpected smeared bands were observed by Western analysis. Moreover, similar patterns of immunoreactions were produced when different antibody probes (Z30, Z69 IgG, and Z69 preimmunoserum) were used in the Western analysis. The smear patterns were not observed in

lanes containing proteins from clone G#8Z28 and plasmid pBluescript only.

3.1.4 Discussion

3.1.4.1 Screening with antiserum Z28

Before other specific polyclonal antisera were available, the goat polyclonal antiserum Z28 was chosen to immunoscreen goat kidney cDNA λZAPII library. Nine positive clones were identified. Six clones showed no sequence homologies as indicated by Southern hybridizations. Restriction enzyme analysis provided different restriction maps among these six clones. Of the remaining three clones, two were identical (G#8 and G#11) and one (G#16) was related to the identical clones. These results suggested that there were seven different groups in the nine positive clones. False positive clones were common during immunoscreening as any fusion proteins containing the same or similar epitopes may be recognized by polyclonal antibodies. Different preparations of antisera raised against the same protein may contain antibodies against different epitopes, therefore a clone, recognized by two separate preparations of antisera, is more likely to represent the given protein (Helfman and Hughes, 1987). By reacting with other polyclonal antisera acquired subsequently, a total of four clones out of the original nine putative clones (G#8Z28, G#11Z28, #16Z28, and G#22Z28) appeared to cross-immunoreact with antisera Z59 and

Z69. However, they also cross-immunoreacted with polyclonal antiserum Z30. Antiserum Z30 was raised against only the 80 kDa peptide from goats and did not react with β-mannosidase. The 80 kDa protein in goat was relatively abundant in crude homogenate and Con A fractions. It did not express β-mannosidase enzyme activity and co-purified in small amounts with β-mannosidase protein (Sopher *et al.*, 1992; 1993). As antiserum Z28 contained some anti-80 kDa antibodies, it was possible that clones G#8Z28 (G#11Z28), G#16Z28, and G#22Z28 were picked up by the contaminated anti-80 kDa antibodies since they did cross-immunoreact to antiserum Z30. However, antiserum Z30 did not recognize β-mannosidase peptides, which suggested that the 80 kDa peptide did not contain similar epitopes to the β-mannosidase protein. The cross-immunoreaction to antisera Z59 and Z69 observed in the above clones implied that these clones were more likely to represent other proteins sharing the same or similar epitopes with β-mannosidase as well as the 80 kDa protein. That these clones were not β-mannosidase candidates was further demonstrated by their negative immunoreaction to affinity purified anti-β-mannosidase antibody and monoclonal antibodies. This conclusion was further supported by the lack of hybridization of these clones with several different β-mannosidase oligonucleotides obtained later on.

3.1.4.2 Screening with antisera Z59 and Z69

Polyclonal antisera Z59 and Z69 were produced by immunizing a rabbit with deglycosylated 100 kDa and 110 kDa bovine peptides (Sopher *et al.*, 1993). They had little cross-immunoreaction to the 80 kDa bovine peptide. The specificity of an antibody is important for successful immunoscreening. The only positive clone recognized by antiserum Z59 upon screening 10^6 plaques from bovine liver λZAPII cDNA library could not be identified by antiserum Z69. Preliminary epitope mapping suggested that Z69 recognized the same β-mannosidase epitopes as Z59 did and recognized more peptides than Z59 (data not shown). Therefore, it was unlikely that the clone, recognized by Z59 but without cross-immunoreaction with the antiserum Z69, would encode β-mannosidase. The whole antiserum Z69 and its IgG fraction appeared to contain a quite large amount of preimmune-antibodies. Most of the positive clones identified by Z69-IgG through rescreening the goat kidney cDNA library showed strong cross-immunoreaction with preimmunoserum of Z69. Two clones containing some cross-immunoreaction with Z59 antiserum and one without the cross-immunoreaction were chosen for further analysis. Although no obvious fusion proteins were seen by Ponceau S staining, multiple smear bands were observed by immunostaining with various antisera. The immunostaining pattern of each clone detected with Z69-IgG and preimmunoserum of Z69 was very

similar. The possibility of inefficient blocking of non-specific protein binding sites was implausible since two internal controls of protein lysates (pBluescript plasmid without insert and clone G#8Z28) prepared at the same time did not show the smear pattern. This result implies that these smear patterns were related to the clones isolated by Z69-IgG.

For successful immunoscreening, there are two crucial factors (Helfman and Hughes, 1987). First, the library should be large enough to contain sufficient copies of the gene of interest because not all the recombinant clones will express fusion proteins. In theory, only one out of six recombinant clones may react with antisera. Secondly, the antibody should be specific in recognizing epitopes of the proteins of interest, and of high titer. The failure to isolate the β-mannosidase gene by polyclonal antisera was probably due to the low abundance of expression of the gene. Screening with oligonucleotides in the following section indicated that the frequency of the β-mannosidase gene in the thyroid library was approximately 0.001%. The activities of the β-mannosidase protein in different tissues have been studied (Lovell *et al.*, 1994). Thyroid displays the highest enzyme activity, followed by kidney and liver which were about four fold less than thyroid. Therefore, the frequency of the β-mannosidase gene in a kidney or a liver library are expected to be lower than 0.001%. This

would limited the success of immunoscreening. The polyclonal antisera used here were able to detect 10 ng purified protein with a suitable signal at about 250-fold dilution. However, antiserum Z28 had a significant amount of cross-immunoreaction with the 80 or 84 kDa peptide. This also affected the success of the immunoscreening.

3.1.5 Summary

Both goat kidney and bovine liver cDNA libraries were screened by polyclonal antibodies specific for β-mannosidase. Twenty nine putative clones from three independent screenings were identified. Cross-immunoreaction with the other polyclonal antibodies different from that used in the screening was evaluated to determine the authenticity of any of these positive clones for β-mannosidase. The results excluded the candidacy of these clones. Subsequently, Southern hybridization of some of these clones with oligonucleotides obtained later further supported that they did not represent the β-mannosidase gene. The failure to isolate β-mannosidase cDNA by immunoscreening was most likely due to the low level of expression of the β-mannosidase gene in goat kidney and bovine liver libraries.

3.2 ISOLATION AND CHARACTERIZATION OF BOVINE β-MANNOSIDASE CDNA

3.2.1 Introduction

After failing to isolate β-mannosidase cDNA by immunoscreening, efforts were directed to peptide sequencing of the β-mannosidase protein. Previous studies suggested that the N-terminus of β-mannosidase was blocked (Sopher, 1992a). Several approaches had been tried before. Bovine β-mannosidase protein was purified by a four-step purification procedure (Sopher *et al.*, 1993). After deglycosylation, the β-mannosidase protein was further purified by SDS-PAGE. The major 86 kDa β-mannosidase was subjected to CNBr digestion. Three peptides were sequenced. However, these peptides yielded a limited sequence information. One produced a sequence with six amino acid residues. The second contained two different peptides, and the third yielded only a partial sequence. By *in situ* digestion with V8 protease on SDS-PAGE followed with transfer to PVDF membrane, multiple amino acid assignments were obtained (Sopher, 1992a). Therefore to obtain large quantities of β-mannosidase peptide, the deglycosylation and gel elution were omitted. β-Mannosidase was purified approximately 15,000 fold using the four-step procedure. This purified protein was digested with CNBr and trypsin and analyzed by the Biotechnology Resource Laboratory at Yale

University. Eventually, we obtained more than a dozen peptide sequences. These amino acid sequences, together with the partial sequence obtained earlier (Sopher, 1992a), enabled us to pursue the isolation of cDNA clones by screening cDNA libraries with synthetic oligonucleotides as well as by PCR methods. Cloning a cDNA using synthetic oligonucleotide probes designed from a known amino acid sequence is a popular approach. There are two different types of oligonucleotide probes: one is a mixed short oligonucleotide and the other one is a guessmer. Both probes have been used successfully to isolate genes.

Recently, a PCR technique, known as mixed oligonucleotides primed amplification of cDNA (MOPAC), has been developed (Lee *et al.*, 1988). In this method, degenerate oligonucleotides based on one or two known peptide sequences are used in PCR reactions. Unlike those used for conventional screening, the mixed oligonucleotide probes used in the MOPAC system can be very degenerate. Mixed primers of up to thousands of combinations have been used to clone a cDNA (Lee and Caskey, 1990).

Another new PCR technique, known as rapid amplification of cDNA ends (RACE), was introduced first by Frohman *et al.* (1988) and has now been commercialized (BRL and Clontech). The RACE system includes 5′ and 3′ RACE systems. Even with limited information of amino acid sequences, the RACE system is capable of cloning cDNA ends rapidly.

The main advantages of the PCR techniques are their speed, ease, and sensitivity in comparison with the library screening methods. In this paper, the MOPAC method was used in an attempt to generate specific cDNA probes and as an alternative approach for cloning cDNAs. The RACE system was used when conventional screening methods were inadequate for cloning the 5′ region of β-mannosidase cDNA.

3.2.2 Experimental procedures

3.2.2.1 Partial amino acid sequencing

β-Mannosidase protein was purified from bovine kidney (Ada Beef Co., Ada, MI) according to the procedures described by Sopher *et al.* (1992 and 1993). Briefly, a crude homogenate was prepared from 2.4 kg of sliced bovine kidney tissue and subjected to a four-step chromatography purification procedure including Con A-Sepharose (Sigma, St. Louis, MO), immunoaffinity, TSK-butyl (Pharmacia LKB Biotechnology Inc. Piscataway, NJ), and cation exchange high performance liquid chromatography (HPLC, Mono S, Pharmacia LKB Biotechnology Inc., Piscataway, NJ). The purity of the protein preparation after the Mono S step was confirmed by Coomassie blue staining and Western analysis. The purified protein was dialyzed against 5 mM ammonium bicarbonate solution, lyophilized, and sent to Keck Foundation Biotechnology Resource Laboratory in Yale University for amino acid sequencing. Approximately 740 pmol of purified

protein (from a total of 1240 pmol) was subjected to CNBr/trypsin digestion. Peptides were separated by C18 reverse phase HPLC. Peptides which yielded a separable and high peak (Fig. 3. 3) were subjected directly to amino acid sequencing or were repurified by reverse phase HPLC prior to sequencing (Table 3.2).

3.2.2.2 Construction of synthetic oligonucleotide probes

Mixed and unique oligonucleotides were synthe18sized on 394 and 380 B DNA synthesizers (Applied Biosystems). Regions with minimal codon redundancy were chosen to construct mixed oligonucleotide probes. Guessmers were constructed according to the typical codon usage frequency of human protein (Lathe, 1985). Inosine was placed at residues with three or four fold codon degeneracies in the preparation of the MJ30 oligonucleotide (Table 3.3). Antisense oligonucleotides were constructed for PCR analysis. More than ten oligonucleotides were des'igned based on peptide sequences (Table 3.3 and 3.4). Gene specific oligonucleotides were designed based on the Primer Program (Scientific & Educational Software, PA) to avoid dimer formations and long stretches of G or C. Synthesized oligonucleotides were reprecipitated by two volumes of ethanol and 2 M NH_4OAc before being used in PCR. The 2 M NH_4OAc solution was substituted by 0.3 M NaOAc (pH 5.6) when the oligonucleotide was going to be labeled at the 5′ end.

Table 3.2 CNBr/Tryptic Peptide Sequences of β-Mannosidase

Peptide	Sequence
142r12	F T P I Y D N Y (M)
Peptide 1*	F T P I Y D
104r86	I I K D P Y Y R R (G)
180	T V F F Y P W K P (M/T)
151r72	L F I Y G A S D L H S D Q Q (M)
Peptide 2*	(G) (A) (S) (D) L (H) (S) F (E) (L) (Y) (P) (T) (K) (D) (Q)
218r24	(S) (F) F A P L L P V G F E D K D (T) (P) (R) (Q/P) (K) (G) (P) (A) (E)
Peptide 3*	? ? Y F A ? ? F F A ? ? L
103	(L) H Y F A R (D)
169r64	F K D T L Y L T Q V (M)
171	(T) E L E Q S F H V T S L A D T Y (A)
169r65	F A C A L Y P T D K D F (M)
169r61	F A C A L Y P T D K D F G

Asterisks are peptide sequences obtained previously (Sopher, 1992a). Amino acids in parentheses are uncertain or best guess residues. Peptides generating limited partial sequences or more than two peptide mixture sequences are not listed.

Table 3.3 Oligonucleotides Used For Screening of cDNA Libraries

Peptide	Probe	Oligonucleotide Sequence (5'→3')
218r24	MJ4	GGN TTT GAG GAT AAG GA C A C A
169r61	MJ63	CC AAA ATC CTT ATC NGT G G T G
104r86	MJ64	ATT AAG GAT CCN TAT TA C A C C A
142r12	MJ48	TA ATT ATC ATA AAT NGG G G G G T
218r24	MJ7	TCC TTC TTT GCC CCC CTG CTG CCT GTG GGC TTT GAG GAC AAG GA *** * * * * *
142r12	MJ30	ATG TTT ACI CCI ATI TAT GAT AAT TA C C C
171	MJ65	GAG CTG GAG CAG TCC TTC CAT GTG ACC TCC CTG GCT GAC ACC TAC * * * * * * *
151r72	MJ23	CTG TTC ATC TAT GGC GCC TCT GAC CTG CAC TCT GAC * * * * * * *

Oligonucleotide probes designed from antisense strands are marked with underlines. Guessmers are marked with double underlines. Nucleotides that did not match with cDNA sequence are marked with asterisks.

Table 3.4 Selected Oligonucleotide Primers Used In PCR Analysis

Location (peptide)	primer	Sequences 5'→ 3'
(103)	MJ66	CTggatccATG CTN CAT TAT TTT GC *Bam HI* C C C
(218r24)	MJ6	ACgtcgacTC CTT ATC CTC AAA NCC *Sal*I T G T G
(180)	MJ5	ACgtcgacGG CTT CCA NGG ATA AAA AAA *Sal*I T G G G
(142r12)	MJ9	CCN ATT TAT GAT AAT TA C C C C A
-1988 bp	MJ74	CCAGATGTCATTGAGCTGCC
-2008 bp	MJ73	GAAGACCAGGAAGGAGCTTG
	MJ81*	CGACAACATCACGAAGGTGGTA
	MJ82*	TGCTTTATCAGGCTGGACTTCA
-836 bp	MJ100	CACAAAGAGCTCAACAGTCCTTTC
-802 bp	MJ101	AGGTCAATGTTGTTTGTCTGCTGT
-658 bp	MJ110	GCtctagaAACATGAAGTAGTTCAGATGACAA *Xba*I

Restriction sites (lowercase) were added in some degenerate primers. MJ81 and MJ82 were designed from a human expressed sequence tag (ETS 01397). They were used with primers MJ73 and MJ74 to produce a fragment covering a cleaved internal *EcoRI* site. Primers MJ66 and MJ6 were used in an attempt to generate a probe for screening. Primers MJ100, MJ101, and MJ110 were applied in the 5' RACE system in order to clone the 5' region of β-mannosidase cDNA. Other degenerate primers were employed to characterize cDNA clones.

3.2.2.3 Labeling probes

The 5′ end labeling reaction was carried out by T4 polynucleotide kinase (Boehringer Mannheim Biotechnicals, Indianapolis, IN). A twenty microliter reaction contained ten pmol of oligonucleotide, 15 pmol of ^{32}P[r-ATP] (6000 Ci/mmole, 10 mCi/ml, NEN/Du Pont, Wilmington, DE), 1 × buffer (Boehringer Mannheim Biotechnicals), and ten units of T4 polynucleotide kinase. The mixture was incubated at 37°C for one hour. Unincorporated nucleotides were removed using an Nuctrap push column (Stratagene) following the manufacturer's instructions. cDNA fragments were labeled by the random primed method (Boehringer Mannheim Biotechnicals) using ^{32}P [α-dCTP] (3000 Ci/mmole, Amersham, Arlington Heights, IL).

3.2.2.4 cDNA library screening

Normal bovine thyroid tissue was provided to Clontech (Palo Alto, CA) to construct a λZAPII cDNA library. The bovine thyroid cDNA library, consisting of 1.2×10^6 independent clones with an average insert size of 2.0 kb, was plated at a density of 1×10^4 plaque forming-units per 150 mm petri dish. After 4-5 hours of incubation, a colony/plaque screen filter (137 mm) (NEN research products, Du Pont, Boston, MA) was overlaid on top agarose for 2-5 min. The filter, with the phage plaques facing up, was then transferred to a fresh plate and incubated overnight

(approximately 12 hours). Phage DNAs were denatured, neutralized by the standard procedures (Sambrook *et al.*, 1989) and fixed onto the filters by baking at 80°C for two hours. The filters were washed in a prewarmed (50°C) solution of 2 × SSC/0.5% SDS/50 mM EDTA (pH 8.0) prior to hybridization. The hybridization procedure was basically as described in Current Protocols in Molecular Biology (Ausubel *et al.*, 1989). The filters were prehybridized (at least two hours) and hybridized (2-3 days) at 46-48°C in a solution containing 3 M TMAC (Aldrich Chemical Co. Inc., Milwankee, WI or Sigma, St. Louis, MO) /0.1 M sodium phosphate buffer, pH 6.8/1 mM EDTA, pH 8.0/5 × Denhardt's solution (1% Ficoll/1% polyvinylpyrrolidone/1% bovine serum albumin) /0.6% SDS/100 μg/ml denatured herring sperm DNA (Boehringer Mannheim Biochemicals) in crystallizing dishes. Approximately 1-2 × 10^6 cpm/ml of end-labelled mixed oligonucleotide probes with specific activity at 2-10 × 10^6 cpm/pmol were used during hybridization. The hybridized filters were washed once with 3 M TMAC/50 mM Tris.Cl (pH 8.0)/0.2% SDS, and then washed at room temperature for 15 min, followed by washing exactly 1 hour with the same solution at 46-50°C. Finally, filters were washed in 2 × SSC/0.1% SDS at room temperature for 2-3 times, each for 10 min. Filters were then wrapped with Saran wrap and exposed to Kodak XOMAT-AR film (Eastman KODAK Co., Rochester, NY) at -80°C for 1-3 days. The hybridized filters were

sequentially probed with different oligonucleotides after removal of the previous hybridized probe. Probes were removed by incubating hybridized filters in 0.4 M NaOH for 30 min at 45°C and 0.1 × SSC solution containing 0.1% SDS and 0.2 M Tris.Cl (pH 7.5) for 30 min at 45°C. Putative positive plaques were purified by several rounds of rescreening at lower densities and excised as pBluescript plasmids according to the manufacture's instructions. In order to isolate a full-length cDNA, the 1.6 kb insert of clone #47MJ4 was isolated and labeled by the random primed method to a specific activity of 1.4×10^9 cpm/μg. Up to 1×10^6 cpm/ml of the denatured probe was added in a 5 × SSPE hybridization solution containing 50% formamide (Boehringer Mannheim Biotechnicals)/0.5% SDS/5 × Denhardt's solution/10 μg/ml denatured herring sperm DNA to reprobe the filters at 42°C. After approximately 20 hours of incubation, the filters were washed in 2 × SSC/0.1% SDS for two times, each 10 min at room temperature, then washed in 1 × SSC/0.1% SDS at 65°C for 30 min. Finally, the filters were washed in 0.1 × SSC/0.1% SDS at 65°C for 30 min.

3.2.2.5 Polymerase chain reaction (PCR)

Ten micrograms of total RNA or 1 μg of poly A^+ mRNA were reverse transcribed into single strand cDNAs. RNA in diethylpyrocarbonate (DEPC) treated water was first heated at 65°C for 5 min, then incubated with 10 μl of 5 × reverse

transcription buffer (Gibco BRL, Gaithersburg, MD), 5 μl of 0.1 M dithiothreitol (DTT), 5 μl of 10 mM dNTP mixture, 1.5 μl of 40 unit/μl ribonuclease inhibitor (rRNasin) (Promega, Madison, WI), 5 to 75 pmol of antisense oligonucleotides and 200-400 units of M-MLV reverse transcriptase (Gibco BRL) in a 50 μl of reaction mixture at 37°C for one hour. The first strand cDNAs were precipitated by adding an equal volume of 4 M ammonium acetate acid and two volumes of ethanol and resuspended in 50 μl of distilled water. An aliquot (1-2 μl) of cDNAs were amplified by AmpliTaq DNA polymerase (Perkin-Elmer Cetus, Norwalk, CT) in the GeneAmp PCR system 9600 (Perkin-Elmer Co.). Generally, 30-35 PCR cycles were performed and each PCR cycle consisted of 45 seconds denaturation at 94°C, 1 min annealing at 46°C-55°C, and a 1 min extension at 72°C of cDNA. A 5 min predenaturation at 95°C and an additional 10 min extension at 72°C were applied before and after the cycle reactions, respectively. One fourth of the amplified product was analyzed by electrophoresis on Nusieve 3:1 agarose gels (FMC BioProducts, Rockland, ME). To perform reamplification and nested PCR, the agarose containing interesting bands was removed using a capillary tube. The agarose was either directly used (5 μl) in PCR reactions or diluted in distilled water, then aliquots of agarose suspension were used in PCR reactions.

To analyze putative clones, PCR was performed using

either plasmid DNA or crude phage lysates as templates. Crude phage lysates were prepared by adding an equal volume of 0.1 M NaOH to an aliquot of phage stock, incubating for 10 min at 95°C, and then neutralizing by adding 1/20 volume of 2 M Tris.Cl solution (pH 7.5). To determine the orientation of an insert, PCR was accomplished by using a gene specific primer (either unique or degenerate) and a M13 primer (reverse or -21 primer).

In order to isolate the 5′ region of β-mannosidase cDNA, 0.5 μg poly A^+ mRNA of bovine thyroid was copied into single strand cDNAs and amplified using the 5′ RACE system kit (Gibco BRL, Gaithersburg, MD) according to the manufacture's instructions. Two gene specific antisense oligonucleotides (MJ100 and MJ101) designed from clone #17MJ48 and a degenerate oligonucleotide MJ48 from peptide 142r12 were used in the 5′ RACE system.

3.2.2.6 RNA isolation and Northern blot hybridization

Total RNA was extracted from various bovine and caprine tissues and from both normal and affected animals according to the procedure as described (Ausubel *et al.*, 1989). Poly A^+ RNA was isolated using poly A^+ quick mRNA kit (Stratagene).

RNA samples (gift from Dr. Karen Friderici) were analyzed by electrophoresis on a 1% agarose gel containing 2.2 M formaldehyde as described (Ausubel *et al.*, 1989) and

blotted on the Hybond-N membrane (Amersham). The hybridization condition is either based on the method described in Current Molecular Protocol (Ausubel *et al.*, 1989), or is the same as the Southern hybridization described in section 2.2.8. Approximately 1-2 × 10^6 cpm/ml of cDNA probe(s) with a specific activity of approximately 1 × 10^9 cpm/μg was used during the hybridization. Filters were washed in a step-wise fashion according to the background signal. The filters were exposed at -80°C for 3-10 days. After removal of the β-mannosidase probe, the blot was rehybridized to a cDNA probe of rat glyceraldehyde-3-phosphate-dehydrogenase (GAPDH) as an internal control.

3.2.2.7 DNA sequencing and computer analysis

Miniplasmid DNAs were prepared using the Magic (Wizard) minipreps DNA purification system (Promega, Madison, WI). Sequencing was carried out by the *Taq* cycling method using dye terminators and dye primers (M13 -21 primer and M13 reverse primer) on a 373A DNA sequencing system (Applied Biosystems). The entire inserts in clones #46MJ4, #47MJ4, #17MJ48, #11MJ101/UAP, and #2MJ48/UAP were sequenced in both orientations using internal oligonucleotide primers. To sequence PCR products, PCR products were separated on 1% Nusieve GTG low melting agarose gel (FMC BioProducts, Rockland, ME), bands of interest were excised under long wave-length UV light, and purified using Magic (Wizard) PCR

DNA purification system according to Promega's instructions. The purified PCR products were sequenced directly by the dye terminator sequencing method. In some cases, the PCR products were subcloned into PCR™ vector using a T/A cloning system (Invitrogene, San Diego, CA). Plasmid DNAs were then prepared by the Magic (Wizard) minipreps system and sequenced by dye dideoxyribonucleotide terminator. DNA sequence analysis and homology search against GenBank were performed using GCG program version 7, April, 1991 (Genetics Computer Group DNA Sequence Analysis Software, Madison, Wisconsin).

3.2.2.8 Southern hybridization of chromosome blot and zoo blot

A DNA panel of 24 human/rodent somatic cell hybrids was obtained from Coriell Cell Repositories, Coriell Institute for Medical Research (Camden, NJ). Except for hybrids NA07299 and NA10478, all hybrids retain one human chromosome under either mouse or Chinese hamster background. Fifteen micrograms of DNA isolated from each of the hybrids were digested in 100 μl reactions by restriction enzyme *PstI* (120 units) at 37°C. After overnight digestion, DNA was reprecipitated by two volumes of ethanol and 0.3 M NaOAc (pH 5.6) and separated in a 1% agarose gel at 25 V for 24 hours. DNA was transferred overnight to Hybond-N membrane (Amersham Co., Arlington Heights, IL). A PCR product of clone #46 MJ4

was labeled by the random primed method to a specific activity of 4 × 10^8 cpm/μg. Approximately 3 × 10^6 cpm/ml of the probe was denatured and added to 15 ml of hybridization solution. Hybridization was carried out in 50% formamide/6 × SSC/5 × Denhardt's/0.5% SDS/100 μg/ml denatured herring sperm DNA for 20 hours at 42°C. The final wash was in 0.2 × SSC/0.1% SDS for 15 min at 65°C (high stringency wash) or 1 × SSC/0.1% SDS for one hour at 65°C (low stringency wash). The blot was exposed to Kodak XOMAT-AR film at -80°C for three to ten days. Genomic DNAs from different species and affected animals were prepared previously by Sopher (1992a). The restriction enzyme digestion, gel electrophoresis, DNA transfer, and hybridization were performed as described above.

3.2.3 Results

3.2.3.1 Peptide sequencing

Bovine protein was purified to approximately 15,000 fold by a four-step-chromatography procedure. The final protein preparation revealed three peptides 84, 100, and 110 kDa as judged by Coomassie blue (Fig. 3.2). There was little, if any, 80 kDa peptide as judged by Western analysis (Fig. 3.2) using the antiserum Z30 which reacts specifically with caprine 80 kDa protein or bovine 80 kDa protein. Approximately 740 pmol of this purified protein was subjected to CNBr/tryptic digestion. Fourteen peptides

including those repurified by reverse phase HPLC were sequenced (Fig. 3.3). This resulted in a total of ten non-overlapping peptides with complete sequences plus additional peptides with incomplete or uncertain sequences. The sequence of the ten peptides is shown on Table 3.2.

3.2.3.2 Isolation and characterization of cDNA clones

After the failure of immunoscreening, preliminary oligonucleotide screening with several oligonucleotide probes by plaque hybridization, and PCR with various combinations of degenerate oligonucleotide primers were performed without success (data not shown). In an effort to increase detection levels, *in situ* amplification of plaques was used to intensify the signal. Furthermore, a library was obtained from bovine thyroid tissue showing the highest expression of β-mannosidase activity. In addition, oligonucleotides (e.g. MJ4) that gave a good signal in preliminary studies or that were derived from peptide sequences (e.g. 142r12) produced by two different sequencing sources were used for screening. By screening approximately 5×10^5 phage from the bovine thyroid cDNA library sequentially with four different degenerate oligonucleotides, a total of 19 positive clones were detected. Among these positive clones, seventeen clones were identified by oligonucleotide probe MJ4 and two by oligonucleotide probe MJ48. No positive clones were found

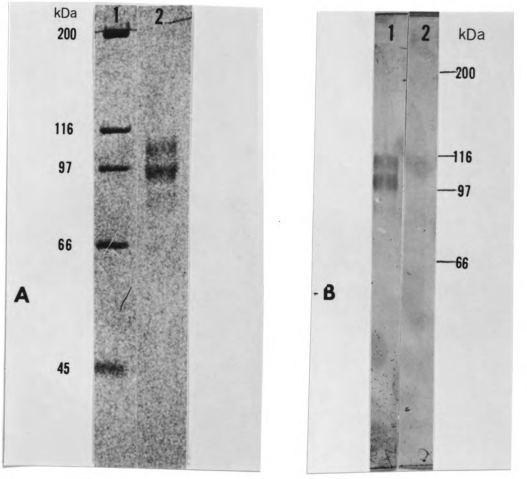


Fig. 3.2 Panel A: Coomassie staining of purified β-mannosidase peptides. Fifty microliters (approximately 1/100) of purified β-mannosidase protein was fractionated on 7.5% SDS-PAGE (lane 2). Lane 1, High range SDS-PAGE standard (Bio-Rad). **Panel B: Western analysis of purified β-mannosidase peptides with antisera Z59 and Z30.** Lane 1 and 2, Containing 1 μg of purified protein were reacted with antisera Z59 and Z30 respectively.

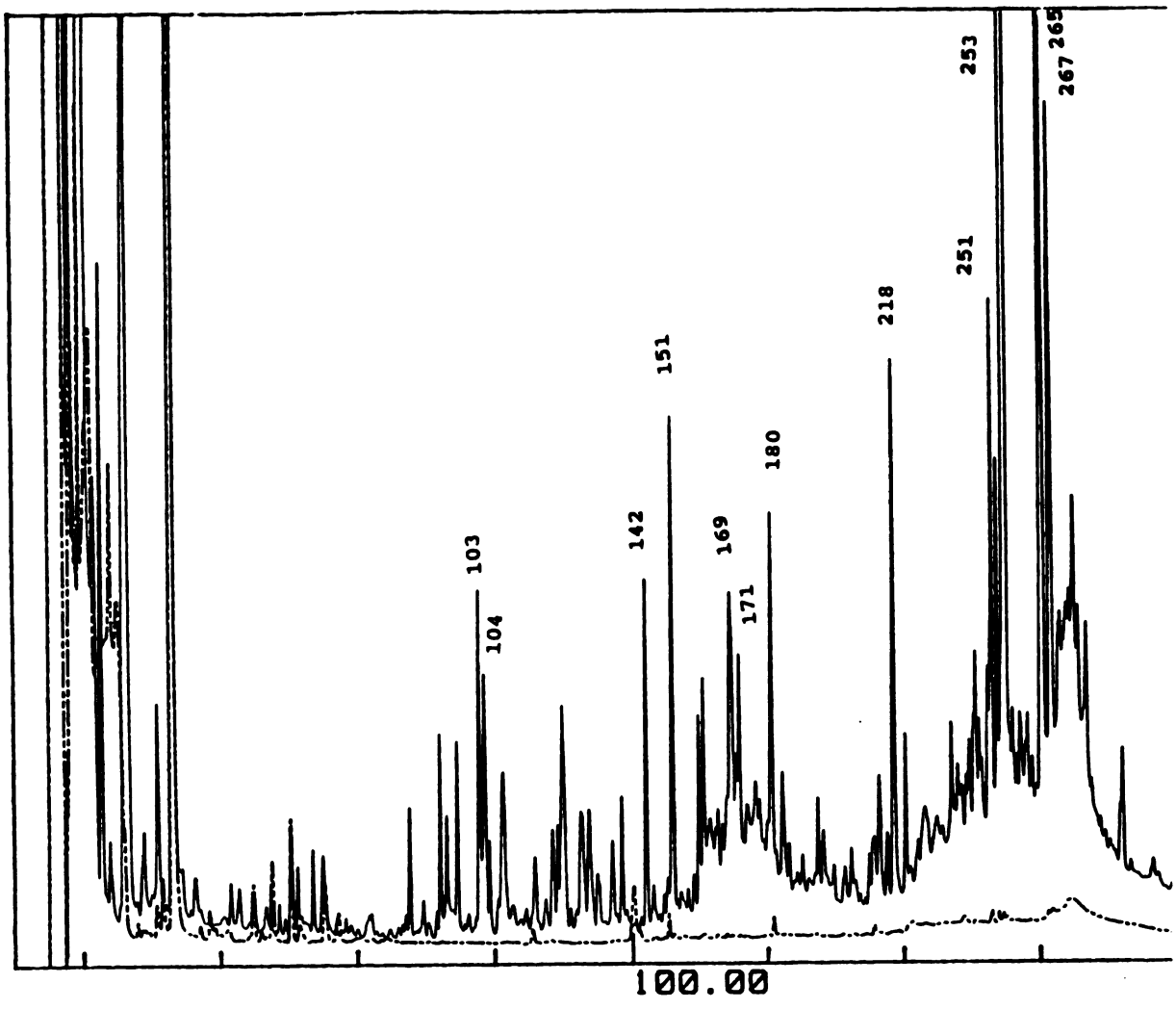


Fig. 3.3 The reverse phase HPLC profile of CNBr/tryptic cleaved peptides of β-mannosidase. Approximately 750 pmol β-mannosidase protein purified by a four-step column purification procedure were subjected to CNBr and trypsin digestions and separated by C18 reverse phase HPLC. Peptides 103, 171, 180, and 253 were sequenced directly. Peptides 104, 142, 151, 169, 218, 251, 265, and 267 were subjected to repurification before sequencing. No sequence was yielded by sequencing peptides 252, 265, and 267. Mixed sequences were obtained from peptide 253.

when reprobing the filters with two other oligonucleotide probes: MJ63 and MJ64. Of the 19 positive clones, clones #43MJ4, #46MJ4, and #47MJ4 were also found to hybridize with guessmers corresponding to three different non-overlapping peptides (i.e. MJ7, MJ23, and MJ65). The three clones were plaque purified, subcloned into pBluescript plasmid, and subjected to further analysis by restriction enzyme digestion and sequencing. The results indicated that clones #43MJ4 and #47MJ4 were identical containing a 1.6 kb insert with identical restriction maps, while clone #46MJ4 contained an 1875 bp insert (Fig. 3.4). Clones #43MJ4, #47MJ4, and #46MJ4 all started at the same 5′ region at a cleaved *EcoR*I site and contained 735 bp of open reading frame. There were two nucleotide differences between the two clones: C at position 2124 was replaced by G in clone #47MJ4, while C at position 2417 was substituted by T in clone #47MJ4 (Fig. 3.5). The C/T substitution was neutral. The C/G substitution changes an amino acid residue from histidine (H) to the aspartic acid (D), found in that position in the direct peptide sequence of peptide 151r72. The homology between clone #46MJ4 and #47MJ4 diverged at 1182 bp from their 5′ ends. The authenticity of the three clones #46MJ4, #47MJ4, and #43MJ4 was established by colinearity of predicted amino acid sequence of these clones with five microsequenced peptide sequences (103, 218r24, 151r72, 180, and 171) (Table 3.2).

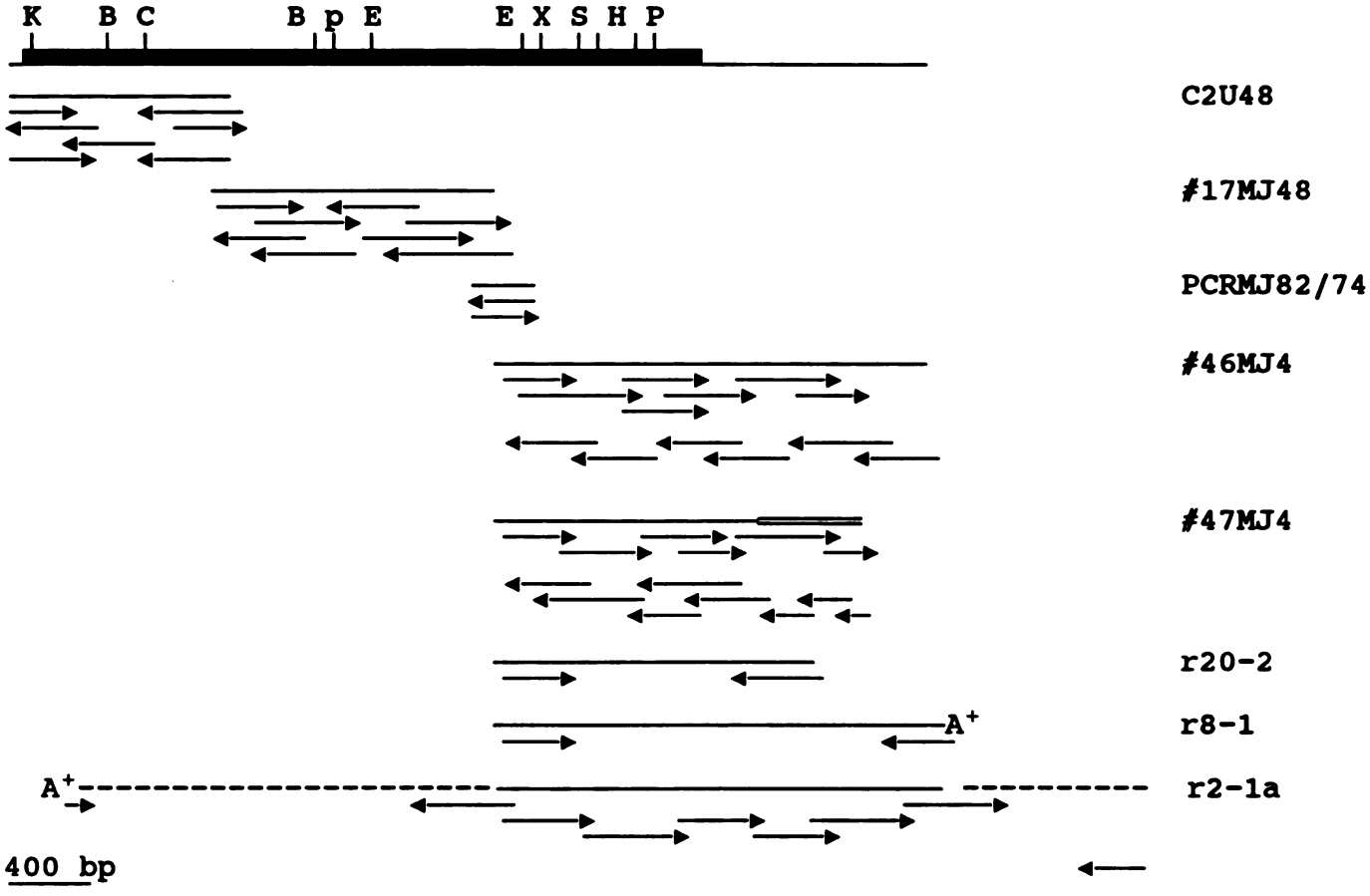


Fig. 3.4 The restriction map and sequencing strategy for β-mannosidase cDNA clones. C2U48, a subclone of 5′ RACE product; PCRMJ82/74, a PCR product generated using the human ETS01379 sequence information; solid bar, coding region; double line and dash line, no homologies with #46MJ4; K, *Kpn*I; B, *Bam*HI; C, *Cla*I; E, *Eco*RI; X, *Xba*I; S, *Sal*I; H, *Hinc*II; p, *Pst*I; P, *Pvu*II. A^+, poly (A) tail

Fig. 3.5 Nucleotide and deduced amino acid sequences of β-mannosidase cDNA. Nucleotides upstream of the predicted initiation codon ATG are given negative numbers. Potential N-glycosylation sites are indicated by ★. Colinear CNBr/Tryptic peptides are underlined. Residues which do not match with peptide sequences determined by microsequencing are marked with [ ]. Two possible polyadenylation sites are marked with ▬▬. Signal peptide sequence is underlined. The arrow indicates the predicted signal peptide cleavage site.

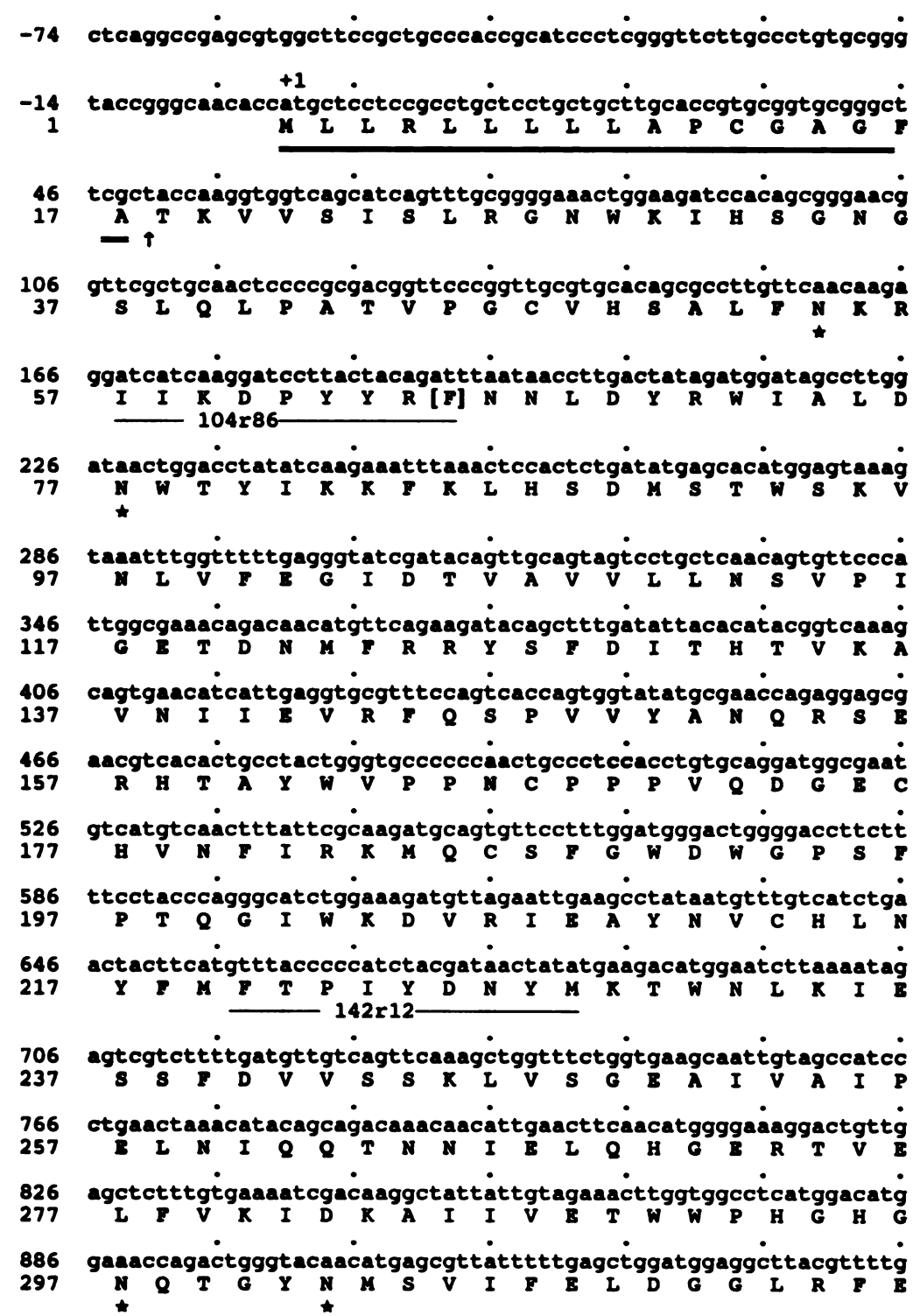

-74 ctcaggccgagcgtggcttccgctgcccaccgcatccctcgggttcttgccctgtgcggg
+1
-14 taccgggcaacaccatgctcctccgcctgctcctgctgcttgcaccgtgcggtgcgggct
1 M L L R L L L L L A P C G A G F
46 tcgctaccaaggtggtcagcatcagtttgcggggaaactggaagatccacagcgggaacg
17 A T K V V S I S L R G N W K I H S G N G
106 gttcgctgcaactccccgcgacggttcccggttgcgtgcacagcgccttgttcaacaaga
37 S L Q L P A T V P G C V H S A L F N K R
166 ggatcatcaaggatccttactacagatttaataaccttgactatagatggatagccttgg
57 I I K D P Y Y R [F] N N L D Y R W I A L D
104r86
226 ataactggacctatatcaagaaatttaaactccactctgatatgagcacatggagtaaag
77 N W T Y I K K F K L H S D M S T W S K V
286 taaatttggtttttgagggtatcgatacagttgcagtagtcctgctcaacagtgttccca
97 N L V F E G I D T V A V V L L N S V P I
346 ttggcgaaacagacaacatgttcagaagatacagctttgatattacacatacggtcaaag
117 G E T D N M F R R Y S F D I T H T V K A
406 cagtgaacatcattgaggtgcgtttccagtcaccagtggtatatgcgaaccagaggagcg
137 V N I I E V R F Q S P V V Y A N Q R S E
466 aacgtcacactgcctactgggtgccccccaactgccctccacctgtgcaggatggcgaat
157 R H T A Y W V P P N C P P P V Q D G E C
526 gtcatgtcaactttattcgcaagatgcagtgttcctttggatgggactggggaccttctt
177 H V N F I R K M Q C S F G W D W G P S F
586 ttcctacccagggcatctggaaagatgttagaattgaagcctataatgtttgtcatctga
197 P T Q G I W K D V R I E A Y N V C H L N
646 actacttcatgtttacccccatctacgataactatatgaagacatggaatcttaaaatag
217 Y F M F T P I Y D N Y M K T W N L K I E
142r12
706 agtcgtcttttgatgttgtcagttcaaagctggtttctggtgaagcaattgtagccatcc
237 S S F D V V S S K L V S G E A I V A I P
766 ctgaactaaacatacagcagacaaacaacattgaacttcaacatggggaaaggactgttg
257 E L N I Q Q T N N I E L Q H G E R T V E
826 agctctttgtgaaaatcgacaaggctattattgtagaaacttggtggcctcatggacatg
277 L F V K I D K A I I V E T W W P H G H G
886 gaaaccagactgggtacaacatgagcgttatttttgagctggatggaggcttacgttttg
297 N Q T G Y N M S V I F E L D G G L R F E

Fig. 3.5

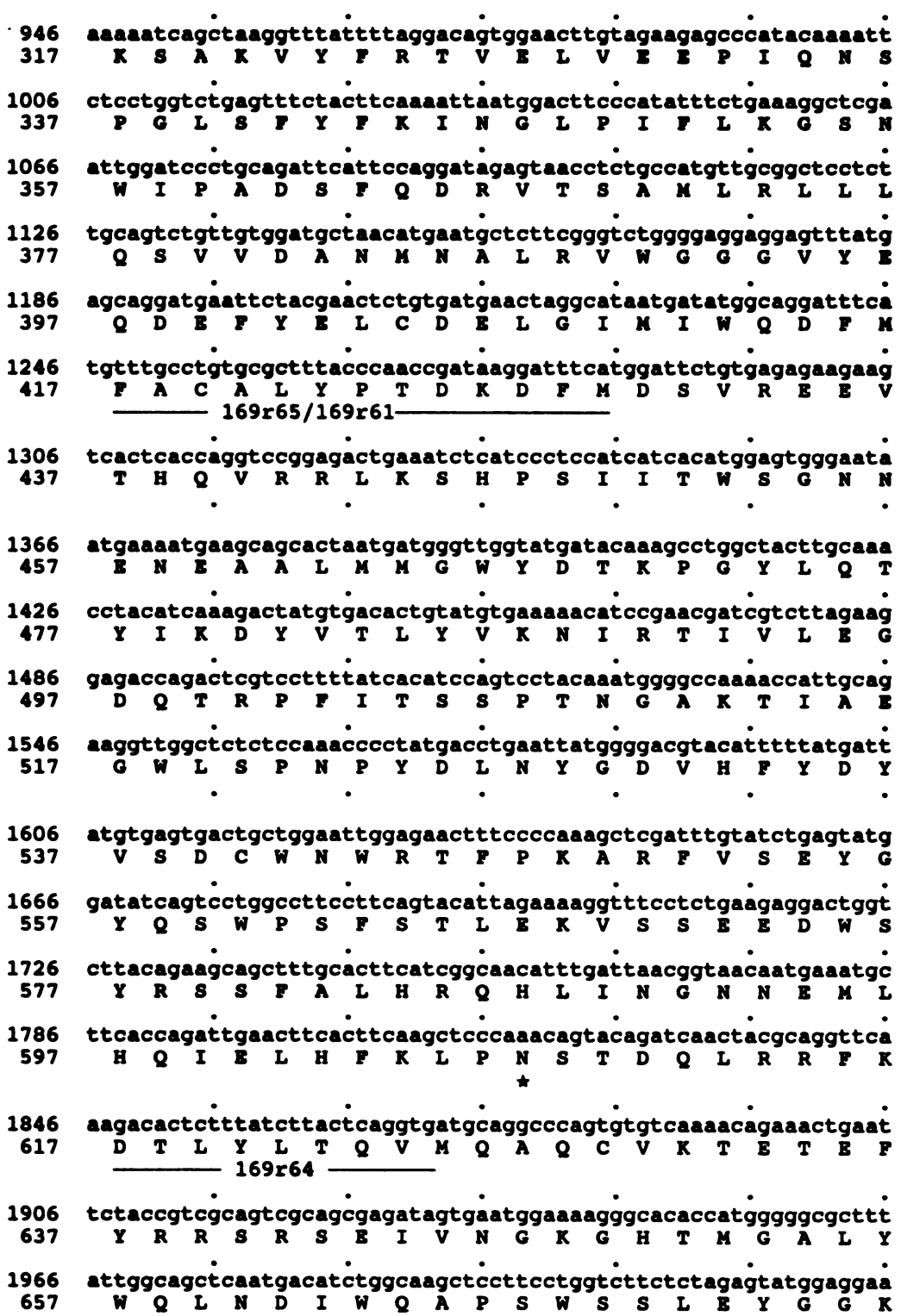

946 aaaaatcagctaaggtttattttaggacagtggaacttgtagaagagcccatacaaaatt
317 K S A K V Y F R T V E L V E E P I Q N S
1006 ctcctggtctgagtttctacttcaaaattaatggacttcccatatttctgaaaggctcga
337 P G L S F Y F K I N G L P I F L K G S N
1066 attggatccctgcagattcattccaggatagagtaacctctgccatgttgcggctcctct
357 W I P A D S F Q D R V T S A M L R L L L
1126 tgcagtctgttgtggatgctaacatgaatgctcttcgggtctggggaggaggagtttatg
377 Q S V V D A N M N A L R V W G G G V Y E
1186 agcaggatgaattctacgaactctgtgatgaactaggcataatgatatggcaggatttca
397 Q D E F Y E L C D E L G I M I W Q D F M
1246 tgtttgcctgtgcgctttacccaaccgataaggatttcatggattctgtgagagaagaag
417 F A C A L Y P T D K D F M D S V R E E V
169r65/169r61
1306 tcactcaccaggtccggagactgaaatctcatccctccatcatcacatggagtgggaata
437 T H Q V R R L K S H P S I I T W S G N N
1366 atgaaaatgaagcagcactaatgatgggttggtatgatacaaagcctggctacttgcaaa
457 E N E A A L M M G W Y D T K P G Y L Q T
1426 cctacatcaaagactatgtgacactgtatgtgaaaaacatccgaacgatcgtcttagaag
477 Y I K D Y V T L Y V K N I R T I V L E G
1486 gagaccagactcgtccttttatcacatccagtcctacaaatggggccaaaaccattgcag
497 D Q T R P F I T S S P T N G A K T I A E
1546 aaggttggctctctccaaaccccctatgacctgaattatggggacgtacatttttatgatt
517 G W L S P N P Y D L N Y G D V H F Y D Y
1606 atgtgagtgactgctggaattggagaactttccccaaagctcgatttgtatctgagtatg
537 V S D C W N W R T F P K A R F V S E Y G
1666 gatatcagtcctggccttccttcagtacattagaaaaggtttcctctgaagaggactggt
557 Y Q S W P S F S T L E K V S S E E D W S
1726 cttacagaagcagctttgcacttcatcggcaacatttgattaacggtaacaatgaaatgc
577 Y R S S F A L H R Q H L I N G N N E M L
1786 ttcaccagattgaacttcacttcaagctcccaaacagtacagatcaactacgcaggttca
597 H Q I E L H F K L P N S T D Q L R R F K
★
1846 aagacactctttatcttactcaggtgatgcaggcccagtgtgtcaaaacagaaactgaat
617 D T L Y L T Q V M Q A Q C V K T E T E F
169r64
1906 tctaccgtcgcagtcgcagcgagatagtgaatggaaaagggcacaccatgggggcgcttt
637 Y R R S R S E I V N G K G H T M G A L Y
1966 attggcagctcaatgacatctggcaagctccttcctggtcttctctagagtatggaggaa
657 W Q L N D I W Q A P S W S S L E Y G G K

Fig 3.5

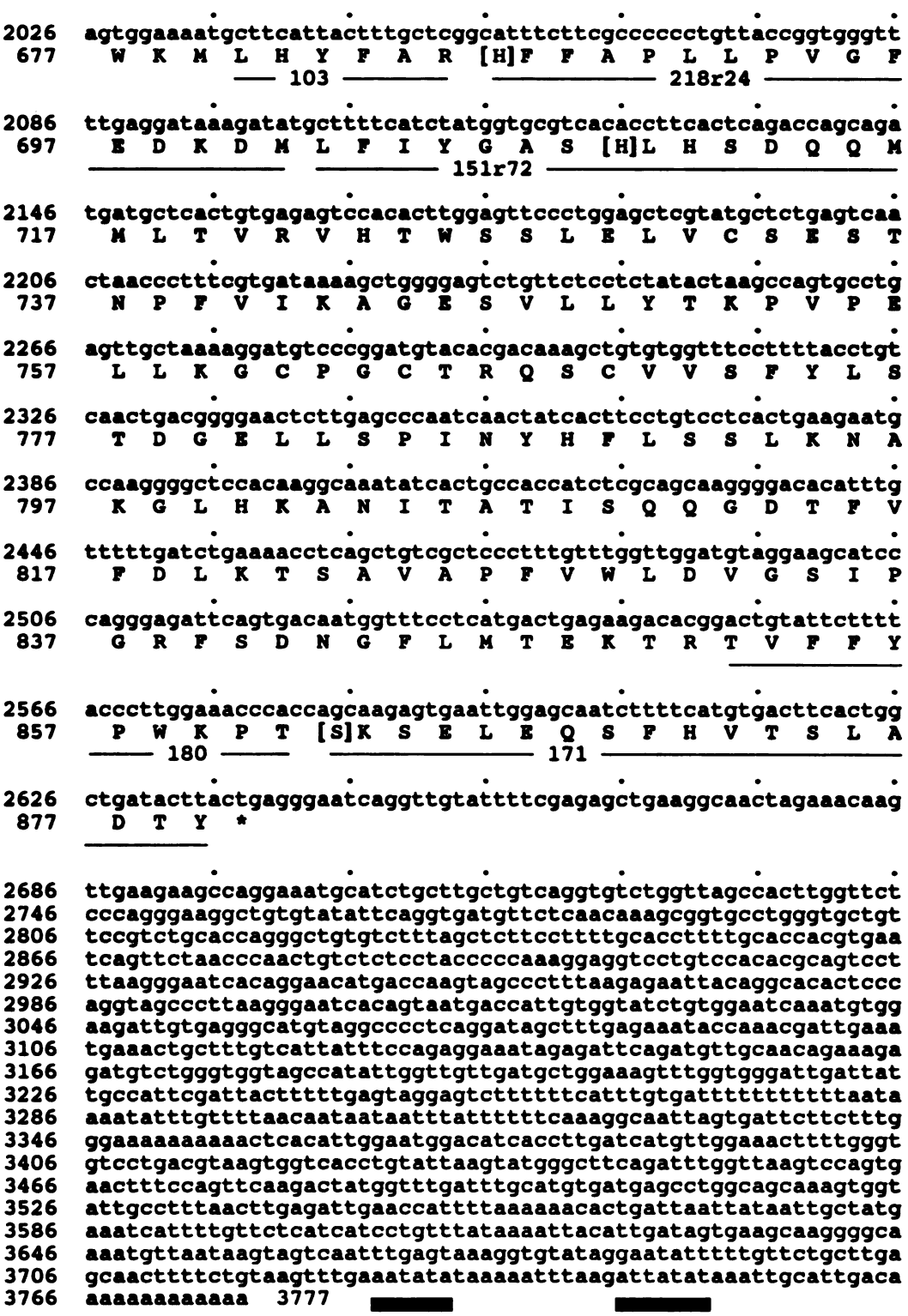

2026 agtggaaaatgcttcattactttgctcggcatttcttcgccccctgttaccggtgggtt
677 W K M L H Y F A R [H]F F A P L L P V G F
103
218r24
2086 ttgaggataaagatatgcttttcatctatggtgcgtcacaccttcactcagaccagcaga
697 E D K D M L F I Y G A S [H]L H S D Q Q M
151r72
2146 tgatgctcactgtgagagtccacacttggagttccctggagctcgtatgctctgagtcaa
717 M L T V R V H T W S S L E L V C S E S T
2206 ctaacccttcgtgataaaagctggggagtctgttctcctctatactaagccagtgcctg
737 N P F V I K A G E S V L L Y T K P V P E
2266 agttgctaaaaggatgtcccggatgtacacgacaaagctgtgtggtttccttttacctgt
757 L L K G C P G C T R Q S C V V S F Y L S
2326 caactgacggggaactcttgagcccaatcaactatcacttcctgtcctcactgaagaatg
777 T D G E L L S P I N Y H F L S S L K N A
2386 ccaaggggctccacaaggcaaatatcactgccaccatctcgcagcaaggggacacatttg
797 K G L H K A N I T A T I S Q Q G D T F V
2446 tttttgatctgaaaacctcagctgtcgctccctttgtttggttggatgtaggaagcatcc
817 F D L K T S A V A P F V W L D V G S I P
2506 cagggagattcagtgacaatggtttcctcatgactgagaagacacggactgtattctttt
837 G R F S D N G F L M T E K T R T V F F Y
2566 acccttggaaacccaccagcaagagtgaattggagcaatcttttcatgtgacttcactgg
857 P W K P T [S]K S E L E Q S F H V T S L A
180
171
2626 ctgatacttactgagggaatcaggttgtattttcgagagctgaaggcaactagaaacaag
877 D T Y *
2686 ttgaagaagccaggaaatgcatctgcttgctgtcaggtgtctggttagccacttggttct
2746 cccagggaaggctgtgtatattcaggtgatgttctcaacaaagcggtgcctgggtgctgt
2806 tccgtctgcaccagggctgtgtctttagctcttccttttgcacctttgcaccacgtgaa
2866 tcagttctaacccaactgtctctcctacccccaaaggaggtcctgtccacacgcagtcct
2926 ttaagggaatcacaggaacatgaccaagtagccctttaagagaattacaggcacactccc
2986 aggtagcccttaagggaatcacagtaatgaccattgtggtatctgtggaatcaaatgtgg
3046 aagattgtgagggcatgtaggcccctcaggatagctttgagaaataccaaacgattgaaa
3106 tgaaactgctttgtcattatttccagaggaaatagagattcagatgttgcaacagaaaga
3166 gatgtctgggtggtagccatattggttgttgatgctggaaagtttggtgggattgattat
3226 tgccattcgattacttttgagtaggagtctttttcatttgtgatttttttttttaata
3286 aaatatttgttttaacaataataatttatttttcaaaggcaattagtgattcttctttg
3346 ggaaaaaaaaaactcacattggaatggacatcaccttgatcatgttggaaacttttgggt
3406 gtcctgacgtaagtggtcacctgtattaagtatgggcttcagatttggttaagtccagtg
3466 aactttccagttcaagactatggtttgatttgcatgtgatgagcctggcagcaaagtggt
3526 attgcctttaacttgagattgaaccattttaaaaaacactgattaattataattgctatg
3586 aaatcattttgttctcatcatcctgtttataaaattacattgatagtgaagcaaggggca
3646 aaatgttaataagtagtcaatttgagtaaaggtgtataggaatatttttgttctgcttga
3706 gcaacttttctgtaagtttgaaatatataaaaatttaagattatataaattgcattgaca
3766 aaaaaaaaaaaa 3777

Fig. 3.5

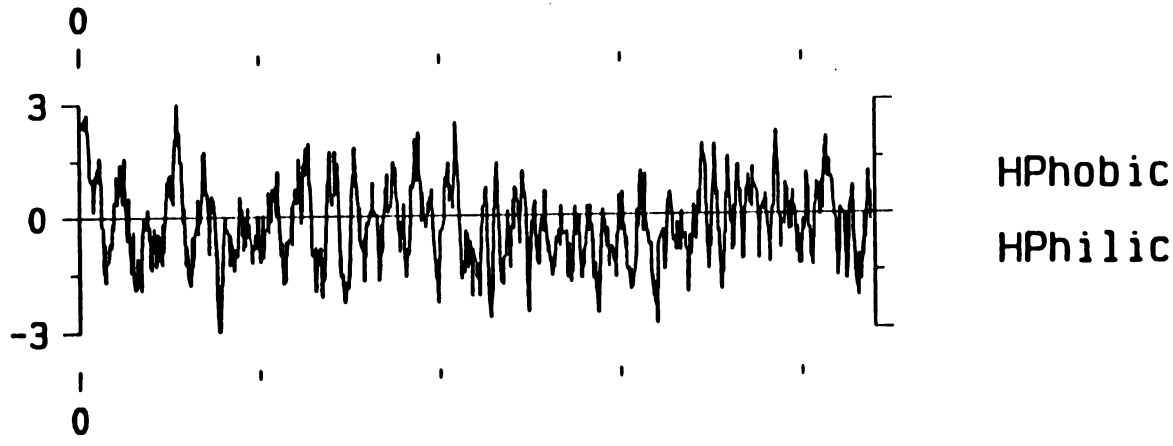


Fig. 3.6 The hydropathy plot of the β-mannosidase polypeptides predicted from the bovine β-mannosidase cDNA. The hydropathy plot was based on Kyte and Doolittle's method (Kyte and Doolittle., 1982). The window size was 7. No membrane spanning peptides were predicted.

In a parallel experiment, PCR was used to generate a gene specific probe for screening the cDNA library. One of earlier peptides generated by sequencing of CNBr cleaved peptides (peptide 3, Table 3.2) of the deglycosylated 86 kDa protein (Sopher, 1992a) was found to overlap two of the peptides (218r24 and 103, Table 3.2) in this study. We suspected that the two peptides 218r24 and 103 were continuous. Therefore, a sense primer MJ66 and an antisense primer MJ6 (Table 3.4) were used to amplify a bovine thyroid cDNA reverse transcribed by oligonucleotide primer MJ6. As was predicted from the amino acid sequence, multiple PCR products were produced, including one with expected size of 81 bp. The 81 bp product was gel purified and reamplified. The single 81 bp product was then repurified and subcloned into PCR™ vector using the T/A cloning system. Several subclones containing the 81 bp product were sequenced. The deduced amino acid sequence showed colinearity with the peptide sequence. This approach was not pursued since these data became available at the same time the large clones described above were obtained.

In order to isolate a full-length cDNA, the 1.6 kb insert of clone #47MJ4 was gel purified and used as a probe to rescreen the original filters. Three additional clones (r2-1a, r8-1, and r20-2) were isolated and excised into pBluescript plasmid. *EcoR*I digestion of the plasmid DNAs indicated clone #r20-2 lacked an *EcoR*I site in one of the

cloning sites and contained an insert size of approximately 1.4 kb. The insert size of clone #r8-1 was close to 1.8 kb (Fig. 3.4). Clone #r2-1a appeared to have a large insert of approximately 4.3 kb, which was confirmed by Southern hybridization of *EcoRI* digested plasmid and phage DNA. To analyze whether any of these clones contained sequences upstream of the 5′ ends of clones #46MJ4 and #47MJ4, a specific oligonucleotide (MJ74) (Table 3.4) was designed according to the sequence located 105 bp downstream from the 5′ end of clone #46MJ4. This gene specific oligonucleotide was used with vector primers close to the cloning sites to prime PCR reactions of either crude phage lysates or plasmid DNAs of clones #r2-1a, #r8-1, and #r20-2. The PCR results indicated that clone #r2-1a contained approximately 1.2 kb more sequence than the 5′ end of clones #46MJ4 and #47MJ4. The other two clones appeared to start at the same internal *EcoRI* site as clones #46MJ4 and #47MJ4 (Fig. 3.4). Further studies by PCR using either gene specific (MJ73, MJ74) or mixed oligonucleotide primers (MJ66, MJ5) located downstream of the *EcoRI* site of clone#46MJ4 suggested that clones #r2-1a, #r8-1, and #r20-2 had sequence homologies with clones #46MJ4 and #47MJ4. Partial sequencing of these clones confirmed that clones #r8-1 and #20-2 did indeed start at the same 5′ position as clones #46MJ4 and #47MJ4, corresponding to a cleaved *EcoRI* site. Their 3′ end sequences were nearly identical to that of clone #46MJ4.

Clone #r8-1 contains additional 20 base pairs including a poly (A) tail (Fig. 3.5). Partial sequencing of clone #r2-1a with M13 reverse and forward primers and some internal primers derived from the DNA sequence of clone #46MJ4 demonstrated that it encompassed most of the sequence of clone #46MJ4. However, the sequence homology diverged at 86 bp upstream of the 3′ end of clone #46MJ4. A long stretch of poly (A) tail was present in the 5′ end of clone r2-1a, and no open reading frame was revealed in the region upstream the *EcoR*I site of clone #46MJ4. These results clearly indicated that the *EcoR*I sites of these clones had not been protected by *EcoR*I methylase during the construction of the bovine thyroid cDNA library.

Since an internal *EcoR*I site of β-mannosidase cDNA was cleaved, these clones were not suitable as probes to rescreen the bovine thyroid cDNA library. Surprisingly, a sequence homology search disclosed that there was high homology between sequence downstream of the *EcoR*I site of a human expressed sequence tag to an unknown gene (EST01397) (Adams *et al.*, 1992) from GenBank and the 5′ ends of clones #46MJ4 and #47MJ4 (Fig. 3.7). The insertion of one bp at positions 284 and 305 of the human clone shifted the reading frame twice in a human expressed sequence tag cDNA (Fig. 3.7). We speculated that the human expressed sequence tag obtained from a hippocampal cDNA library was derived from human β-mannosidase. Using this information, two

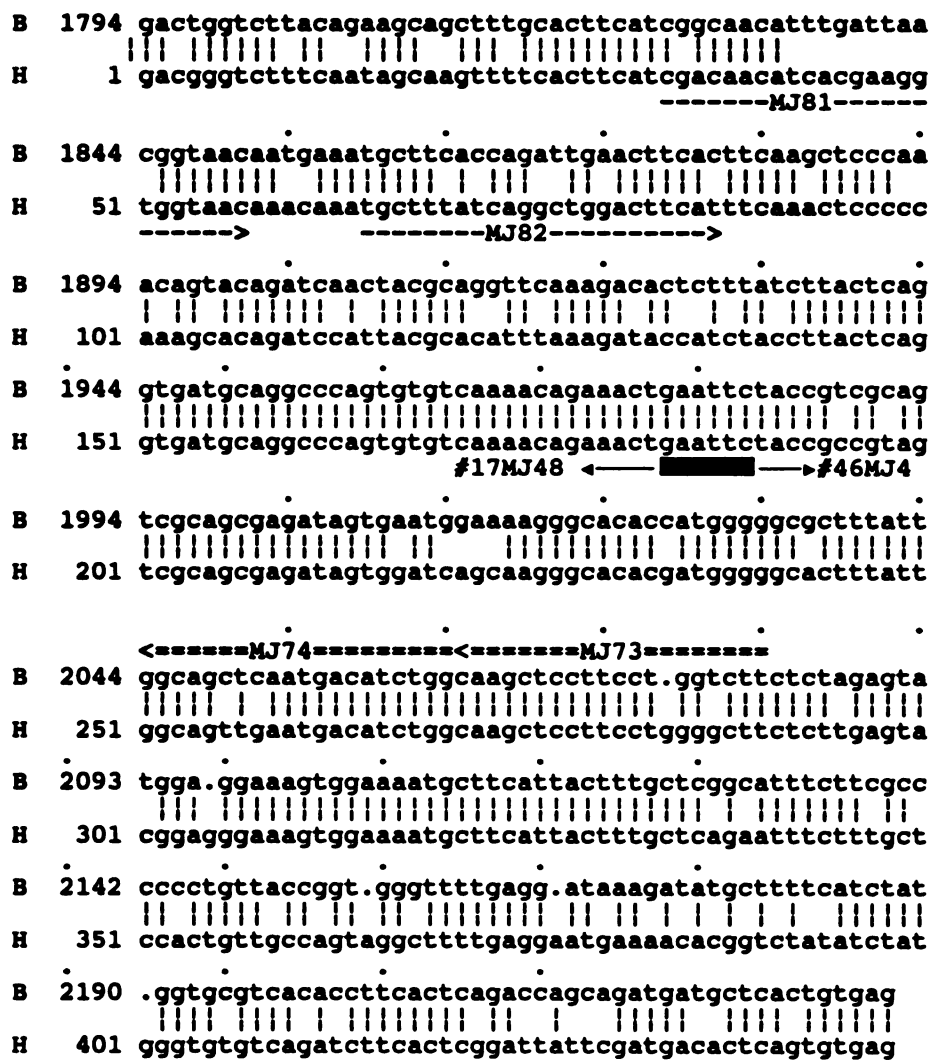


Fig. 3.7 DNA sequence comparison between human expressed sequence tag (EST01397) and bovine β-mannosidase cDNA. B represents the bovine sequence and H represents the human sequence. There is 81.49% identity between the two sequences. The internal *EcoRI* site is marked with ▬▬. Sequence upstream of the *EcoRI* site is from the 3' end of clone #17MJ48, and downstream is from the 5' end of clone #46MJ4. The two sense human primers (MJ81 and MJ82) and the two antisense bovine primers (MJ73 and MJ74) are marked with single dash lines and double dash lines, respectively.

oligonucleotide primers MJ81 and MJ82 were designed based on the 5′ region of the human sequence in order to generate a suitable probe to rescreen the library. Bovine thyroid total RNA was reverse transcribed into first strand cDNA using an antisense primer (MJ73) located downstream of the *EcoR*I site in clone #46MJ4. Amplification was done by priming the first strand cDNA with antisense primer MJ73 and sense primer MJ81. A specific product of approximately 215 bp was generated by reamplification of the PCR products with two nested primers (MJ82 and MJ74). The above 215 bp of PCR product was gel purified and directly sequenced from both directions. The sequence flanking the *EcoR*I site had high homology to the human taq sequence (EST01397) at both DNA and amino acid levels (Fig. 3.7).

In the meantime, the discovery of the failure of *EcoR*I methylation led us to reevaluate the other two clones, which had previously been identified by a mixed oligonucleotide probe (MJ48). The peptide sequence corresponding to probe MJ48 was not found in the clones #46MJ4 or r8-1, so it was possible that the clones identified by MJ48 corresponded to sequences upstream of the cleaved *EcoR*I site. To quickly determine the authenticity of these clones, PCR reactions were carried out on crude phage lysates from clones #9MJ48 and #17MJ48 using vector primers (M13 forward or reverse primer) and various oligonucleotide primers (e.g. MJ9, MJ48, MJ63, MJ82, and MJ64) with various combinations (Table 3.2

and 3.3). Specific PCR products were produced from clone #17MJ48. This clone was subjected to sequencing analysis. The sequence data showed that clone #17MJ48 contained an insert of 1119 bp and encoded 373 amino acids. Four peptide sequences (169r64, 169r65, 169r61, and 142r12) were found to exactly match with predicted amino acid sequences. The 3′ region showed high sequence homology with the human EST01397 sequence (Fig 3.7). The 215 PCR product (amplified by the human and bovine primers described above) spanned the 3′ region of clone 17MJ48 and the 5′ region of clone 46MJ4. Besides the expected 3′ cloning *EcoRI* site, an internal *EcoRI* site was found in this clone. The *EcoRI* site at the 5′ cloning site was defective by missing a C nucleotide.

The composite cDNA of clone 17MJ48 and r8-1 was approximately 3 kb long, however, about 1.2 kb of the 5′ region was still missed as indicated by the Northern analysis (see the following section). In order to isolate the missing 5′ end of the β-mannosidase gene, 5′ RACE was adopted (Fig. 3.8). Two synthetic oligonucleotide primers (MJ100 and MJ101) based on the 5′ region of clone #17MJ48 were prepared and used in the 5′ RACE system. The DNA template was first strand cDNA transcribed from bovine thyroid poly(A)$^+$ RNA by MJ100 oligonucleotide and tailed with homopoly-C. Under a smear background, a discrete band at approximately 950 bp was observed in a PCR reaction that was carried out by a gene specific oligonucleotide primer

(MJ101) and an anchor primer (Fig. 3.9). The 950 bp product was gel purified and reamplified by a universal amplification primer (UAP) and a nested oligonucleotide primer (MJ101) and by UAP and a nested degenerate primer (MJ48). A PCR product of approximately 950 bp was produced using the first pair of oligonucleotide primers. Two PCR products with a major product being 770 bp, were observed using the second pair of oligonucleotide primers (Fig. 3.9). The reamplified 950 bp product and 770 bp product (Fig. 3.9) were gel purified and subcloned into PCR™ vectors. To identify positive subclones, putative subclones were subjected to PCR using a nested gene-specific oligonucleotide (MJ110) and M13 reverse or forward primer. The authenticity of the 5′ RACE product was confirmed by both direct sequencing of the PCR products and sequencing positive subclones. A peptide sequence (104r86) was found to be colinear with the deduced amino acid sequence of the 5′ RACE products. In addition, sequencing of the 5′ RACE products also revealed a possible translation initiation codon at nucleotide 75 followed by an open reading frame. The nucleotides flanking the ATG (ACCATGC) were in good agreement with the consensus sequence for the eukaryotic initiator codon: A/GCCATGG (Kozak, 1986). Furthermore, the 17 amino acid residues following the initiator codon exhibited features characteristic of a signal sequence, i.e. a basic N-terminal region (M-L-L-R), a central hydrophobic

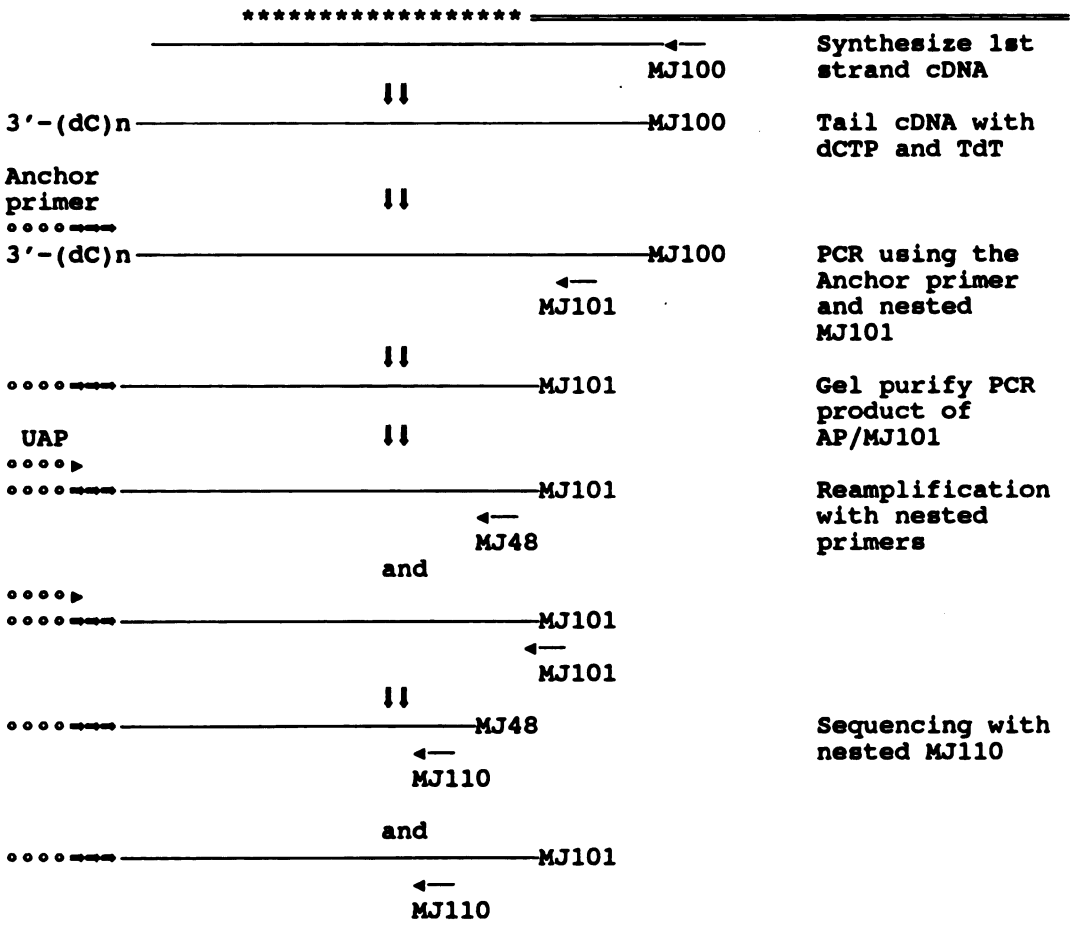


Fig. 3.8 The strategy of the 5' RACE for the isolation of the 5' region of β-mannosidase cDNA. The asterisks represents the 5' region of β-mannosidase. The double line represents cDNA of clone #17MJ48. The open circles represent universal sequence primer (UAP): 5' CUACUACUACUAGGCCACGCGTCGACTAGTAC 3'. The open circles plus the double line arrow represent anchor primer (AP): 5' CUACUACUACUAGGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG 3'.

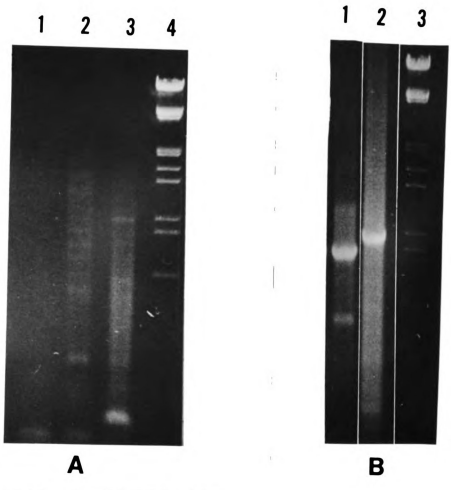


Fig. 3.9 Analysis of the 5′ RACE products. Panel A: cDNA reverse transcribed from bovine thyroid 0.5 μg poly $(A)^+$ RNA was amplified by a gene specific primer MJ101 and an anchor primer. After 35 cycles (94°C 1 min, 57°C 30 seconds, and 72°C 2 min for each cycle) of amplification, twenty microliters of 50 μl reactions were loaded into 1% agarose gel. Lane 1: negative control without cDNA. Lane 2: PCR of non-tailed cDNA. Lane 3: PCR of poly C tailed cDNA. Lane 4: 1 μl of marker III (lambda DNA/*EcoRI*/*Hind*III). Panel B: the major product of 940 bp in lane 3 at panel A was excised and suspended in 30 μl water. Aliquots of the suspension were reamplified with nested primers in a total of 200 μl reaction. DNA was reprecipitated, fractionated on 1% Nusieve GTG agarose gel in order to purify the products for subcloning. Lane 1: PCR with MJ48 and UAP. Lane 2: PCR with MJ101 and UAP. Lane 3 DNA marker III.

region with (L-L-L-L-L-A), and a more polar C-terminal region.

3.2.3.3 Northern analysis

A single transcript of approximately 4.2 kb was observed in both normal and affected tissues as well as in both caprine and bovine tissues (Fig. 3.10). The amount of transcript in affected tissues was significantly decreased compared to their normal counterparts. The tissue distribution of β-mannosidase transcript seemed to be consistent with the distribution of enzyme activity of β-mannosidase (Lovell *et al.*, 1994), i.e, Thyroid > kidney, liver > spleen, brain.

3.2.3.4 Southern genomic blot analysis

A PCR product of 544 bp generated from clone #46MJ4 using MJ66 and MJ5 (Table 3.4) was labeled and hybridized with genomic DNAs from different species including human, cattle, goat, mouse, rat, and Chinese hamster (Fig. 3.11). Under high stringency wash (Fig. 3.11 Panel A), i.e. in 0.2 × SSC/0.1% SDS at 65°C, *EcoR*I digested DNA revealed a single band of 7 kb in goat, a band of 3.2 kb in human, and two bands of 5 kb and 7.5 kb in bovine DNA. With *Pst*I digestion, there were two bands of 3.7 and 5.5 kb in goat, one band of 1.7 kb in human and five bands of 1.1, 1.8. 2.7, 3.1, and 8 kb in bovine samples. With *Xba*I digestion, two

bands of 2.8 and 3.4 kb in goat, one band of 1.9 kb in human and three bands of 3, 3.4, and over 10 kb in bovine samples were observed. No bands were observed in DNA from rat, mouse, and Chinese hamster. However, when the filter was washed at 45°C in a solution containing 1 × SSC/0.1% SDS, a smear background in human DNA was found and a new band was appeared in caprine and bovine DNA (Fig. 3.11 Panel B). Under the low stringency wash, some faint bands could be seen in lanes containing DNAs of rat, mouse, and hamster. Southern hybridization of *Pst*I cleaved DNAs from a panel of 25 human/rodent hybrids showed a 1.7 kb-band in the hybrid NA10115 (Fig. 3.12, lane 4). Its human origin was demonstrated by the observation of a band with the same size as in the control human DNA. Ninety-seven percent of cells from the hybrid NA10115 contain chromosome 4. Two bands of larger size were also found in several other hybrids (Fig. 3.12). The human control showed a smear background too. Similar results were produced with different bovine β-mannosidase cDNA probes (data not shown).

3.2.4 Discussion

In previous studies (Sopher, 1992a), the bovine β-mannosidase peptide of 100 kDa was deglycosylated and gel purified by SDS-PAGE. The peptide was subjected to CNBr digestion and sequencing. Three peptide sequences (peptide 1, 2, and 3) were obtained (Sopher, 1992a). Only peptide 1

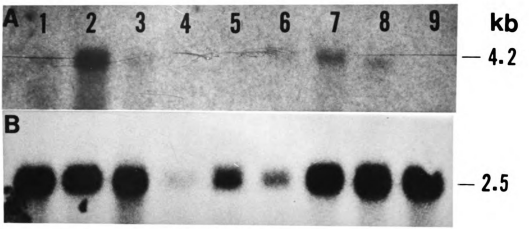


Fig. 3.10 Northern hybridization analysis of normal tissues and affected animals. Panel A: Poly A^+ RNA samples isolated from various bovine and caprine tissues were hybridized to a cDNA probe generated by PCR of clone #46MJ4 using primers MJ66 and MJ5 and a *EcoRI* fragment of clone #17MJ48. The blot was hybridized for two days and washed finally in 0.5 × SSC/0.1% SDS at 42°C for 30 min. The film was exposed for 10 days at -80°C. Lane 1 to 9: affected bovine thyroid ; normal bovine thyroid; affected bovine kidney; normal bovine kidney; normal goat brain; normal goat spleen; normal goat liver; normal goat kidney; affected goat kidney. Panel B: The blot was rehybridized to a rat GAPDH cDNA probe after removal of β-mannosidase probe. The film was exposed for 1 day at -80°C

Fig. 3.11 Southern hybridization of genomic DNA from various species. Panel A: Genomic DNA samples from hamster (lane 1, 7, and 13), mouse (lane 2, 8, and 14), rat (lane 3, 9, and 15), goat (lane 4, 10, and 16), cattle (lane 5, 11, and 17), and human (lane 6, 12, and 18) were digested with *EcoR*I (Lane 1-6), *Pst*I (lane 7-12), and *Xba*I (lane 13-18). The blot was hybridized with a cDNA probe generated by PCR of #46MJ4 using MJ66 and MJ5 and washed at high stringency condition (0.2 × SSC/0.1% SDS at 65°C for 1 hour) and exposed for 11 days at -80°C. Panel B: The same blot was stripped, rehybridized, and washed at low stringency condition (1 × SSC/0.1% SDS at 42°C for 1 hour) and exposed for 10 days at -80°C.

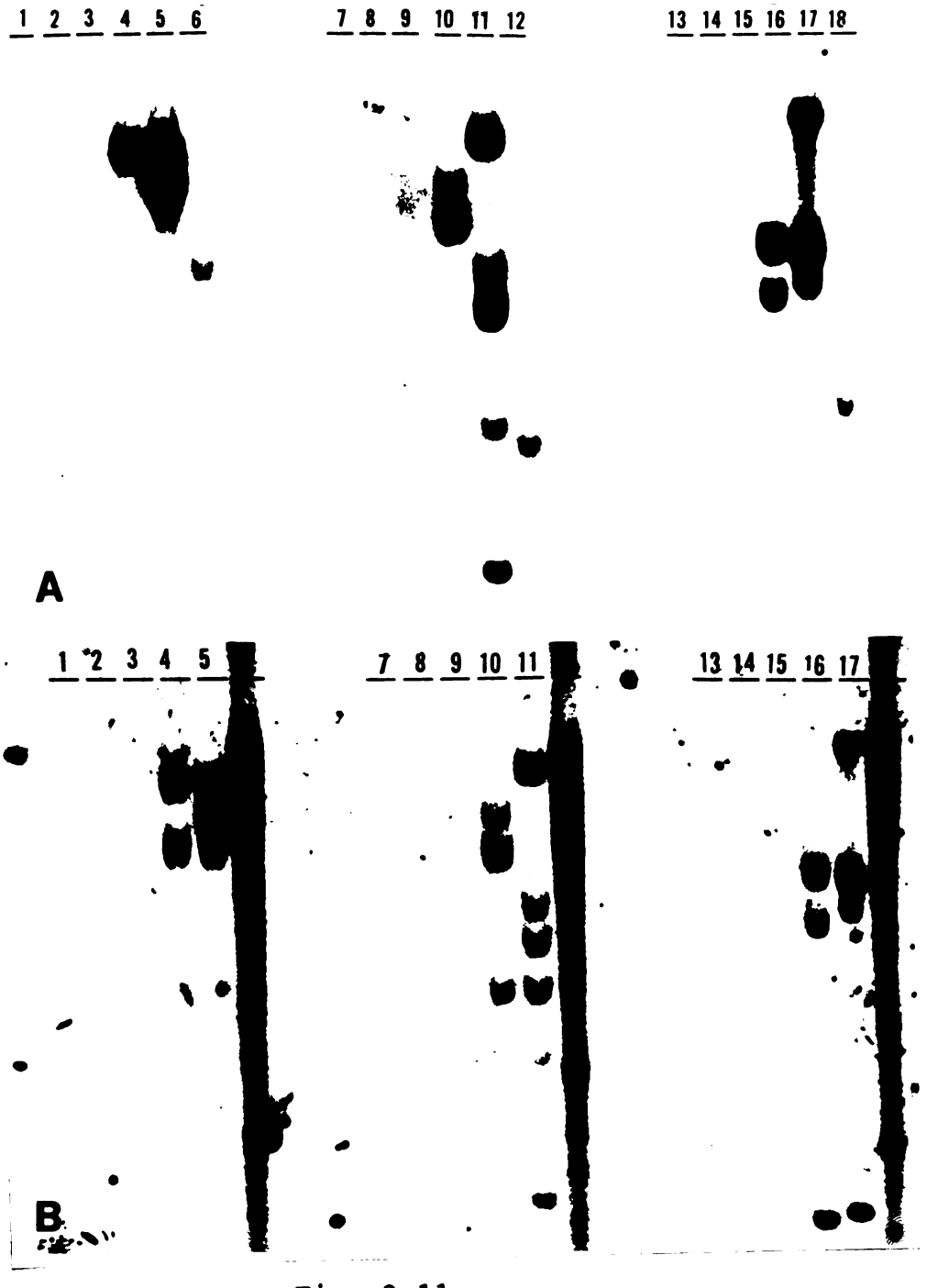

1 2 3 4 5 6
7 8 9 10 11 12
13 14 15 16 17 18
A
1 2 3 4 5
7 8 9 10 11
13 14 15 16 17
B

Fig. 3.11

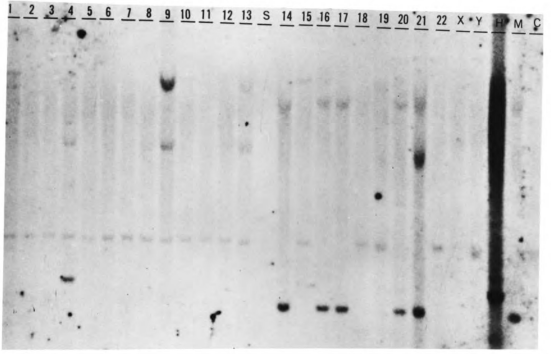


Fig. 3.12 Chromosome localization of β-mannosidase cDNA. Approximately 15 μg of *Pst*I digested genomic DNA from 24 human/rodent somatic cell hybrids were hybridized with a cDNA probe generated by PCR of plasmid DNA of clone #46MJ4 using primers MJ66 and MJ5. The hybridized blot was washed in 1 × SSC/0.1% SDS for 1 hour at 42°C and exposed at -80°C for 10 days. The number on top represents a human chromosome retained in that somatic cell line. H: human control DNA. M: mouse control DNA. C: Chinese hamster control DNA. S: DNA molecular marker III.

yielded a complete sequence with six amino acid residues and the other two produced either mixed assignments or incomplete sequence. The sequence of peptide 1 was too degenerate to design a mixed oligonucleotide probe suitable for screening. More peptide sequences were therefore required. The protein used for peptide sequencing was highly purified and contained three bands (84, 100, and 110 kDa). All three peptides were β-mannosidase protein (Sopher *et al.*, 1993) with the 100 kDa as a major peptide. The 84 kDa was thought to be a proteolytic product from the larger peptides and little or no 80 kDa protein was present in this preparation as judged by Coomassie blue staining and Western analysis. Previous studies by Sopher *et al.* (Sopher, 1992a) indicated that the N-terminal of β-mannosidase protein was blocked. In addition, trypsin cleaved peptides appeared to be relatively insoluble. Therefore, internal peptides created by combined CNBr and trypsin digestion were sequenced. The sequence of peptide 142r12 was found to match the sequence of peptide 1 but with two additional residues at its C-terminus. Peptide 151r72 appeared to match the mixed sequence of peptide 2 with differences in the last two residues. The incomplete sequence from peptide 3 appeared to overlap peptides 103 and 218r24. Searches for all the β-mannosidase peptide sequences obtained revealed no significant homologies with existing proteins in the GenBank database.

These results indicated that these peptide sequences were derived from the bovine β-mannosidase protein. Knowing the β-mannosidase peptide sequences made it possible to clone the β-mannosidase cDNA. Six independent but incomplete cDNA clones were identified and confirmed by screening a total of 5×10^5 plaque-forming units from the bovine thyroid cDNA library. Their relationship is shown in Fig. 3.4. The 3′ non-coding region of clone r8-1 may represent the true sequence for β-mannosidase cDNA since its sequence was found in clones 46MJ4 (only missing the final 20 bp) and r2-1a (except for the final 106 bp). The final parts of 3′ non-coding regions of clone #47MJ4 (approximately 450 bp) and r2-1a (approximately 1 kb) were not found in other clones and were most likely cloning artifacts. Analysis of genomic DNA, or Northern analysis, or genomic cloning will solve this problem. These clones comprised approximately 70% sequence of the full-length cDNA (4.2 kb). With the addition of the 5′ RACE system, cDNA sequence information close to a full-length of β-mannosidase was obtained.

Three pieces of evidence support the authenticity of the cDNA for β-mannosidase. First, the deduced amino acid sequence from the nucleotide sequence of this cDNA is colinear with that of all ten β-mannosidase peptides (a total of 105 residues) determined by direct amino acid sequencing. Second, this cDNA is located on human

chromosome 4, which is in agreement with that of previous report (Fisher *et al.*, 1987). Third, the RNA transcript of this cDNA in affected β-mannosidosis animals was much lower than in normal animals.

The cDNA contains 3852 base pairs. There is a 74 bp 5′ non-coding region, followed by a 2637-bp coding region encoding 879 amino acids, then a 1141-bp of 3′ non-coding region and finally a 13-bp poly (A) tail. The first in-frame initiation codon is followed by 17 amino acids containing the characteristic features of a signal peptide sequence, i.e. a positively charged amino acid (lysine) within the first 5 amino acids, a hydrophobic core (L-L-L-L-L-A), and a more polar C-terminal region.

Since the β-mannosidase protein was N-terminal blocked as suggested by previous studies (Sopher, 1992a), the precise cleavage site of the peptide sequence is unknown. Based on the (-3, -1) rule (von Heijne, 1986; 1990) we predict that the signal peptide is cleaved between residues 17 (A) and 18 (T).

Besides the signal peptide sequence, several hydrophobic regions (e.g. residues 96 to 114 and 406 to 422) are predicted according to Kyte and Doolitte (1982) (Fig. 3.6), however, none of them is likely to be a membrane spanning peptide. The hydropathy profile of the β-mannosidase polypeptide shows dispersed hydrophobic or hydrophilic regions except around region 490 to 640 which is

mainly hydrophilic.

Two possible poly (A) signal sequences (AATATA and ATTATA) were found at 15 bp and 32 bp before the poly (A) tail. None of them is a consensus sequence. Various non-consensus poly (A) signal sequences have been reported in several lysosomal enzymes (Stoltzfus *et al.*, 1992; Pohlman *et al.*, 1988; Stein *et al.*, 1989; Proia *et al.*, 1986).

The deduced peptide sequence from the cDNA matches with all peptide sequences including those containing incomplete or mixed sequences. Although there are four discrepancies between the microsequenced amino acid sequences from CNBr/tryptic peptides and those predicted from the cDNA, two are at residues with uncertainty and thus are possibly due to peptide sequence artifacts. Other two may reflect natural polymorphisms. Previous studies (Sopher *et al.*, 1993) demonstrated that the size of bovine β-mannosidase was decreased to 86 and 91 kDa from 100 and 110 kDa, respectively, after deglycosylation with N-glycosidase F. This suggested that seven to nine complex type oligosaccharides may be present. There are six potential glycosylation sites in the cDNA (Fig. 3.5). We predict therefore that all six glycosylation sites may be occupied. The overestimation of glycosylation sites by the deglycosylation study might be due to inaccurate estimation by SDS-PAGE since the presence of carbohydrate chains was found to shift the protein migration (Mahuran *et al.*, 1988).

The 2586 bp coding region (after removal of 17 amino acid residues of the signal peptide) encodes 862 amino acids. This would give a predicted molecular mass of approximately 103 kDa. Therefore, an additional 12 to 17 kDa peptide is presumably cleaved from the β-mannosidase precursor protein as well as the signal peptide sequences. The majority of lysosomal enzymes have been demonstrated or predicted to be involved in proteolytic processing including trimming the N-terminal or C-terminal sequences, or cleaving internal peptides (Gottschalk *et al.*, 1989; Mahuran *et al.*, 1988; Erickson and Blobel, 1983; Yamamoto *et al.*, 1990; Quinn *et al.*, 1987; Hoefsloot *et al.*, 1988). So far, arylsulfatase A is the only lysosomal enzyme whose maturation is restricted to the cleavage of its signal peptide sequence (Stein *et al.*, 1989). A peptide (171) sequence was found to be located next to the stop codon in bovine β-mannosidase cDNA. Therefore, it is likely that the proteolytic processing of bovine β-mannosidase does not involve the C-terminal cleavage. The expression of the β-mannosidase cDNA should determine the biosynthesis and processing of β-mannosidase and the relationship between the three β-mannosidase peptides.

A single transcript species of approximately 4.2 kb was revealed in both normal and β-mannosidosis animals and in both bovine and caprine tissues. There is a slight size difference between the constructed cDNA (3845 bp + 200 bp

poly (A) = 4045) of β-mannosidase and the RNA transcript (4.2 kb), which probably reflects some missing 5′ non-coding region. There was no size difference between the normal mRNA and β-mannosidosis mRNA. However, the amount of messenger RNA was much lower in affected animals compared to normal after standardizing the RNA loading with GAPDH. The observation of normal size generated from β-mannosidosis animals implies that the β-mannosidosis in ruminants is most likely to be caused by point mutations or small deletions or insertions (1-10 bp). The low yield of transcript observed in affected ruminants indicated that there might be a point mutation producing a premature stop codon (Mahuran, 1991; Zhang, 1994, and Cheng and Maquat, 1993) or mutations in the promoter region affecting the transcription initiation. Southern analysis with several restriction enzyme digestions revealed no gross gene rearrangements in affected and carrier β-mannosidosis animals (data not shown). The same sized RNA transcript was demonstrated in both caprine and bovine tissues. These results were consistent with a previous study which demonstrated that cattle and goats had β-mannosidase peptides of the same size after deglycosylation (Sopher *et al.*, 1992; 1993). This laboratory has already documented that goats and cattle affected with β-mannosidosis have very similar phenotypes including clinical manifestations, pathological defects, physiological dysfunction, and biochemical perturbation

(Jones and Abbitt, 1993; Jones *et al.*, 1992; Lovell *et al.*, 1991; Patterson *et al.*, 1991).

Hybridization of restriction enzyme digested genomic DNA with a cDNA fragment of bovine β-mannosidase revealed 2-6 bands of approximately 1 to 10 kb in the bovine species, two bands of approximately 2.8 to 7 kb in the caprine species, and a single band in humans. Under the conditions of a low stringency wash, a new band in caprine DNA and bovine DNA and a smear background in human appeared. A similar pattern was observed upon hybridization of genomic DNA from a panel containing 24 human/rodent cell hybrids with the same bovine cDNA probe and with a different β-mannosidase cDNA probe. Similar complex results of genomic Southern analysis were observed in studies of the β-subunit of human β-hexosaminidase (O'Dowd *et al.*, 1985). These results may reflect the presence of closely related gene sequences, e.g. gene families or pseudogenes. However, this speculation remains to be proven by genomic cloning studies. The reason for this smear is unclear now. A 1.75 kb fragment was located only on human chromosome 4. The 1.75 kb was the only band presented under the condition of a high stringency wash. Our result supported the previous chromosome mapping (Fisher *et al.*, 1987; Lundin, 1987). The two larger bands of approximately 4 and 20 kb that appeared under low stringency washing were seen in several cell hybrids. The incomplete restriction enzyme digestion might

account for this result.

No significant sequence homologies between the β-mannosidase and other lysosomal enzymes were found by searching against GenBank. However, a striking homology between a human expressed sequence tag (EST01397) to unknown gene (Adams *et al.*, 1992) was observed. There was 80.6% identity in a 454 bp overlap. High homology at amino acid level was observed mostly in the central region of the human cDNA. However, the open reading frame in the human sequence is shifted by adding a nucleotide at 284 bp and 305 bp. We believe that the human tag sequence actually represents a partial cDNA sequence of human β-mannosidase. The reading frame shift in the human expressed sequence tag is most likely due to sequence errors since it occurs in regions containing G stretches.

β-Mannosidase had not previously been cloned from any species. The availability of the cDNA encoding bovine β-mannosidase enables us to isolate β-mannosidase cDNA from other species including human and goat. The expression of the cDNA allows studies of β-mannosidase processing and transport, as well as possible association with other proteins in the lysosome. The availability of the β-mannosidase cDNA should also permit us to characterize the gene structure and gene regulation. In addition, the cloning of β-mannosidase cDNA will facilitate the identification of molecular lesions underlying β-

mannosidosis in humans, goats, and cattle. Finally, it opens the door to gene therapy of several early onset neurodegenerative disorders.

3.2.5 Summary

Approximately three dozen positive clones were identified by screening a bovine thyroid cDNA library with several degenerate oligonucleotide probes. Among them, two independent clones #46MJ4 and #47MJ4 (#43MJ4) showed cross-hybridization with other oligonucleotide probes. Rescreening the library with the whole insert of clone #47MJ4 generated three additional clones r2-1a, r8-1, and r20-2. Except for clone r2-1a, the other four clones started at the same 5′ end with a cleaved *EcoRI* site. Sequencing and PCR analysis demonstrated that there was sequence homology between these clones. Clones #46MJ4, #47MJ4, r20-2, and r2-1a have an incomplete 3′ region. However, clone #47MJ4 contained approximately 450 bp of different sequence at its 3′ end, while clone r2-1a possesses additional sequence in both ends which appeared to be gene cloning artifacts. Besides clones #46MJ4 and #47MJ4, clone #17MJ48 was also identified from the original screening but showed no cross-hybridization with other oligonucleotide probes. This clone was proved later to be β-mannosidase cDNA and was located adjacent to the 5′ ends of other clones. The remaining 5′ region of the bovine β-

mannosidase cDNA was cloned by using 5′ RACE technique. The whole cDNA contains 3852 base pairs, which includes a 74-bp 5′ non-coding region, a 2637-bp coding region encoding 879 amino acids, a 1141-bp 3′ non-coding region and a 13-bp poly (A) tail. A 17-residue sequence after the predicted initiation codon contains the characteristics of a signal peptide sequence. Six possible glycosylation sites were predicted. All ten peptide sequences determined by amino acid sequencing were found in the predicted amino acid sequence from the bovine cDNA, comprising a total of 105 amino acids. A few mismatches were present. A striking sequence similarity was observed between the β-mannosidase cDNA and a human expression tag sequence to an unknown gene.

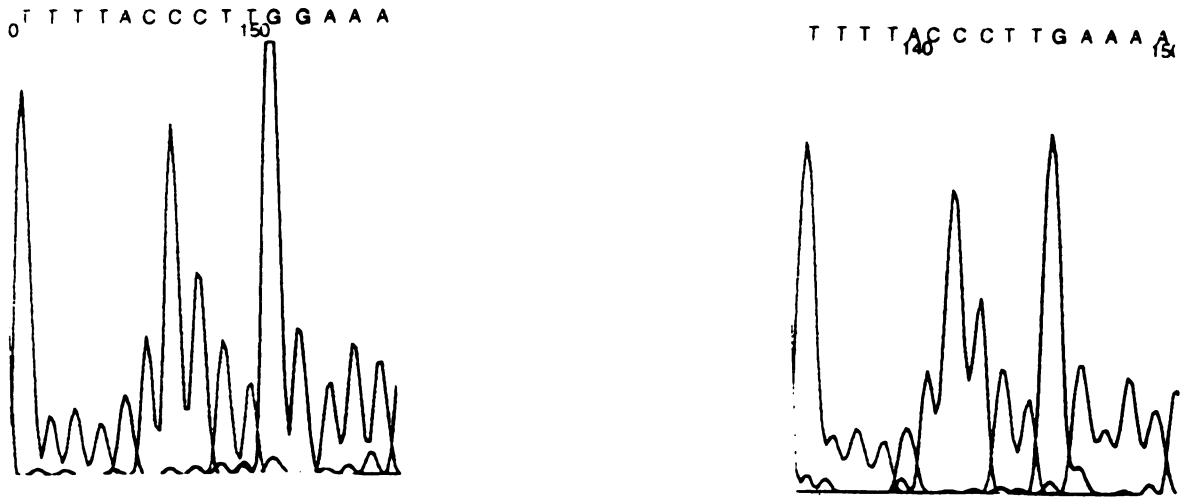
Northern blot analysis demonstrated a single transcript of 4.2 kb in bovine and caprine tissues. The size of the transcript in the affected animal did not change, but the yield was reduced. Southern blot analysis indicated that there were no gross gene rearrangements in animals with β-mannosidosis. The β-mannosidase gene was mapped on human chromosome 4 by Southern hybridization of DNAs from 24 rodent/human somatic cell hybrids.

3.4 BEYOND THE CLONING OF β-MANNOSIDASE CDNA

The availability of β-mannosidase cDNA enabled immediate initiation of several studies on which there has already been significant progress made in this laboratory: (1) mutation analysis; (2) fluorescence *in situ* hybridization (FISH) analysis; (3) Cloning and sequencing analysis of the normal caprine β-mannosidase cDNA; and (4) isolation of human β-mannosidase cDNA (a separate project undertaken by a colleague). For mutation analysis, cDNA was prepared by reverse transcription of affected bovine or caprine total RNA. PCR was performed using internal primers designed from the cloned bovine β-mannosidase cDNA. Overlapping PCR products were generated to cover the whole encoding region and directly sequenced. Eleven overlapping PCR fragments covering the whole encoding region and the partial 3′ non-coding region of a bovine affected RNA were analyzed. No mutation, except for a single bp change at 2573 (g/a), was found. This g/a change introduced a stop codon 22 residues before the normal stop codon (Fig. 3.13). Preliminary PCR analysis of genomic DNAs indicated that affected β-mannosidase calves were homozygous for this nonsense mutation and β-mannosidase carriers were heterozygous, which was consistent with the autosomal recessive inheritance pattern for β-mannosidosis. Samples from more Salers cattle with β-mannosidosis will be

analyzed. The identification of this mutation in samples from β-mannosidosis calves further supports the conclusion that the isolated cDNA encodes β-mannosidase. Mutation analysis of affected goats is still underway. Approximately 2 kb cDNAs from both normal and affected goats were already amplified and sequenced. A missense mutation (A/D) was observed (Fig. 3.14), which introduced an *Alf1* III restriction site at this position. PCR analysis of genomic DNA will be performed to verify this A/D change. The partial normal cDNA for caprine β-mannosidase showed a stop codon at the same position as bovine cDNA and 95% sequence homologies with the bovine counterpart at both DNA and amino acid levels, reflecting the close evolutionary relationship between the two species.

The FISH technique was developed more than a decade ago (Rudkin and Stollar, 1977). Its applications include prenatal diagnosis, tumor biology, and gene mapping (Poggensee and Lucas, 1992). A FISH analysis system was used to verify the β-mannosidase gene mapping determined by Southern analysis and to generate a regional assignment. The preliminary studies using a random primed labeled β-mannosidase cDNA probe produced a high background. A probe labeled by the nick translation method would be more s98uitable due to its small size. Hybridization and washing 7conditions should be analyzed further. Human β-mannosidase clones would be better probes for use in FISH analysis



n.t ttttacccttggaaa
a.a F Y P W K

Normal sequence

n.t ttttacccttgaaaa
a.a F Y P * ▲

Mutant sequence

Fig. 3.13 A nonsense mutation in β-mannosidosis calves. PCR products generated from normal and affected animals were directly sequenced by the dye terminator method. n.t, nucleotide; a.a amino acid; *, stop codon. The single base change is marked by an arrow.

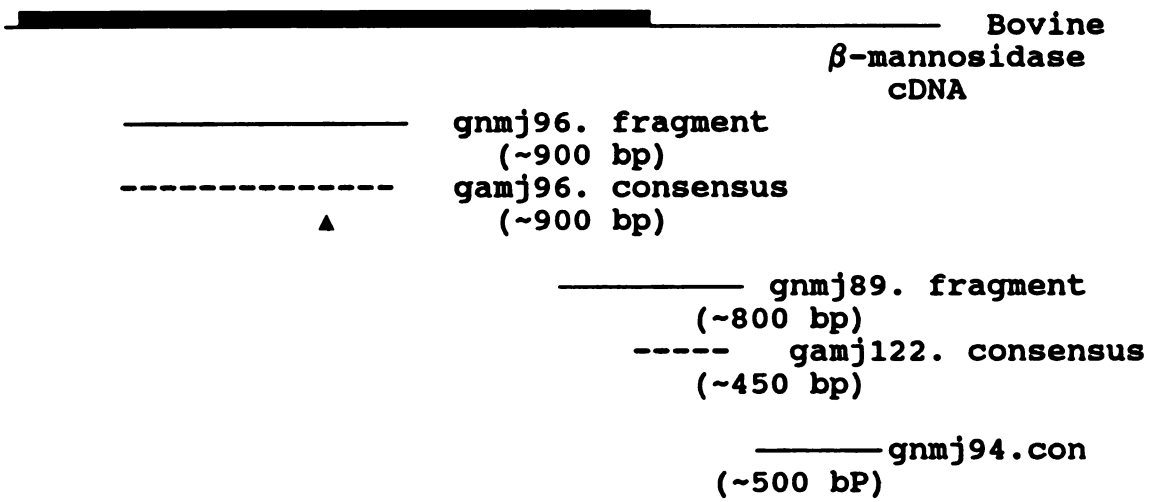


Fig. 3. 14 Summary of PCR analyses of cDNAs from normal and affected β-mannosidosis goats. Dash lines represent sequences from an affected goat. Lines represent normal caprine sequences obtained. Solid bar represents the coding region of bovine β-mannosidase cDNA. All caprine sequences were obtained by PCR amplifications of normal or affected caprine cDNAs using bovine primers derived from the cloned bovine β-mannosidase cDNA. A possible missense mutation (A/D change) is marked with ▲.

because high stringency conditions can be used.

There is a significant contrast in the clinical manifestations and biochemical perturbations between human and ruminant β-mannosidoses. Isolation and characterization of human β-mannosidase cDNA will help define the genetic basis of these differences. The successful isolation and characterization of bovine β-mannosidase cDNA enabled us to pursue cloning of the human counterpart. Although screening a human cDNA library using bovine β-mannosidase cDNA probes did not produce positive clones, an alternative screening approach using a PCR product generated by a bovine β-mannosidase internal primer and a human primer from the human tag sequence has identified several putative clones. Further studies are underway.

Besides the three studies described above, isolation of a genomic clone encoding the β-mannosidase gene and expression of the bovine cDNA in *COS* cells can be done because of the successful cloning and sequencing of β-mannosidase cDNA. However, a full-length construct has to be made before initiating *in vitro* expression. Cloning of a genomic DNA for β-mannosidase will provide the information regarding the structural organization, regulation, and evolution of the β-mannosidase gene. The *in vitro* expression study will determine whether this bovine cDNA does encode a functional gene, and more importantly it will help us to understand the biosynthesis and processing

procedures of the β-mannosidase protein and how the protein is transported. All these studies will facilitate the establishment of an animal model for gene therapy studies and make the final goal of gene therapy more reality.

BIBLIOGRAPHY

REFERENCES

Abbitt, B., Jones, M.Z., Kasari, T.R., Storts, R.W., Templeton, J.W., Holland, P.S., and Castenson, P.E. (1991) β-Mannosidosis in twelve Salers calves. *J. Am. Vet. Med. Assoc.* **198**, 109-113.

Adams, M.D., Dubnick, M., Kerlavage, A.R., Moreno, R., Kelley, J.M., Utterback, T.R., Nagle, J.W., Fields, C., and Venter, J.C. (1992) Sequence identification of 2,375 human brain genes. *Nature* **335**, 632-634.

Aronson, N.N. and Kuranda, M.J. (1989) Lysosomal degradation of Asn-linked glycoproteins. *FASEB J.* **3**, 2615-2622.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Smith, J.A., Seidman, J.G. and Struhl, K. (eds) (1987) *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley-Interscience, New York.

Bapat, B., Ethier, M., Neote, K., Mahuran, D., and Gravel, R.A. (1988) Cloning and sequence analysis of a cDNA encoding the β-subunit of mouse β-hexosaminidase. *FEBS Lett.* **237**, 191-195.

Bernard, M., Sioud, M., Percheron, F., and Foglietti, M.J. (1986) β-Mannosidase in human serum and urine: a comparative study. *Intl. J. Biochem.* **18**, 1065-1068.

Bernstein, P., and Ross, J. (1989) Poly (A), poly (A) binding protein and the regulation of mRNA stability. *TIBS.* **14**, 373-377.

Bischoff, J., Moremen, K., and Lodish, H.F. (1990) Isolation, characterization, and expression of cDNA encoding a rat liver endoplasmic reticulum α-mannosidase. *J. Biol. Chem.* **265**, 17110-17117.

Bishop, D.F., Calhoun, D.H., Bernstein, H.S., Hanizopoulos, P., Quinn, M., and Desnick, R.J. (1986) Human α-galactosidase A: nucleotide sequence of a cDNA clone encoding the mature enzyme. *Proc. Natl. Acad. Sci. USA* **83**, 4859-4863.

Bishop, D.F., Kornreich, R., and Desnick, R.J. (1988)

Structure 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.

Boyer, P.J. and Lovell, K.(1990) Investigation of dysmyelinogenesis in caprine β-mannosidosis: in vitro characterization of oligodendrocytes. *Glia* **3**, 222-227.

Boyer, P.J., Jones, M.Z., Nachreiner, R.F., Refsal, R.F., Common, K.R., Kelley, R.S., and Lovell, K.L.(1990) Caprine β-mannosidosis; abnormal thyroid structure and function in a lysosomal storage disease. *Lab. Invest.* **63**, 100-106.

Bryan, L., Schmutz, S., Hodges, S.D., and Suyder, F.F.(1990) β-Mannosidase deficiency. *Biochem. Biophys. Res. Commun.* **173**, 491-496.

Burg, J., Conzelmann., Sandohoff, K., Solomon, E., and Swallow, D.M. (1985) Mapping of the gene coding for the human G_{m2} activator protein to chromosome 5. *Annu. Human. Genet.* **49**, 41-45.

Calhoun, D.H., Bishop, D.F., Bernstein, H.S., Quinn, M., Hantzopoulos, P., and Desnick, R.J. (1985) Fabry disease: isolation of a cDNA clone encoding human α-galactosidase A. *Proc. Natl. Acad. Sci. USA* **82**, 7364-7368.

Camirand, A., Heysen, A., Grondin, B., and Herscovics, A. (1991) Glycoprotein biosynthesis in Saccharomyces cerevisiae: isolation and characterization of the gene encoding a specific processing α-mannosidase. *J. Biol. Chem.* **266**, 15120-15127.

Casella, J.F., Casella, S.J., Hollands, J.A., Caldwell, J.E., and Cooper, J.A. (1989) Isolation and characterization of cDNA encoding the α subunit of Cap Z $_{(36/32)}$, an actin-capping protein from the Z line of skeletal muscle. *Proc. Natl. Acad. Sci. USA* **86**, 5800-5804.

Cavanagh, K., Dunstan, R.W., Jones, M.Z. (1982) Plasma α- and β-mannosidase activities in caprine β-mannosidosis. *Am. J. Vet. Res.* 43, 1058-1059.

Cavanagh, K.T., Fisher, R.A., Legler, G., Herrchen, M., Jones, M.Z., Julich, E., Sewell-Alger, R.P., Sinnott, M.L., and Wilkinson, F.E. (1985) Goat liver β-mannosidases: molecular properties, inhibition and inactivation of the lysosomal and nonlysosomal forms. *Enzyme*. **34**, 75-82.

Cavanagh, K. T., Jones, M.Z., Abbitt, B., and Skinner, R. (1992) Bovine plasma β-mannosidosis activity and its

potential use for β-mannosidosis carrier detection. *J. Vet. Diagn. Invest.* **4**, 434-440.

Chatterjee, S., and Gosh, N. (1989) Neutral sphingomyelinase from human urine. *J. Biol. Chem.* **264**, 12554-12561.

Cheng, J. and Maquat, L.E. (1993) Nonsense codons can reduce the abundance of nuclear mRNA without affecting the abundance of pre-mRNA or the half-life of cytoplasmic mRNA. *Mol. Cell. Biol.* **13**, 1892-1902.

Cooper, A., Sardharwalla, I.B., and Roberts, M.M. (1986) Human β-mannosidase deficiency. *N. Eng. J. Med.* **315**, 1231.

Cooper, A., Hatton, C., and Sardharwalla, I.B.(1987) Acid β-mannosidase: influence of age and sex on enzyme activity. *J. Inher. Metab. Dis.* **10**, 229-233.

Cooper, A., Wraith, J.E., Savage, w.J., Thornley, M., and Noronha, M.J. (1991) β-Mannosidase deficiency in a female infant with epileptic encephalopathy. *J. Inher. Metab. Dis.* **14**, 18-22.

Colin, B., Bernard, M., Foglietti, M.J., and Percheron, F. (1987) β-D-Mannosidase in human polymorphonuclear leukocytes and lymphocytes: a comparative study. *J. Biochem.* **19**, 395-397.

Dahl, D.L., Warren, C.D., Rathke, E.J.S. and Jones, M.Z. (1986) β-Mannosidosis: prenatal detection of caprine allantoic fluid oligosaccharides with thin layer gel permeation and high performance liquid chromatography. *J. Inher. Metab. Dis.* **9**, 1-5.

D'Amore, M.A., Gallagher, P.M., Korfhagen, T.R., and Ganschow, R.E. (1988) Complete sequence and organization of the Murine β-glucuronidase gene. *Biochem.* **27**, 7131-7140.

Dawson, G. (1982) Evidence for two distinct forms of mammalian β-mannosidase. *J. Biol. Chem.* **257**, 3369-3371.

de Wet, J.R., Fukushima, H., Dewji, N.N., Wilcox, E., O'Brien, J.S., and Helinski, D.R. (1984) Chromogenic immunodetection of human serum albumin and α-L-fucosidase clones in a human hepatoma cDNA expression library. *DNA* **3**, 437-447.

Dorland, L., Duran, M., Hoefnagels, F.E.T., Breg, J.N., Fabery de Jonge, H., van Eeghan-cransberg, K. van Spring, F.J., and van Diggelen, O.P. (1988) β-Mannosidosis in two brothers with hearing loss. *J. Inher. Metab. Dis.* **11**, 255-258.

Dunstan, R.W., Cavanagh, K., Jones, M.Z. (1983) Caprine α- and β-mannosidase activities; effects of age, sex, and reproductive status and potential use in heterozygote detection of β-mannosidosis. *Am. J. Vet. Res.* **44**, 685-689.

Dynan, W.S (1986) Promoters for housekeeping genes. *Trends Genet.* **2**, 196-197.

Erickson, A.H. and Blobel, G. (1983) Carboxyl-terminal proteolytic processing during biosynthesis of the lysosomal enzymes β-glucuronidase and cathepsin D. *Biochem.* **22**, 5201-5205.

Fisher, R.A., Povey, S., Cavanagh, K., Dupuis, M., Jones, M.Z. (1987) Studies to determine the chromosome assignment of human lysosomal β-mannosidase (MANBA). *Am. J. Hum. Genet.* **41**, A165.

Fisher, K.J and Aronson,Jr. N.N. (1989) Isolation and sequence analysis of a cDNA encoding rat liver α-L-fucosidase. *Biochem. J.* **264**, 695-701.

Frei, J.I., Cavanagh, K.T., Fisher, R.A., Hausinger, R.P., Dupuis, M., Rathke, E.J.S., and Jones, M.Z. (1988) Partial purification of goat kidney β-mannosidase. *Biochem. J.* **249**, 871-875.

Frohman, M.A., Dush, M.K., and Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002.

Frohman, M.A. Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: thermal RACE. *Methods Enzymol.* **218**, 340-356.

Fukushima, H., de Wet, J.R., and O'Brien, J.S. (1985) Molecular cloning of cDNA for human α-L-fucosidase. *Proc. Natl. Acad. Sci. USA* **82**, 1262-1265.

Geier, C., von Figura, K., and Pohlmann, R. (1989) Structure of the human lysosomal acid phosphatase gene. *Eur. J. Biochem.* **183**, 611-616.

Ginns, E.I., Choudary, P.Y., Martin, B.M., Winfield, S., Stubblefield, B., Mayor, J., Merkle-Lehman, D., Murry, G.J., Bowers, L.A., and Barranger, J.A. (1984) Isolation of cDNA clones from human β-glucocerebrosidase using the λgt11 expression system. *Biochem. Biophy. Res. Commun.* **123**, 574-580.

Gonzalez, D.H. and Chan, R.L. (1993) Screening cDNA

libraries by PCR using λ sequencing primers and degenerate oligonucleotides. *Trends Genet.* **9**, 231-232.

Gottschalk, S., Waheed, A., Schmidt, B., Laidler, P., and von Figura, K. (1989) Sequential processing of lysosomal acid phosphatase by a cytoplasmic thiol proteinase and a lysosomal aspartyl proteinase. *EMBO J.* **8**, 3215-3219.

Grabowski, G.A., Gatt, S., and Horowitz, M. (1990) Acid ß-glucosidase: enzymology and molecular biology of Gaucher disease. *Biochem. Mol. Biol.* **25**, 385-414.

Graham, T.R., Zassenhaus, P., and Kaplan. A. (1988) Molecular cloning of the cDNA which encodes β-N-acetylhexosamindase A from Dictyostelium discoideum. *J. Biol. Chem.* **263**, 16823-16829.

Hancock, L.W., Jones, M.Z., and Dawson, G. (1986) Glycoprotein metabolism in normal and β-mannosidase-deficient cultured goat skin fibroblasts. *Biochem. J.* **234**, 171-183.

Hancock, L.W. and Dawson, G. (1987) Evidence for two catabolic endoglycosidase activities in β-mannosidase-deficient goat fibroblasts. *Biochem. Biophy. Acta.* **928**, 13-21.

Harlow, E. and Lane, D. (eds) (1988) *Antibodies: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor Press, New York.

Hartley, W.J. and Blakemore, W.F. (1973) Neurovisceral storage and dysmyelinogenesis in neonatal goats. *Acta Neuropathol.* **25**, 325-333.

Healy, p.J., Slaman, J.T., Gardner, I.A., and Sewell, C.A. (1981) β-Mannosidase deficiency in Anglo Nubian goats. *Aust. Vet. J.* **57**, 504-507.

Helfman, D.M. and Hughes, S.H. (1987) Use of antibodies to screen cDNA expression libraries prepared in plasmid vectors. *Methods Enzymol* **152**, 451-457.

Hers, H.G. (1965) Inborn lysosomal disease. *Gastroenterol.* **48**, 625-633.

Hers, H.G. (1973) The concept of inborn lysosomal storage disease. In: *Lysosomes and Storage Diseases* (Hers, H.G. and van Hoof, F; eds) pp. 147-171, Academic Press, New york

Hieber, V.C. (1982) Cloning of a cDNA complementary to rat preputial gland β-glucuronidase mRNA. *Biochem. Biophy. Res.*

Commun. **104**, 1271-1278.

Hoefsloot, L.H., Hoogeveen-Westerveld, M., Kroos, M.A., Beeuman, J.V., Reuser, A.J.J., and Oostra, B.A. (1988) Primary structure and processing of lysosomal α-glucosidase: homology with the intestinal sucrase-isomaltose complex. *EMBO. J.* **7**, 1697-1704.

Hoefsloot, L.H., Hoogeveen-Westerveld, M., Reuser, A.J.J., and Oostra, B.A. (1990) Characterization of the human lysosomal α-glucosidase gene. *Biochem. J.* **272**, 493-497.

Hokke, C.M., Duran, M., Dorland, L., van Pelt, J., and van Sprang, F.J.(1990) Novel storage products in human β-mannosidosis. *J. Inher. Metab. Dis.* **13**, 273-276.

Holtzman, E. (ed) (1989) *Lysosomes* pp 363-416. Plenum Press, New York.

Horwich, A.L., Kraus, J.P., and Rosenberg, L.E. (1984) Strategies for the molecular cloning of low abundance messenger RNAs. in *Molecular basis of lysosomal storage disorders* (Barranger, J.A., and Brady, R.O. eds) pp.365-385, Academic Press, Inc. Orlando.

Horowitz, M., Wilder, S., Horowitz, Z., Reiner, O., Gelbart, T., and Beutler, E. (1990) The human glucocerebrosidase gene and pseudogene structure and evolution. *Genomics* **4**, 87-96.

Hu, P., Wenger, D.A., van Diggelen, O.P., and Kleijer, W.J. (1991) Complementation studies in human and caprine β-mannosidosis. *J. Inher. Metab. Dis.* **14**, 13-17.

Ikonen, E., Baumann, M., Grön, K., Syvänen, A.C., Enomaa, N., Halila, R., Aula, P., and Peltonen, L. (1991) Aspartylglucosaminuria: cDNA encoding human aspartylglucosaminidase and the missense mutation causing the disease. *EMBO J.* **10**, 51-58.

Iwasaki, Y., Tsuji, A., Omura, K., and Suzuki, Y. (1989) Purification and characterization of β-mannosidase from human. *J. Biochem.* **106**, 331-335.

Jackson, R., and Standart, N. (1990) Do the poly (A) tail and 3′ untranslated region control mRNA translation. *Cell* **62**, 15-24.

Jacobs, K.A., Rudersdorf, R., Neill, S.D., Dougherty, J.P., Brown, E.L., and Fritsch, E.F. (1988) The thermal stability of oligonucleotide duplexes is sequence independent in tetraalkylammonium salt solution: application to identifying recombinant DNA clones. *Nucl. Acids Res.* **16**, 4637-4650.

Jolly, R.D., Thompson, K.G., Bayliss, S.L., Vidler, B.M., Orr, M.B., Healy, P.J. (1990) β-Mannosidosis in a Salers calf: a new storage disease of cattle. *N. Z. Vet. J.* **38**, 102-105.

Jones, M.Z and Dawson, G. (1981) Caprine β-mannosidosis: inherited deficiency of β-D-mannosidase. *J. Biol. Chem.* **256**, 5185-5188.

Jones, M.Z. and Laine, R.A. (1981) Caprine oligosaccharide storage disease: accumulation of β-mannosyl (1→4) β-N-acetylglucosaminyl (1→4) β-N-acetylglucosamine in brain. *J. Biol. Chem.* **256**, 5181-5184.

Jones, M.Z., Cunningham, J.G., Dade, A.W., Dawson, G., Laine, R. A., Williams, C. S. F., Alessi, D. M., Mostosky, U.V., and Vorro, J.R. (1982) Caprine β-mannosidosis. in: *Animal Models Inher. Metab. Dis.* (Desnick, R.F., Patterson, D.F., Scarpelli D.G. (eds). Alan R. Liss, Inc., New york, p165-176.

Jones, M.Z., Cunningham, J.G., Dade, A.W., Alessi, D.M., Mostosky, U.V., Vorro, J.R., Benitez, J.T., and Lovell, K.L. (1983) Caprine β-mannosidosis: clinical and pathological features. *J. Neuropathol. Exp. Neurol.* **42**, 268-285.

Jones, M.Z., Rathke, E.J.S., and Cavanagh, K. (1984) β-mannosidosis: prenatal biochemical and morphological characteristics. *J. Inher. Metab. Dis.* **7**, 80-85.

Jones, M.Z. (1986) β-Mannosidosis. *Comparative Pathol. Bull.* **XXI**, 2.

Jones, M.Z., Rathke, E.J.S., Gage, D.A., Costello, C.E., Murakami, K., Ohta, M., and Matsuura, F. (1992) Oligosaccharides accumulated in the bovine β-mannosidosis kidney. *J. Inher. Metab. Dis.* **15**, 57-67.

Jones, M.Z. and Abbitt, B. (1993) Animal model of human disease: bovine β-mannosidosis. *Am. J. Path.* **142**, 957-960.

Kleijer, W.J. Hu, P., Thoomes, R. Boer, M. Huijmans, J.G.M., Blom, W., van Diggelen, O.P., Seemanova, E., and Macek, M. (1990) β-Mannosidase deficiency, heterogeneous manifestations in the first female patient and her brother. *J. Inher. Metab. Dis.* **13**, 867-872.

Kleyn, P.W., Brzustowicz, L. M., Wilhelmsen, K.C., Freimer, N.B ., Miller, J.M., Munsat, J.M., and Gilliam, T.C. (1991) Spinal muscular atrophy is not the result of mutation at the beta-hexosaminidase or GM2-activator locus. *Neurology* **41**, 1418-1422.

Konings, A. (1984) cDNA cloning of acid alpha-glucosidase. in: *Molecular basis of lysosomal storage disorders* (Barranger, J.A., and Brady, R.O. eds) pp. 405-408, Academic Press, Inc. Orlando.

Kornfeld, S. (1987) Trafficking of lysosomal enzymes. *FASEB J.* **1**, 462-468.

Korneluk, R.G., Mahuran, D.J., Neote, K., Klavins, M.H., O'Dowd, B.F., Tropak, M., Willard, H.F., Anderson, M.J., Alexander Lowden, J., and Gravel, R.A. (1986) Isolation of cDNA clones coding for the α-subunit of human β-hexosaminidase. *J. Biol. Chem.* **261**, 8407-8413.

Kozak, M. (1986) Point mutation define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283-292.

Kretz, K.A., Cripe, D., Carson, G.S., Fukushima, H., and O'Brein. J.S. (1992) Structure and sequence of the human α-L-fucosidase gene and pseudogene. *Genomics* **12**, 276-280.

Kreysing, J., von Figura, K., and Gieselman, V. (1990) Structure of the arylsulfatase A gene. *Eur. J. Biochem.* **191**, 637-631.

Kumar, K., Jones, M.Z., Cunningham, G.C., Kelley, J.A., and Lovell, K.L. (1986) Caprine β-mannosidosis: phenotypic features. *Vet. Record* **118**, 325-327.

Kyosaka, S., Murata, S., Nakamura, F., and Tanaka, M. (1985) Purification and kinetic properties of guinea pig liver β-mannosidase. *Chem. Pharm. Bull.* **33**, 256-263.

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

Lathe, R. (1985) Synthetic oligonucleotide probes deduced from amino acid sequence data: theoretical and practical considerations. *J. Mol. Biol.* **183**, 1-12.

Lee, C.C. and Caskey, C.T. (1990) Degenerate primers for DNA amplification. in: *PCR Protocols: a guide to methods and applications* (Innis M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. eds). pp 39-45. Academic Press, Inc., San Diego.

Lee, C.C., Wu, X., Gibbs, R., Cook, R.G. Muzny, D.M., and Caskey, T. (1988) Generation of cDNA probes directed by amino acid sequences: cloning of urate oxidase. *Science* **239**, 1288-1291.

Lee, S., Zambas, E.D., Laurence-Marsh, W., and Redman, C.M. (1991) Molecular cloning and primary structure of Kell blood group protein. *Proc. Natl. Acad. Sci. USA* **88**, 6353-6357.

Levade, T., Graber, D., Flurin, V., Carriere, J., Delisle, M., and Salvayre, R. (1991) Human β-mannosidase: a new clinical form. *Proceedings of European Science Guild on Lysosomal Storage Diseases,* Les Pensieres Workshop, Annecy, France.

Levade, T., Graber, D., Flurin, V., Delisle. M.B., Pieraggi, M.T., Testut, M.F., Carriere, J.P., and Salvayre, R. (1993) Human β-mannosidase deficiency associated with peripheral neuropathy. *Annu. Neurol.* **35**, 116-119.

Lewin, B. (1990) Genes IV pp.578-609, Oxford University Press, Oxford.

Lovell, K.L. and Jones, M.Z. (1983) Distribution of central nervous system lesions in β-mannosidosis. *Acta Neuropathol.* **62**, 121-126.

Lovell, K.L. and Jones, M.Z. (1985) Axonal and myelin lesions in β-mannosidosis: ultrastructural characteristics. *Acta Neuropathol.* **65**, 293-299.

Lovell, K.L. and Boyer, P. (1987) Dysmyelinogenesis in caprine β-mannosidosis: ultrastructural and morphometric studies in fetal optic nerve. *Int. J. Dev. Neuroscience* **5**, 245-253.

Lovell, K.L. (1990) Caprine β-mannosidosis: development of glial and myelin abnormalities in optic nerve and corpus callosum. *Glia* **3**, 26-32.

Lovell, K.L., Jones, M.Z., Patterson, J., Abbitt, B., and Castenson, P. (1991) Thyroid structure and function in bovine β-mannosidosis. *J. Inher. Metab. Dis.* **14**, 228-230.

Lovell, K.L., Kranich, R.J., and Cavanagh, K.T. (1994) Biochemical and histochemical analysis of lysosomal enzyme activities in caprine β-mannosidosis. *Mol. Chem. Neuropath.* **21**, 61-74.

Lundin, L.G. (1987) A gene (bmn) controlling β-mannosidase activity in the mouse is located in the distal part of chromosome 3. *Biochem. Genet.* **25**, 603-610.

Mahuran, D.J., Neote, K., Klavins, M.H., Leung, A., and Gravel, R.A. (1988) Proteolytic processing of pro-α and pro-β precursors from human β-hexosaminidase: generation of the

mature α and $\beta_a\beta_b$ subunits. *J. Biol. Chem.* **263**, 4612-4618.

Mahuran, D.J. (1991) The biochemistry of HEXA and HEXB gene mutations causing G_{M2} gangliosidosis. *Biochem. Biophy. Acta.* **1096**, 87-94.

Martiniuk, F., Mehler, M., Pellicer, A., Tzall, S., Badie, G.L., Hobart, C., Ellenbogen, A., and Hirschhorn, R. (1986) Isolation of a cDNA for human acid α-glucosidase and detection of genetic heterogeneity for mRNA in three α-glucosidase-deficient patients. *Proc. Natl. Acad. Sci. USA* **83**, 9641-9644.

Martiniuk, F., Mehler, M., Tzall, S., Meredith, G., and Hirschhorn, R. (1990) Sequence of the cDNA and 5′-flanking region for human acid α-glucosidase, detection of an intron in the 5′ untranslated leader sequence, definition of 18-bp polymorphisms, and differences with previous cDNA and amino acid sequences. *DNA Cell Biol.* **9**, 85-94.

Matsuura, F., Laine, R.A., and Jones, M.Z. (1981) Oligosaccharides accumulated in the kidney of a goat with β-mannosidosis: mass spectrometry of intact premethylated derivatives. *Arch. Biochem. Biophys.* **211**, 465-493.

Matsuura, F. and Jones, M.Z. (1985) Structural characterization of novel complex oligosaccharides accumulated in the caprine β-mannosidosis kidney. *J. Biol. Chem.* **260**, 15239-15245.

McCabe, N.R. and Dawson, G. (1990) A study of the heterogenous structure of guinea pig lysosomal β-mannosidase using a polyclonal antibody. *Biochem. Biophy. Acta.* **1077**, 133-140.

Menon, K.P., Tieu, P.T., and Neufeld, E.F. (1992) Architecture of the canine IDUA gene and mutation underlying canine mucopolysaccharidosis I. *Genomics* **14**, 763-768.

Miller, R.D., Hoffmann, J.W., Powell, P.P., Kyle, J., Shipley, J.M., Bachinsky, D.R., and Sly, W.S. (1990) Cloning and characterization of the human β-glucuronidase gene. *Genomics* **7**, 280-283 .

Moremen, K.W. (1989) Isolation of a rat liver Golgi mannosidase II clone by mixed oligonucleotide-primed amplification of cDNA. *Proc. Natl. Acad. Sci. USA* **86**, 3276-3280.

Morreau, H., Galjart, N., Gillemans, N., Willemamsen, R., van der Horst, G.T.J., and d'Azzo, A. (1989) Alternative

splicing of β-galactosidase mRNA generates the classic lysosomal enzyme and a β-galactosidase-related protein. *J. Biol. Chem.* **264**, 20655-20663.

Morreau, H., Bonten, E., Zhou, X.Y., and D'Azzo, A. Organization of the gene encoding human lysosomal β-galactosidase. *DNA Cell Biol.* **10**, 495-504.

Moskowitz, S.M., Menson, K.P., Tieu, P.T., and Neufeld, E.F. (1992) Architecture of the gene encoding alpha-L-iduronidase and some mutations underlying human and canine mucopolysaccharidosis I (MPS I). *Am. J. Hum. Genet.* **51**, supple. A680.

Myerowitz, R. and Proia, R. (1984) cDNA clone for the α-chain of human β-Hexosaminidase : deficiency of α-chain mRNA in Ashkenazi Tay-Sachs fibroblasts. *Proc. Natl. Acad. Sci. USA* **81**, 5394-5398.

Myerowitz, R., Piekarz, R., Neufeld, E.F., Shows, T.B., and Suzuki, K. (1985) Human β-Hexosaminidase α-chain; coding sequence and homology with the β-chain. *Proc. Natl. Acad. Sci. USA* **82**, 7830-7834.

Nanba, E., and Suzuki, K. (1991) Organization of the mouse acid β-galactosidase gene. *Biochem. Biophys. Res. Commun.* **178**, 158-164.

Neote, K., Bapat, B., Dumbrille-Ross, A., Troxel, C., Schuster, S.M., Mahuran, D.J., and Gravel, R.A. (1988) Characterization of the human HEXB gene encoding lysosomal β-hexosaminidase. *Genomics* **3**, 279-286.

Neufeld, E.F. (1991) Lysosomal storage diseases. *Annu. Rev. Biochem.* **60**, 257-280.

Nishimura, Y., Rosenfeld, M.G., Kreibich, G., Gubler, U., Sabatini, D.D., Adesnic, M., and Andy, R. (1986) Nucleotide sequence of rat preputial gland β-glucuronidase cDNA and in vitro insertion of its encoded polypeptide into microsomal membranes. *Pro. Natl. Acad. Sci. USA* **83**, 7292-7296.

Noeske, C., and Mersmann, G. (1983) β-D-Mannosidase from human placenta: properties and partial purification. *Hoppe-Seyer's Z. Physiol. Chem.* **364**, 1645-1651.

O'Brien, J.S., de Wet, J., Fukishima, H., Wilcox, E., Dewji, N., Mcgee, J., Warner, T., Yoshida, A., Fluharty, A., Hill, F., and Helinski, D. (1984) Cloning of lysosomal genes. In: *Molecular Basis of lysosomal storage disorders*. (Barranger, J.A and Brady, R.O. eds). pp 387-403. Academic Press, Inc. Orlando.

Occhidoro, T., Beckmann, K.R., Morris, C.P., and Hopwood, J.J. (1989) Human α-L-fucosidase: complete coding sequence from cDNA clones. *Biochem. Biophy. Res. Commun.* **164**, 439-445.

O'Dowd, B., Quan, F., Willard, H.F., Lamhonwah, A.M., Korneluk, R.G., Alexander Lowden, K.J., Gravel, R.A., and Mahuran, D. (1985) Isolation of cDNA clones coding for the β-subunit of human β-hexosaminidase. *Proc. Natl. Acad. Sci. USA* **82**, 1184-1188.

Ohara, O., Dorit, R., and Gilbert, W. (1989) One-sided polymerase chain reaction: the amplification of cDNA. *Proc. Natl. Acad. Sci. USA* **86**, 5673-5677.

Orr, M. (1990) Possible β-mannosidosis in a Salers calf. *Surveillance*. **17**, 9.

Oshima, A., Kyle, J.W., Miller, R.D., Hoffmann, J.W., Powell, P.P., Grubb, J.H., Sly, W.S., Tropak, M., Guise, K.S., and Gravel, R.A. (1987) Cloning, sequencing, and expression of cDNA for human β-glucuronidase. *Proc. Natl. Acad. Sci. USA* **84**, 685-689.

Oshima, A., Tsuji, A., Nagao, Y., Sakuraba, H., Suzuki, Y. (1988) Cloning, sequencing, and expression of cDNA for human β-galactosidase. *Biochem. Biophy. Res. Commun.* **157**, 238-241.

Ottile, S., Chernoff, J., Hanning, G., Hoffman, C.S., and Erikson, R.L. (1991) A fission-yeast gene encoding a protein with features of protein-tyrosine-phosphatase. *Proc. Natl. Acad. Sci. USA* **88**, 3455-3459.

Ovnic, M., Swank, R.T., Fletcher, C., Zhen, L., Novak, E.K., Baumann, H., Heintz, N., and Ganschow, R.E. (1991) Characterization and function expression of a cDNA encoding egasyn (Esterase-22): the endoplasmic reticulum-targeting protein of β-glucuronidase. *Genomics* **11**, 956-967.

Patterson, J.S., Jones, M.Z., Lovell, K.L, and Abbitt, B. (1991) Neuropathology of bovine β-mannosidosis. *J. Neuropath. Exp. Neur.* **30**, 538-546.

Pearce, R.D., Callahan, J.W., Little, P.B., Armstrong, D.T., Kiehm, D., and Clarke, J.T.R. (1987) Properties and prenatal ontogeny of β-mannosidase in selected goat tissues. *Biochem. J.* **243**, 603-609.

Percheron, F., Foglietti, M.J., Bornard, M., and Ricard,B. (1992) Mammalian β-D-mannosidase and β-mannosidosis. *Biochimie*. **74**, 5-11.

Peters, C., Schmidt, B., Rommerskirch, W., Rupp, K., Zuhlsdorf, M., Vingron, M., Meyer, H.E., Pohlmann, R., and von Figura, K. (1990) Phylogenetic conservation of arylsulfatases-cDNA cloning and expression of human arylsulfatase B. *J. Biol. Chem.* **265**, 3374-3381.

Pohlmann, R., Krentler, C., Schmidt, B., Sthroder, W., Lorkowski, G., Culley, J., Mersmann, G., Geier, C., Waheed, A., Gottschalk, S., Grzeschik, K.H., Hasilik, A., and von Figura, K. (1988) Human lysosomal acid phosphatase: cloning, expression and chromosomal assignment. *EMBO. J.* **7**, 2343-2350.

Poenaru, L., Akli, S., Rocchiccioli, F., Eydonx, P., and Zamet, E.P. (1992) Human β-mannosidosis: a 3-year-old boy with speech impairment and emotional instability. *Clin. Genet.* **41**; 331-334.

Poggensee, M. and Leucas, J.N. (1992) A cytogenetic technologist looks at fluroresecense *in situ* hybridization. *Applied Cytogenetics* **18**, 1-5.

Powell, P.P., Kyle, J.W., Miller, R.D., Pantano, J., Grubb, J.H., and Sly, W.S. (1988) Rat liver β-glucuronidase. *Biochem. J.* **250**, 547-555.

Proia, R.L., and Soravia, E. (1987) Organization of the gene encoding the human β-hexosaminidase α-chain. *J. Biol. Chem.* **262**, 5677-5681.

Proudfoot, N. Poly (A) signals. (1990) *Cell* **64**, 671-674.

Quinn,M., Hantzopoulos, P., Fidanza, V., and Calhoun, D.H. (1987) A genomic clone containing the promoter for the gene encoding the human lysosomal enzyme, α-galactosidase A. *Gene.* **38**, 177-188.

Quintern, L.E., Schuchman, E.H., Levran, O., Suchi, M., Ferlinz, K., Reinke, H., Sandhoff, K., and Desnick, R.J. (1989) Isolation of cDNA clones encoding human acid sphingomyelinase: occurrence of alternatively processed transcripts. *EMBO. J.* **8**, 2469-2473.

Reiner, O., Wilder, S., Givol, D., and Horowitz, M. (1987) Efficient *in vitro* and *in vivo* expression of human glucocerebrosidase. *DNA* **6**, 101-108.

Reiner, O., Wigderson, M., and Horowitz, M. (1988a) Structural analysis of the human glucocerebrosidase genes. *DNA* **7**, 107-116.

Reiner, O. and Horowitz,M. (1988b) Differential expression

of the human glucocerebrosidase-coding gene. *Gene* 73, 469-478.

Robertson, D.A., Freeman, c., Nelson, P.V., Morris, C.P., and Hopwood, J.J. (1988) Human glucosamine-6-sulfatase cDNA reveals homology with steroid sulfatase. *Biochem. Biophy. Res. Commun.* 157, 218-224.

Robertson, D.A., Freeman, C., Morris, C.P., Hopwood, J.J. (1992) A cDNA clone for human glucosamine-6-sulphatase reveals differences between arylsulphatases and non-arylsulphatases. *Biochem. J.* 288, 539-544.

Rorman, E.G., and Grabowski, G.A. (1989) Molecular cloning of a human co-β-glucosidase cDNA: evidence that four sphingolipid hydrolase activator proteins are encoded by single genes in human and rats. *Genomics* 5, 486-492.

Ross, J. (1988) Messenger RNA turnover in eukaryotic cells. *Mol. Biol. Med.* 5, 1-14.

Rudkin, G. and Stollar, B. (1977) High resolution detection of DNA/RNA hybrids by indirect immuofluoresence. *Nature* 265, 472-473.

Sasaki, H., Yamada, K., Akasaka, K., Kawasaki, H., Suzuki, K., Saito, A., Sato, M., and Shimada, H. (1988) cDNA cloning, nucleotide sequence and expression of the gene for arylsulfatase in the sea urchin (Hemicentrotus pulcherrimus) embryo. *Eur. J. Biochem.* 177, 9-13.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (eds) (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor Press. New York.

Schatzle, J., Bush, J., and Cardelli, J. (1992) Molecular cloning and characterization of the structural gene coding for the developmentally regulated lysosomal enzyme, α-mannosidase, in Dictyostelium discoideum. *J. Biol. Chem.* 267, 4000-4007.

Schuchman, E., Jackson, C.E., and Desnick, R.J. (1990) Human arylsulfatase B: MOPAC cloning, nucleotide sequence of a full-length cDNA, and regions of amino acid identity with arylsulfatase A and C. *Genomics* 6, 149-158.

Schuchman, E.H., Levran, O., Pereira, L.V., and Desnick, R.J. (1992) Structural organization and complete nucleotide sequence of the gene encoding human acid sphingomyelinase (SMPD1). *Genomics* 12, 197-205.

Scott, H.S., Anson, D.S., Orsborn, A.M., Nelson, P.V.,

Clements, P.R., Morris, C.P., and Hopwood, J.J. (1991) Human α-L-iduronidase: cDNA isolation and expression. *Proc. Natl. Acad. Sci. USA.* **88**, 9695-9699.

Segundo, B.S., Chan, S.J., and Steiner, D.F. (1985) Identification of cDNA clones encoding a precursor of rat liver cathepsin B. *Proc. Natl. Acad. Sci. USA* **82**, 2320-2324.

Sheets, M., Ogg, S.C., and Wickens, M.P. (1990) Point mutation in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation *in vitro*. *Nucl. Acids Res.* **18**, 5799-5805.

Shipley, J.M., Miller, R.D., Wu, B.M., Grubb, J.H., Christensen, S.G., Kyle, J.W., and Sly, W.S. (1991) Analysis of the 5′ flanking region of the human β-glucuronidase gene. *Genomics* **10**, 1009-1018.

Shipley, J.M., Klinkenberg, M., Wu, B.M., Bachinsky, D.R., Grubb, J.H., and Sly, W.S. (1993) Mutational analysis of a patient with mucopolysaccharidosis type VII, and identification of pseudogenes. *Am. J. Hum. Genet.* **52**, 517-526.

Smith, C.W.J., Patton, J.G., and Nadal-Ginard, B. (1989) Alternative splicing in the control of gene expression. *Annu. Rev. Genet.* **23**, 527-577.

Snyder, M., Elledge, S., Sweetser, D., Young, R.A., and Davis, R.W. (1987) λgt11: gene isolation with antibody probes and other applications. *Methods Enzymol.* **154**, 107-128.

Sopher, B.L. (1992a) Purification and characterization of caprine and bovine β-mannosidase/Characterization of a cDNA cloned with anti-β-mannosidase serum. *Thesis*, Michigan State University.

Sopher, B.L., Traviss, C.E., Cavanagh, K.T., Jones, M.Z., and Friderici, K.H. (1992) Purification and characterization of goat lysosomal β-mannosidase using monoclonal and polyclonal antibodies. *J. Biol. Chem.* **267**, 6178-6182.

Sopher, B.L., Traviss, C.E., Cavanagh, K.T., Jones, M.Z., and Friderici, K.H. (1993) Bovine kidney β-mannosidase: purification and characterization. *Biochem, J.* **289**, 343-347.

Sorge, J., West, C., Westwood, B., and Beutler, E. (1985) Molecular cloning and nucleotide sequence of human glucocerebrosidase cDNA. *Proc. Natl. Acad. Sci. USA* **82**, 7289-7293.

Sorge, J., Gross, E., West, C., and Beutler, E. (1990) High level transcription of the glucocerebrosidase pseudogene in normal subjects and patients with Gaucher disease. *J. Clin. Invest.* **86**, 1137-1141.

Stein C., Gieselmann, V., Kreysing, J., Schmidt, B., Pohlmann, R., Waheed, A., Meyer, H.E., O'Brien, J.S., and von Figura, K. (1989) Cloning and expression of human arylsulfatase A. *J. Biol. Chem.* **264**, 1252-1259.

Stoltzfus, L.J., Sosa-Pineda, B., Moskowitz, S.M., Menon, K.P., Dlott, B., Hooper, L., Teplow, D.B., Shull, R.M., and Neufeld, E.F. (1992) Cloning and characterization of cDNA encoding canine α-L-iduronidase. *J. Biol. Chem. 267*, 6570-6575.

Strub, K. and Walter,P. (1989) Isolation of a cDNA clone of the 14-kDa subunit of the signal recognition particle by cross-hybridization of differently primed polymerase chain reactions. *Proc. Natl. Acad. Sci. USA* **86**, 9471-9751.

Sukeno, T., Tarentino, A.L., Plummer, T.H., and Maley, F. (1972) Purification and properties of α-D- and β-D-mannosidases. *Biochem.* **11**, 1493-1550.

Suzuki, K., and Nanba, E. (1991) Organization of the mouse acid β-galactosidase gene. *Biochem. Biophys. Res. Commun.* **178**, 158-164.

Taylor, J.F., Abbitt, B., Walter, J.P., Davis, S.K., Taques, J.T., and Ochoa, R.F. (1992) Estimation of genotype distributions and posterior genotype probabilities for β-mannosidosis in Salers cattle. *Genetics* **135**, 855-868.

Tsuji, S., Martin, B.M., Kaslow, D.C., Migeon, B.R., Choudary, P.V., Stubblefield, B.K., Mayor, J.A., Murray, G.J., Barranger, J.A., and Ginns, E.I. (1987) Signal sequence and DNA-mediated expression of human lysosomal α-galactosidase A. *Eur. J. Biochem. 165,* 275-280.

Tsuji, S., Yamaguchi, Y., Hiraiwa, M., Isobe, T., Okuyama, T., Sakimura,T., Takahashi, Y., Nishizawa, N., Uda,Y., and Miyatake,T. (1989) Molecular cloning of a full-length cDNA for human α-N-acetylgalactosaminidase (α-galactosidase B). *Biochem. Biophy. Res. Commun.* **163**, 1498-1504.

van Pelt, J., Dorland, L., and Duran, M. (1990) Sialyl-α2→6-mannosyl-β1→4-N-acetylglucosamine, a novel compound occurring in urine of patients with β-mannosidosis. *J. Biol. Chem.* **265**, 19685-19689.

Vanin, E.F. (1985) Processed pseudogenes: characteristics

and evolution. *Annu. Rev. Genet.* **19**, 253-272.

Vervoort, R., Lissens, W., and Liebaers, I. (1993) Molecular analysis of a patient with hydrops fetalis caused by β-glucuronidase deficiency, and evidence for additional pseudogenes. *Hum. Mutation* **2**, 443-445.

von Figura, K., Hasilik, A., and Steckel, K. (1984) Lysosomal storage disorders caused by instability of the missing enzymes. in: *Molecular Basis of Lysosomal storage disorders*. (Barranger, J.A and Brady, R.O. eds). pp 133-146. Academic Press, Inc., Orlando.

von Heijne, G. (1983) Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* **133**, 17-21.

von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. *Nucl. Acids Res.* **14**, 4683-4690.

von Heijne, G. (1990) The signal peptide. *J. Membrane Biol.* **115**, 195-201.

Wang, A.M., Bishop, D.F., and Desnick, R.J. (1990) Human α-N-acetylgalactosaminidase: molecular cloning, nucleotide sequence, and expression of a full-length cDNA. *J. Biol. Chem.* **265**, 21859-21866.

Watson, G., Felder, M., Rabinow, L., Moore, K., Labarca, C., Tietze, C., Molen, G.V., Bracey, L., Brabant, M., Cai, J., and Paigen, K. (1985) Properties of rat and mouse β-glucuronidase mRNA and cDNA, including evidence for sequence polymorphism and genetic regulation of mRNA levels. *Gene* **36**, 15-25.

Watts, R.W.E. and Gibbs, D.A (eds). (1986) Glycoproteinases. In: *Lysosomal storage diseases: biochemical and clinical aspects*. pp171-180. Taylor & Francis, London.

Wenger, D.A., Sujansky, E., Fennessey, P.V., and Thompson, J.T. (1986) Human β-mannosidase deficiency. *N. Engl. J. Med.* **315**, 1201-1204.

Wenger, D.A., Chen, Y.Q., Rafi, M.A., Chakraborty, S., and de Gala, G. (1993) Molecular genetics of galactocerebrosidase and acid beta-galactosidase deficiencies. Molecular Biology of Lysosomal Disease (Abstracts). The second international duodecium symposium. Majvik, Finland.

Wickens, M. (1990) How the messenger got its tail: addition of poly (A) in the nucleus. *TIBS*. **15**, 277-281.

Wijburg, H., de Jong, J., Wevers, R., Bakkeren, J., Trijbels, F., and Sengers, R. (1991) Beta-mannosidosis and ethanolaminuria in a female patient. *Eur. J. Pediatr.* **151**, 311.

Wilson, R.J., Morris, C.P., Anson, D.S., Occhiodoro, T., Bielicki, J., Clements, P.R., and Hopwood, J.J. (1990) Hunter syndrome: isolation of an iduronate-2-sulfatase cDNA clone. *Proc. Natl. Acad. Sci. USA* **87**, 8531-8535.

Wood, W.I., Gitschier, J., Lasky, L.A., and Lawn, R.M. (1985) Base composition-independent hybridization in tetramethylammonium chloride: a method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci. USA.* **82**, 1585-1588.

Wozney, J.M. (1990) Using purified protein to clone its gene. *Methods Enzymol* **182**, 735-750.

Xie, B., Kennedy, J.L., Mcinnes, B., Auger, D., and Mahuran, D. (1992) Identification of a processed pseudogene related to the funtional gene encoding the Gm2 activator protein: localization of the pseudogene to human chromosome 3 and the functional gene to human chromosome 5. *Genomics* **14**, 796-798.

Yamamoto, Y., Hake, C.A., Martin, B.M., Kretz, K.A., Ahern-Rindell, A.J., Naylor, S.L., Mudd, M., and O'Brien, J.S. (1990) Isolation, characterization, and mapping of a human acid β-galactosidase cDNA. *DNA Cell Biol.* **9**, 119-127.

Yamauchi, T., Hiraiwa, M., Kobayashi, H., Uda, Y., Miyatake, T., and Tsuji, S. (1990) Molecular cloning of two species of cDNAs for human α-N-acetylgalactosaminidase and expression in mammalian cells. *Biochem. Biophys. Res. Commun.* **16**, 231-237.

Young, R.A. and Davis, R.W. (1983) Yeast RNA polymerase II genes: isolation with antibody probes. *Science* **222**, 778-782.

Zhang, Z.X., Wakamatsu, N., Mules, E.H., Thomas, G.H., and Gravel, R.A. (1994) Impact of premature stop codons on mRNA levels in infantile Sandhoff disease. *Hum. Mol. Genet.* **3**, 139-145.