Order Number 9304692

Characterization of the genetic lesions causing Types A and B Niemann-Pick Disease in the Ashkenazi Jewish population

Lev-Ran, Orna, Ph.D.

City University of New York, 1992

U·M·I

300 N. Zeeb Rd.
Ann Arbor, MI 48106

CHARACTERIZATION OF THE GENETIC LESIONS CAUSING TYPES A AND B NIEMANN-PICK DISEASE IN THE ASHKENAZI JEWISH POPULATION

BY

ORNA LEV-RAN

A dissertation submitted to the Graduate Faculty in Biomedical Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1992

This manuscript has been read and accepted by the Graduate Faculty in Biomedical Sciences in satisfaction of the dissertation requirements for the degree of Doctor of Philosophy.

4/17/92
Date

Desnick, R. J., M.D. Ph.D.
Chair of Examining Committee

5/8/92
Date

Krulwich, T. A., Ph.D.
Executive Officer

Schuchman, E. H., Ph.D.

Wetmur, J. G., Ph.D.

Eng, C. M., M.D.

Rattazzi, M., M.D.

Supervisory Committee

The City University of New York

Abstract

Characterization of the Genetic Lesions Causing Types A and B Niemann-Pick Disease in the Ashkenazi Jewish Population

by

Orna Lev-Ran

Advisor: Edward H. Schuchman, Ph.D.

Acid sphingomyelinase (ASM; sphingomyelin phosphodiesterase, E.C. 3.1.4.12) is the lysosomal enzyme responsible for the degradation of the phospholipid sphingomyelin. A deficiency of this enzyme results in Types A and B Niemann-Pick Disease (NPD). This thesis describes the isolation and characterization of the cDNA and genomic sequences encoding ASM and the identification of three genetic lesions causing NPD. Human urinary ASM was purified and two groups of cDNA clones were isolated using oligonucleotide mixtures based on tryptic peptide sequences. DNA sequencing studies revealed that group-2 cDNAs did not have an internal 172 bp sequence, but had instead an in-frame 40 bp sequence which was not present in group-1 cDNAs. These initial results suggested the occurrence of alternative RNA splicing, a finding that was confirmed when the ASM genomic region was isolated and sequenced. The ASM gene was small (~5 kb) and contained six exons. The alternatively spliced 172 bp sequence was encoded by exon 3, whereas the 40 bp sequence was located at the 5′ end of intron

2. An *Alu 1* element in the reverse orientation was identified in intron 2. The regulatory region was GC rich and contained putative promoter elements including Sp1, TATA, CAAT, NF-1 and AP-1 binding sites.

In order to identify the molecular lesions causing Types A and B NPD, reverse-transcribed total RNA, and/or genomic DNA from Ashkenazi Jewish patients were amplified by the polymerase chain reaction (PCR) using ASM-specific primers. In Type A patients two mutations have been identified: a G to T transversion of nucleotide 3599, (designated R496L), and a T to C transition at nucleotide 1370 (designated L302P). In Type B patients, a three base in-frame deletion of nucleotides 3933-3935, (designated ΔR608) was identified. Screening other NPD patients and normal individuals with allele-specific oligonucleotides (ASOs) excluded the possibility that these nucleotide changes were common polymorphisms and revealed that the R496L and the L302P mutations occurred in about 32% and 24% of the Ashkenazi Jewish NPD Type A alleles, respectively. The R496L mutation also occurred as one allele in an Ashkenazi Jewish Type B patient and the second allele of this patient was the ΔR608 mutation. This small deletion was homoallelic in another unrelated Ashkenazi Jewish Type B patient, but was not found in Type A patients. To provide additional evidence for the authenticity of these mutations, the mutant cDNAs constructs were transiently expressed in *COS-1* cells. As expected, they did not express catalytically active ASM. These findings should facilitate molecular carrier detection and prenatal diagnosis for NPD in the Ashkenazi Jewish population and provide initial genotype/phenotype correlations for this disease.

Acknowledgements

I am indebted to Dr. Edward Schuchman for allowing me to complete my dissertation under his supervision, and for accepting and respecting me for who I am. I would like to express my gratitude to Dr. Robert Desnick who believed in me and gave me the opportunity to become a part of a remarkable scientific community.

I would like to thank Drs. Mariko Suchi, Lothar Quintern and Tsutomu Takahashi, for their invaluable contribution to this project. Dr. David Bishop, Ms. Lygia Periera, Ms. Safiana Kats and Mr. Constantine Zamfirescu provided me with an expert technical and computer assistance, and I thank them for that.

I am grateful to Efrat Rorman, Dr. Anne Wang and Dr. Cristine Jackson for standing by me throughout the whole process.

Finally, I want to express my deepest love and appreciation to my family, here and in Israel and to my friends, for their unfailingly love, patience and support.

This work is dedicated to the memory of my friend Idit Gal who inspired me with her love for the mystery of nature.

Forward

Portions of this thesis have been presented in the following publications:

1. 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 Journal* 9, 2469-2473.

2. Levran, O., Desnick, R. J. and Schuchman, E. H. (1991): Niemann-Pick disease: A frequent missense mutation in the acid sphingomyelinase-encoding gene of Ashkenazi Jewish Type A and B patients. *Proc. Natl. Acad. Sci. USA* 88, 3748-3752.

3. Schuchman, E. H., Levran, O., Suchi, M. and Desnick, R. J. (1991): An *MspI* polymorphism in the human acid sphingomyelinase gene (SMPD1) at 11p15.1-p15.4. *Nucl. Acid Res.* 19(11), 3160.

4. Levran, O., Desnick, R. J. and Schuchman, E. H. (1991): Niemann-Pick Type B disease: Identification of a single codon deletion in the acid sphingomyelinase gene and genotype/phenotype correlations in Type A and B patients. *J. Clin. Invest.* 88, 806-810.

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

6. Levran, O., Desnick, R. J. and Schuchman, E. H. (1992): Identification and expression of a common missense mutation (L302P) in the acid sphingomyelinase gene of Ashkenazi Jewish Type A Niemann-Pick disease patients. (submitted to *Blood*).

Table of Contents

Title Page	i
Approval Page	ii
Abstract	iii
Acknowledgements	v
Forward	vi
Table of Contents	vii
List of Figures and Tables	xii
Abbreviations	xiv
Chapter One: Background	1
Lysosomes and Lysosomal Enzymes: Overview	2
Acid Sphingomyelinase and Sphingomyelin	3
Niemann-Pick Disease	6
Chapter Two: Isolation of cDNA clones encoding ASM	11
Abstract	12
Introduction	13
Experimental Procedures:	
Amino Acid Microsequencing	13
Synthesis of Oligonucleotide Mixture	14
cDNA Library Screening	14
DNA Sequencing and Computer Analysis	15

Results:

Amino Acid Sequencing and Oligonucleotide Synthesis 15

Isolation of cDNA Clones Encoding ASM 16

Discussion 17

Chapter Three: Structural Organization and Complete Nucleotide Sequence of the Human ASM Gene 24

Abstract 25

Experimental Procedures:

Genomic Library Screening 26

DNA Sequencing and Computer Analysis 27

Detection of the 506 Polymorphism 27

Results:

Isolation of Human ASM Genomic Clones 28

Organization of the Human ASM Gene 28

Intron/Exon Junctions 30

Analysis of the ASM Promotor Region 31

ORFs in the Human ASM Gene 31

*Msp*I Polymorphism 32

Discussion 32

Chapter Four: A Frequent Missense Mutation (R496L) in the ASM Gene of Ashkenazi Jewish Type A and B Patients 52

Abstract 53

Introduction 54

Experimental Procedures:

Cell Lines 55

Enzyme and Protein Assays 56

cDNA and Genomic Amplification and Sequencing of the Mutant Allele 56

Dot-Blot Analysis 58

Results:

Identification of the R496L Mutation in an Ashkenazi Jewish Type A NPD Patient 59

Occurrence of the R496L Mutation in Ashkenazi Jewish and Non-Jewish NPD Type A Families 60

Occurrence of the R496L Mutation in Ashkenazi Jewish and Non-Jewish NPD Type B Families 61

Discussion 61

Chapter Five: A Common Missense Mutation (L302P) in the ASM Gene of Ashkenazi Jewish Type A NPD Patients 71

Abstract 72

Introduction 73

Experimental Procedures:

Cell Lines 73

Identification of the L302P Mutation 73

Confirmation of the L302P Mutation by Dot-Blot Hybridization 74

Results:

Identification of the L302P Mutation in Proband 2 75

Occurrence of the L302P Mutation in Types A and B NPD 76

Discussion 77

Chapter Six: A Single Codon Deletion (ΔR608) in the ASM Gene and Genotype/Phenotype Correlations in Type A and B Patients 84

Abstract 85

Introduction 86

Experimental Procedures:

Cell Lines 87

cDNA and Genomic Amplification and Sequencing 87

Dot-Blot Analysis of the Mutant Alleles 88

Results:

Identification of the ΔR608 Mutation in an Ashkenazi Jewish Patient with Type B NPD 89

Occurrence of the ΔR608 Mutation in Types A and B NPD 90

Comparison of the Residual ASM Activities in Type A and B NPD Patients 90

Discussion 91

Chapter Seven: Transient Expression of the Three ASM Mutations in *COS-1* Cells 99

Abstract 100

Introduction 102

Experimental Procedures:

Site-Directed Mutagenesis 106

R496L Construct 107

L302P Construct 108

ΔR608 Construct 108

Transient Expression of the Mutant Constructs in *COS-1* Cells 109

Results:

Transient Expression of the Mutant Constructs 109

Discussion 110

Bibliography 120

List of Figures and Tables

Chapter Two:

Table I	Tryptic peptide sequence of human ASM and corresponding oligonucleotide mixtures	21
Figure 1	Nucleotide and amino acid sequence of pASM-1 and pASM-2	22

Chapter Three:

Figure 2	Nucleotide sequence of the gene encoding human ASM	37
Table II	Intron/exon junctions in the ASM gene	40
Figure 3	Schematic representation of the human ASM gene	41
Figure 4	*Alu* 1 sequences within intron 2 of the ASM gene	43
Figure 5	Putative promotor region of the human ASM gene	45
Figure 6	Other open reading frames (ORF) in the ASM gene	47
Figure 7	Scheme of alternative splicing in the ASM gene	49
Table III	Putative promotor regions of genes encoding lysosomal enzymes	51

Chapter Four:

Figure 8	Partial sequence of the amplified ASM cDNA from proband 1	66
Figure 9	Dot-blot hybridization with allele-specific oligonucleotides	68

Table IV	Frequency of the R496L alleles in Ashkenazi Jewish and non-Jewish NPD families	70

Chapter Five:

Figure 10	Partial sequence of the amplified ASM cDNA from proband 2	79
Figure 11	Identification of the L302P mutation in proband 2 and family members	81
Table V	Synthetic oligonucleotides for PCR-amplifications	83

Chapter Six:

Figure 12	Three base deletion in the ASM gene from proband 3	94
Figure 13	Genotype analysis of proband 3 and family members by dot-blot hybridization	96
Table VI	Frequency of the ΔR608 mutation in Ashkenazi Jewish and non-Jewish families with Types A and B NPD	98
Table VII	Genotype/phenotype correlation in Types A and B NPD	98

Chapter Seven:

Table VIII	Synthetic oligonucleotides for cDNA amplification of the mutant constructs	112
Figure 14	Construction of the p91023(B) expression vector containing the R496L mutation	113
Figure 15	Construction of the p91023(B) expression vector containing the L302P mutation	115
Figure 16	Construction of the p91023(B) expression vector containing the ΔR608 mutation	117
Table IX	Transient expression of ASM mutations in *COS-1* cells	119

Abbreviations

ASM	Acid Sphingomyelinase
ASO	Allele-Specific Oligonucleotide
bp	Base-pairs
BSA	Bovine Serum Albumin
CRIM	Cross Reactive Immunological Material
kb	Kilobases
l	Liter
MW	Molecular Weight
NPD	Niemann-Pick Disease
NSM	Neutral Sphingomyelinase
nt	Nucleotide
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
(PSA11-SPM)	*N*-[11 (1-pyrene)sulphonyl-aminoundecanyl] sphingomyelin
RER	Rough Endoplasmic Reticulum
U	Unit
Y	a pyrimidine nucleoside
N	a nucleoside
R	a purine nucleoside

Chapter One

BACKGROUND

Lysosomes and Lysosomal Enzymes: Overview

Lysosomes are membrane-bounded cytoplasmic organelles which contain hydrolases that have acidic pH optima and participate in intracellular digestion. They are found in all cell types, excluding the mammalian erythrocytes, and may make up to as much as 10% of the volume of certain cell types which specialize in phagocytosis (e.g. macrophages) (Kornfeld and Mellman, 1989). More than 50 lysosomal enzymes are known to participate in the degradation of glycoproteins, glycolipids and glycosaminoglycans. Lysosomal enzymes are synthesized as prepropolypeptides in the rough endoplasmic reticulum (RER). They contain a hydrophobic amino terminal signal peptide which initiates their transport into the lumen of the RER, where they subsequently undergo extensive modifications, including glycosylation of selected asparagine residues. Following signal peptide cleavage, the enzymes are transported to the Golgi and extensive oligosaccharide processing begins (Farquhar *et al.*, 1985). The critical modification for these hydrolyses is the acquisition of specific phosphomannosyl residues that are recognized by mannose-6-phosphate receptors. The interaction of the enzymes with these receptors is the major, but not the only, mechanism by which the lysosomal enzymes are specifically sorted and targeted to lysosomes. The disassociation of the enzyme/receptor complex occurs upon acidification of the lysosomal compartment and the receptor is recycled to the Golgi (Dahn *et al.*, 1989). Some acid hydrolyses are stimulated by small, heat-stable proteins which interact with either the substrate or the enzyme; the so called "activator proteins" (O'Brian and Kishimoto, 1991).

Due to the recent advances in molecular biology, it has been possible to determine

the molecular structure of many of the genes encoding lysosomal enzymes and to characterize the molecular lesions that cause some lysosomal storage diseases. Lysosomal storage diseases are caused by the absence or malfunction of specific hydrolases. The specific substrates gradually accumulate and interfere with the normal cellular activity (von Figura *et al.*, 1984). Most of these disorders are very severe and some are lethal (Spence and Callahan, 1989).

Acid Sphingomyelinase and Sphingomyelin

Acid sphingomyelinase (ASM; HGMW-approved symbol, SMPD1; sphingomyelin phosphodiesterase E.C 3.1.4.12) is a lysosomal enzyme responsible for the degradation of the phospholipid sphingomyelin. The degradation products are ceramide and phosphoryl choline (Levade *et al.*, 1986). The deficiency of this enzymatic activity results in the accumulation of sphingomyelin and other lipids, such as cholesterol and bis-phosphate, in the lysosomes of Niemann-Pick disease (NPD) patients.

Sphingomyelin is a membrane phospholipid that contains sphingosine as a backbone. Sphingosine is an amino alcohol that contains a long, unsaturated hydrocarbon chain. The amino group of the sphingosine backbone is linked to a fatty acid by an amide bond. The primary hydroxyl group is esterified to phosphoryl choline. The phosphoryl choline is the hydrophilic unit, also called the polar head group. The hydrocarbon chain and the fatty acid are the hydrophobic unit of this amphiphatic molecule (Callahan *et al.*, 1983). Sphingomyelin biosynthesis *de novo* involves ceramide formation from serine and fatty acyl-CoA followed by addition of the phosphoryl choline

(Sribney and Kennedy, 1958; Stoffel and Melzner, 1980). It has been hypothesized that ceramide may also be hydrolyzed to form free sphingosine by the lysosomal enzyme, acid ceramidase. Sphingosine can potentially inhibit protein kinase C activity and effect this important second messenger pathway (Merril and Jones, 1990). Sphingomyelin turnover involves the action of sphingomyelinases and ceramidases; both of which have been found in acidic and neutral forms (Schneider and Kennedy, 1967). Human ASM is a glycoprotein that is optimally active in acidic pH optimum (4.6-5.0) in all tissues (Gatt, 1970). The catalytically active form of ASM is probably monomeric. Subunit size has been estimated between 70-105 kDa (Pentchev *et al.*, 1977; Jones *et al.*, 1981; Sakuragawa, 1982; Quintern *et al.*, 1987; Quintern et *al.*, 1989). The predicted MW of the ASM polypeptide according to Quintern *et al.* is about 61 kDa.

In addition to ASM at least three other sphingomyelinases have been identified in humans. These include a Mg^{++}-dependent neutral sphingomyelinase (NSM), with a pH optimum of about 7.5, primarily located in brain gray matter (Chatterjee and Ghosh, 1989). The NSM activity is normal in NPD patients (Gatt *et al.*, 1978). There has been a report of immunological cross-reactivity between ASM and NSM from brain (Maruyama and Arima, 1989). A Zn^{++}-dependent ASM also has been found in serum 0(Spence *et al.*, 1989), and a placental sphingomyelinase (Kurth and Stoffel, 1991) with a different amino acid sequence and a slightly different pH optimum than the acid sphingomyelinase that was purified from urine (Quintern *et al.*, 1987; Quintern et *al.*, 1989). Sphingomyelinases have been described in non-mammalian species as well. For example, sphingomyelinases from *Leptospira intrrogans* and *Bacillus cereus* have been

isolated (Segers *et al.*, 1990; Yamada *et al.*, 1988; Gilmor *et al.*, 1989). The gene encoding the sphingomyelinase from *Bacillus cereus* was found to form a cluster with the phosphatidylcholine phospholipase C gene. These genes include a putative signal sequence, exhibit a unique hemolytic activity in the presence of Ca^{++} and Mg^{++}, and show a high degree of similarity.

Biochemical studies of ASM have been facilitated by the availability of chromogenic, fluorescent and radioactive substrates. Two colored derivatives of sphingomyelin were synthesized (Gatt *et al.*, 1978; Gal and Fash, 1976). In addition, substrate with fluorescent yield (PSA11-SPM) was synthesized by replacing the fatty acid with pyrene dodecanoic acid (Klar *et al.*, 1988). Radiolabeled sphingomyelin was prepared by reducing the double bonds in ceramide with [^{3}H] or the incorporation of [^{3}H] methyl or [^{14}C] methyl into the choline moiety (Brady *et al.*, 1966).

The activator protein SAP D was found to specifically stimulate ASM *in vitro* (Morimoto *et al., 1988*). Lysosomal storage diseases can be caused by activator deficiencies, as has been shown in a Gaucher Disease variant (SAP-2 deficiency) (Christomanou *et al.*, 1986). However, to date no cases of NPD have attributed to faulty or absent activator protein.

Niemann-Pick Disease

Classification and Clinical Heterogeneity

The Niemann-Pick group of diseases can be divided into two groups according to the recent classification of Spence and Callahan (1989). Both are characterized by hepatosplenomegaly, foamy reticuloendothelial cells in bone marrow and increased amounts of sphingomyelin, cholesterol, glycosphingolipids and bis-phosphate in visceral organs. The clinical and biochemical variability in both types indicates genetic heterogeneity. NPD Type I includes the Type A and B forms, which are named according to the older, more widely used classification (Brady, 1983). Patients with Type I disease show dramatic ASM deficiency (<10% of the normal activity) leading to massive storage of sphingomyelin and cholesterol in their lysosomes. Affected patients are characterized by hepatosplenomegaly, foamy reticuloendothelial cells in their bone marrow and variably increased amounts of sphingomyelin, cholesterol, glycosphingolipids and bis phosphate in visceral organs (Croker and Farber, 1958). Type A NPD is the acute form, and is characterized by hepatosplenomegaly and fatal neurodegeneration. Onset is signaled by failure to thrive in the first six months of life. Death usually results by three years of age due to pulmonary infections. Type B NPD is the subacute form, characterized by hepatosplenomegaly, slow progression and little or no neurologic manifestations (Pavone *et al.*, 1986). Affected patients may survive into late adolescence or adulthood. There is variability in the severity of the clinical course of this type of disease. NPD Type II includes patients with lysosomal accumulation of one or more of

the lipids described above. ASM activity is variable and it has been demonstrated that some patients may have a defect in cholesterol esterification (Blanchette-Mackie *et al.*, 1988). This thesis discusses only the Type A and B forms of NPD.

History and Pathology

NPD was first described in 1914 by A. Niemann (Niemann, 1914) and identified later by L. Pick (Pick, 1927) as a disease that is distinct from Gaucher Disease. In 1966, Brady (Brady *et al.*, 1966) described the ASM deficiency in NPD patients. The cellular pathology of both Types A and B NPD is characterized by reticuloendothelial cells that accumulate sphingomyelin, cholesterol and related lipids. These "Foam Cells", are different from "Gaucher Cells", and have been referred to as "Niemann-Pick cells". The appearance of these cells is caused by lipid droplets that are formed in their cytoplasm. Basically, all the tissues of NPD patient have "Foam Cells", particularly the reticuloendothelial organs.

Enzymatic Activity in NPD

When measured *in vitro,* cultured fibroblasts or lymphoblasts, as well as some organ extracts (e.g. liver, kidney and brain) from Type A NPD patients, have <3% of ASM activity. In Type B patients, the residual activities usually range from 0-10% of normal. When ASM activity is measured *in situ* by cell loading assays which use apolipoprotein E to transport the fluorescent substrate directly to the lysosomes (Agmon *et al.*, 1991), the results for Type A patients is similar to those obtained *in vitro.*

However, the results are more variable and the residual activities are higher for Type B patients. Some patients with Types A and B NPD were shown to synthesize ASM in normal amounts and normal size (Rousson *et al.*, 1986). There was no restoration of enzyme activity upon fusion of fibroblasts from patients with NPD Types A and B, providing evidence for allelic, intragenic defects in these forms of NPD (Besley *et al.*, 1980). Limited immunological studies have been performed on Type A and B patients, and show normal levels of ASM cross reactive immunological material (CRIM) (Rousson *et al.*, 1987). Prenatal enzymatic diagnosis (Patrick *et al.*,1977; Maziere *et al.* 1978), and enzymatic diagnosis of affected patients with either Type A or B NPD can be made reliably, but the enzymatic detection of obligate heterozygotes has proven problematic, particularly using peripheral leukocytes as the enzyme source (Vanier *et al.*, 1985).

It has been proposed that the accumulation of lysosphingomyelin, as well as other lysosphingolipids, can be the cause of the cytotoxic effects and cell dysfunction found in NPD and other lysosomal storage diseases. ASM degrades lysosphingomyelin as well as the parental sphingomyelin, so in its deficiency this derivate accumulates as well. The lysosphingolipid derivates are also potent inhibitors of proteinase kinase C, so they may also have an effect on signal transduction (Hannun and Bell, 1989).

Mode of Inheritance and Incidence

The inheritance of both Types A and B NPD is autosomal recessive. The overall distribution of NPD is panethnic, but Types A and B NPD occur at least ten times more frequently among individuals of Ashkenazi Jewish ancestry than in the general population.

The incidence of Type A NPD is estimated at about 1 in 40,000, with a gene frequency of about 1 in 200 and an heterozygote frequency (2pq) of about 1 in 100 (Goodman, 1979). Type B NPD is estimated at about 1 in 80,000 in this population. Thus, the combined heterozygote frequency for Types A and B NPD has been estimated to be about 1 in 70 among individuals of Ashkenazi Jewish decent.

Treatment

Currently, there is no specific treatment for NPD. Liver transplantation has performed in Type A patients (Daloze *et al.*, 1977; Delvin *et al.*, 1974), with little success. Implantations of human amniotic membrane were performed on Type B patient (Scaggiante *et al.*, 1987). Epithelial amniotic cells are not immunogenic and can possibly secrete ASM in significant quantities. There was clinical improvement in the patient and the hepatomegaly decreased, but repeated implantations were necessary, probably due to the limited time of survival of the implanted cells. Bone marrow transplantation was carried out on a Type B patient, in order to provide a lifelong source of ASM and to clear the accumulated sphingomyelin. The storage material was cleared from the liver and the bone marrow. This type of treatment was thus suggested for Type B NPD patients with early and severe liver disease (Vellodi *et al.*, 1987). Enzyme replacement therapy has been considered as a possible treatment for NPD as well as other lysosomal storage disorders. Such a treatment is being performed for Type I (non-neurologic) Gaucher disease with encouraging results (Barton *et al.*, 1990; Parker *et al.*, 1991). In order to achieve permanent enzyme replacement, investigators have considered the use of somatic

gene transfer (Gilboa *et al.*, 1986). Retroviral vectors carrying normal human genes have been used to transfect bone marrow cells to achieve long term expression of those genes. When this treatment involves bone marrow transplantation, it is limited by the need for histocompatible donor, so the ideal therapy would be to introduce the normal genes directly to the patient's bone marrow. Studies of this kind are underway in humans with adenosine deaminase deficiency.

Animal Models

Animal models have been described for NPD in three species: canine (Banzai *et al.*, 1979), feline (Wenger *et al.*, 1980) and murine (Miyawaki *et al.*, 1986; Sakiyama *et al.*, 1986; Kitagawa, 1987). A breeding colony has only been established for the murine models, which should be useful for studies of enzyme replacement and gene transfer.

Chapter Two

ISOLATION OF cDNA CLONES ENCODING HUMAN ASM : OCCURRENCE OF ALTERNATIVELY PROCESSED TRANSCRIPTS

Abstract

Twelve tryptic peptides (128 residues) from the purified human urinary ASM (Quintern *et al.*, 1987; Quintern *et al.*, 1989) were microsequenced. Based on regions of minimal codon redundancy, four oligonucleotide mixtures were synthesized and oligonucleotide mixture 1 (20 mer; 256 mix) was used to screen 3×10^6 independent recombinants from a human fibroblast cDNA library. 92 Putative positive clones were purified and analyzed by Southern hybridization with oligonucleotide mixtures 2-4. These studies revealed two groups of clones: group I (80 clones; inserts ranging from ~1.2 to 1.6 kb) hybridized with oligonucleotide mixtures 1-4, while group II (12 clones; inserts ranging from ~1.2 to 1.4 kb) hybridized with oligonucleotide mixtures 1-3. Several group II clones had larger inserts than those in group I, but did not hybridize with oligonucleotide mixture 4. Representative clones from group I (pASM-1) and group II (pASM-2) were sequenced. pASM-1 contained a 1879 bp insert which was collinear with 95 microsequenced amino acids, while the pASM-2 1382 bp insert was collinear with 62 microsequenced residues. Notably, pASM-2 did not have an internal 172 bp sequence encoding 57 amino acid residues, but had instead an in-frame 40 bp sequence encoding 13 amino acids which was not present in pASM-1. These findings demonstrate the presence of two distinct ASM transcripts in human fibroblasts and suggest the occurrence of alternative processing of the mRNA encoding this lysosomal hydrolase.

Introduction

Recently, human urinary ASM was purified ~30,000-fold from urine by the method of Quintern *et al.* (1987), and shown to be a monomeric protein with a MW of ~72 kDa (Quintern *et al.*, 1989). The purified ASM was treated with glycopeptidase F and the MW of the denatured, deglycosylated monomeric enzyme was ~61 kDa. From these results it was estimated that the purified urinary form of human ASM is composed of ~550 amino acid residues. This finding predicted a cDNA coding sequence of at least 1650 bp. The purified enzyme had both sphingomyelinase and phospholipase C activity towards the respective substrates (Quintern *et al.*, 1989).

This chapter describes the isolation and characterization of two different groups of cDNA clones encoding human ASM, providing evidence for the occurrence of alternative splicing.

Experimental Procedures

Amino Acid Microsequencing

Highly enriched ASM (~0.5 mg) was obtained from Dr. Konard Sandhoff's laboratory, reduced with dithiothreitol, alkylated with iodoacetamide and hydrolyzed with trypsin (trypsin:protein ratio of 1:100). The tryptic peptides were separated on an HIBAR Lichrospher column (500 CH-8;10 um) with a 0-70% acetonitrile gradient in 0.05% trifluroacetic acid, and the single peptide peaks were then rechromatographed on the same

column with a 0-70% acetonitrile gradient in 25 mM ammonium acetate buffer, pH 6.0. Only peptides which were eluted with high yields were sequenced by automated Edman degradation using an Applied Biosystems gas-phase amino acid sequencer with an on-line PTH analyzer (Beyreuther *et al.*, 1983; Stone and Williams, 1986).

Synthesis of Oligonucleotide Mixtures

Oligonucleotide mixtures corresponding to amino acid sequences of minimal codon redundancy were synthesized on an Applied Biosystems DNA Synthesizer using phosphoramidite chemistry (Itakura *et al.*, 1984; Caruthers, 1985). After synthesis, the mixtures were analyzed on 20% polyacrylamide - 8 M urea gels prior to use and, if necessary, purified by isolating the band and eluting the DNA overnight. 5′- end labeling reactions were carried out using [γ-^{32}P] ATP (5000 Ci/mmol) using T4 polynucleotide kinase (Maniatis *et al.*, 1982).

cDNA Library Screening

The pcD human fibroblast cDNA library (Okayma and Berg, 1983) was plated at a density of ~10^4 colonies/150mm^2 petri dish and screened by standard methods (Maniatis *et al.*, 1982). Prehybridization (6×SSPE, 10×Denhardts solution, 0.5% SDS) and hybridization (6×SSPE, 10×Denhardts solution, 0.5% SDS containing 50 pmol γ-^{32}P labeled oligonucleotide) were preformed at 6° C below the lowest T_m for the

oligonucleotide mixture. Filters were washed at the lowest T_m for each oligonucleotide mixture in 6×SSC, 0.1% SDS for 3-6 h.

DNA Sequencing and Computer Analyses

cDNA inserts (and/or restriction fragments) from putative positive clones were subcloned into M13 (Messing *et al.*, 1981) and/or Bluescript (Stratagene, Inc.) vectors and sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977). Sequencing primers were synthesized on an Applied Biosystems DNA Synthesizer as described above. Computer analyses were performed using the University of Wisconsin Genetics Computer Group DNA Sequence Analysis Software.

Results

Amino Acid Sequencing and Oligonucleotide Synthesis

Since the N-terminus of ASM was blocked, the purified urinary enzyme was digested with trypsin and the tryptic peptides were purified by HPLC. Twelve peptides were microsequenced (~ 20% of the purified enzyme) and four were chosen for the construction of corresponding oligonucleotide mixtures based on regions of minimal codon redundancy. No similarities were found with the ASM sequences and other protein sequences in the NBRF protein database. Table I shows the four tryptic peptides and the

corresponding synthetic oligonucleotide mixtures which were constructed for cDNA library screening.

Isolation of cDNA Clones Encoding ASM

Using oligonucleotide mixture 1 as a probe, ~3 x 10^6 independent recombinants from a human fibroblast pcD library (Okayama and Berg, 1983) were screened. The 92 putative positive clones were isolated, purified, and characterized by restriction enzyme and Southern analyses with oligonucleotide mixtures 2-4. Based on these results, two groups of clones were identified. Group I (80 clones with inserts ranging from ~ 1.2 to 1.6 kb) hybridized with oligonucleotide mixtures 1-4, while group II (12 clones with inserts ranging from ~1.2 to 1.4 kb) hybridized with oligonucleotide mixtures 1-3. Intriguingly, some of the group II clones were larger than the group I clones, despite the fact that they did not hybridize with oligonucleotide mixture 4. Subsequently, a human placental cDNA library was screened with a radiolabeled 450 bp cDNA fragment obtained from a group I insert. A total of 21 putative clones were purified and analyzed, resulting in the identification of 18 additional group I and 3 additional group II inserts.

Based on these findings, representative clones from group I (pASM-1) and group II (pASM-2) were selected for DNA sequencing by the dideoxy chain termination method (***Figure 1***). The sequence of clone pASM-1 (1879 bp), obtained from the placental cDNA library, was collinear with 96 microsequenced amino acid residues from human ASM, demonstrating its authenticity. There were three discrepancies between the microsequenced amino acid sequences from the tryptic peptides and those predicted from

the cDNA sequence (residues 123, 252 and 467). No poly (A) tail was found in pASM-1, although a consensus polyadenylation signal (AATAAA) was present at the 3′ terminus of the insert. pASM-2 (1382 bp) was collinear with 78 microsequenced amino acid residues and had the consensus polyadenylation signal and a poly (A) tail (*Figure* 1). The polyadenylation consensus sequence was 27 bp from the poly (A) tract in pASM-2. Of particular interest, pASM-2 did not have an internal 172 bp sequence encoding 57 ASM amino acids which were present in pASM-1; this deleted sequence encoded the peptide corresponding to oligonucleotide mixture 4 (*Figure 1*; T-4). In place of this sequence, pASM-2 contained an in-frame 40 bp sequence encoding 13 amino acids which were not present in pASM-1.

Discussion

The recent purification of human ASM permitted the isolation of tryptic peptides for microsequencing and the synthesis of oligonucleotide mixtures based on regions of low codon redundancy. Using the oligonucleotide mixtures as probes, a total of 92 putative positive cDNA clones for human ASM were isolated and purified from a pcD human fibroblast library. Southern hybridization analyses revealed two groups of cDNA clones. Group I clones had inserts which hybridized with all four oligonucleotide mixtures whereas the inserts of group II clones only hybridized with oligonucleotide mixtures 1-3, despite the fact that some of the group II inserts were larger than those found in group I. The nature of this apparent discrepancy was revealed by sequencing

the longest cDNA insert from each group. pASM-1 (group I) had an internal 172 bp sequence encoding 57 amino acids that was not present in pASM-2. In place of this sequence, pASM-2 had a 40 bp in-frame sequence which encoded 13 amino acid residues which were not present in pASM-1.

The finding of two authentic pASM cDNA clones which differed by the presence of two unique internal nucleotide sequences provided evidence for alternative splicing of the ASM transcript. The possibility that these results were caused by a cloning artifact was excluded by hybridization studies and sequencing of additional cDNA clones. Both libraries, placental and fibroblast, contained group I and II cDNA clones. Presumably, group I clones corresponded to the normally processed transcript that encodes human ASM, since they contained peptide T-4 and represented about 90% of the clones identified in the fibroblast and placental libraries. The fact that pASM-2 was missing 172 nucleotides but contained 40 in-frame nucleotides suggests that ~10% of the ASM transcripts are alternatively spliced. Analysis of the unique 40 bp sequence in pASM-2 revealed that the first and the last two nucleotides (GT and AG, respectively) are the 5′ donor and 3′ acceptor consensus dinucleotides for intron splicing, so this transcript may includes an intronic sequence that can serve as a cryptic 3′ donor splice site. The fact that the putative splicing event deletes 172 bp which are replaced by 40 bp (encoding 13 amino acids) without altering the reading frame suggests that both transcripts may encode functional proteins. The two transcripts might produce enzymes that recognize and preferentially bind different substrates. Alternatively, this region of the enzyme may be important for subcellular targeting since the sequences which are deleted in pASM-2

contain a potential *N*-glycosylation site. The presence of two human transcripts also may provide insights into the clinical heterogeneity of NPD. A mutation in the sequence which is deleted in pASM-2 would only affect the expression of the group I transcripts, leaving ~10% of the transcripts unaffected. If the group II transcripts express a functional protein this would result in residual ASM activity. On the other hand, the alternatively spliced transcript may not be functional. In fact, alternative splicing of the human β-glucuronidase transcripts supports this possibility (Oshima *et al.*, 1987). Of the two β-glucuronidase cDNAs isolated, one had a 153 bp deletion (corresponding to an entire exon) which expressed an enzyme protein that was not catalytically active. Transient expression of full-length ASM group I and II cDNAs will determine the functional integrity of the ASM transcripts.

Finally, since the nucleotide sequences of pASM-1 and pASM-2 are identical, except for the spliced region, it is unlikely that the two groups of clones represent transcripts of different ASM genes or pseudogenes. However, a complete understanding of the molecular nature of the putative alternative splicing event will require additional studies, including S1 nuclease analysis of the ASM RNA in different tissues and the isolation and sequencing of the implicated genomic region(s).

While this work was in progress, Dr. Suchi and Dr. Takahashi joined Dr. Schuchman's laboratory and continued with the cDNA cloning aspects of this project. Some of this work is relevant to this thesis and will be discussed below. Northern hybridization analyses revealed the presence of a single ~2.5 kb transcript in placental

poly(A)+ RNA. Subsequent RNase protection studies demonstrated the occurrence of both type 1 and 2 transcripts. Full-length type 1 cDNAs were isolated including the 2347 bp cDNA, pASM-1FL. The size of the pASM-1FL insert was consistent with the occurrence of the ~2.5 kb transcript. The ~150 bp difference was presumably due to the length of the poly(A) tract. The full-length cDNA had an open reading frame of 1890 bp which contained two in-frame potential initiation codons. It is possible that both initiation ATGs are being used and studies are underway to address this question. The optimal signal peptide cleavage would probably occur after amino acid residue 46. The predicted 14 amino acids before the signal peptide cleavage site have a particular hydrophobic core consisting of five leucine/alanine repeats. Screening of hepatoma and testis libraries did not identify any type 2 clones, but a third 1475 bp ASM cDNA (type 3; pASM-3) was identified which lacked both the unique type 1 and 2 sequences. In order to determine the functional integrity of the three different ASM cDNAs, reconstruction of full-length coding sequences for the type 2 and 3 cDNAs has been done. Transient expression in *COS-1* cells of pASM-1FL and the reconstructed type 2 and 3 cDNAs demonstrated that only the type 1 cDNA was functional in this system (Schuchman *et al.*, 1991).

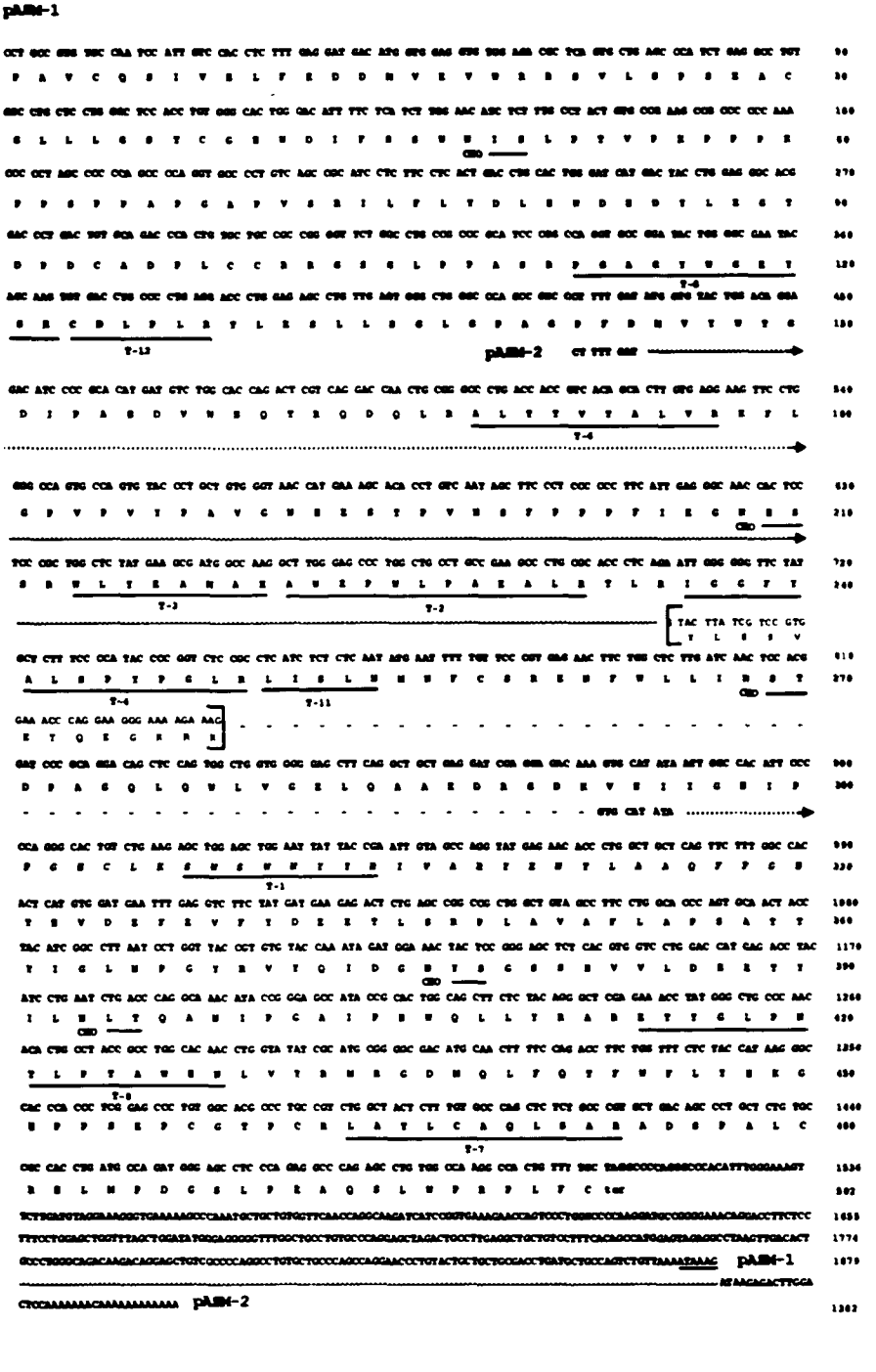
Table I **Tryptic peptide sequences of human ASM and corresponding oligonucleotide mixtures**

Tryptic peptide/Oligonucleotide	Sequence
T-1	-S - W - S - W - N - Y - Y - R -
Oligo 1 (20mer; 256 mix)	TGG ACN TGG AAT TAT TAT AG TG C C CC
T-2	-A - W - E - P - W - L - P - A - E - L - R-
Oligo 2A[a] (17mer; 64 mix)	CGN ACC CTT GGN ACC AA C G
Oligo 2B[a] (17mer; 64 mix)	ACC AAT GGN CGN CTT CG C C
Oligo 2C[a] (17mer; 128 mix)	ACC GAN GGN CGN CTT CG C
T-3	-(X)- L - Y - E - A - M - A - K -
Oligo 3 (17mer; 64 mix)	TAT GAA GCN ATG GCN AA C G
T-4	-I - G - G - F - Y - A - L - S - P - Y - P - G - L - R -
Oligo 4 (17mer; 192 mix)	ATT GGN GGN TTT TAT GC C A C C

[a]Probes constructed as complement strand.

Figure 1 **Nucleotide and amino acid sequences of pASM-1 and pASM-2.**

Collinear tryptic peptides are underlined; those residues which differed from the predicted amino acid sequence are double underlined. The dashed line indicates sequences which were identical in pASM-1 and pASM-2. Residues which were deleted in pASM-2 are indicated with an *.


pAM4-1
pAM4-2
pAM4-1
pAM4-2

Chapter Three

STRUCTURAL ORGANIZATION AND COMPLETE NUCLEOTIDE SEQUENCE OF THE HUMAN ASM GENE

Abstract

Using the full-length cDNA (pASM-1FL) (Schuchman *et al.*, 1991) as a probe, a genomic clone containing the entire ASM genomic region was isolated and the complete nucleotide sequence of the human ASM gene, including 1116 and 468 nucleotides upstream and downstream from the ASM coding region, respectively was determined. This housekeeping gene contained six exons ranging in size from 77 to 773 base pairs (bp) and five introns ranging in size from 153 to 1059 bp. Exon 2 was unusually large and encoded 258 amino acids, or about 44% of the mature ASM polypeptide. The alternatively spliced 172 bp type 1-specific sequence was encoded by exon 3, whereas the type 2-specific sequence was located at the 5′ end of intron 2. An analysis of the intron/exon junctions revealed that there was a weak donor splice site (AAA gtgagg) at the exon 3/intron 3 junction which occasionally leads to alternative splicing of exon 3 and the occurrence of the type 2 and 3 ASM transcripts. A single *Alu 1* element in the reverse orientation was in intron 2, immediately downstream from the type 2-specific sequence. The regulatory region upstream of the ASM coding sequence was GC rich and contained putative promoter elements including SP1, TATA, CAAT, NF-1 and AP-1 binding sites. Intriguingly, the ASM genomic region encoded three other long open reading frames (ORFs) which predicted polypeptides of 101, 104 and 158 amino acid residues, respectively. Analysis of genomic DNA from 110 unrelated individuals for a polymorphism identified at codon 506 (Gly or Arg), revealed that this polymorphism occurred with a frequency of 0.84 for the Gly allele and 0.16 for the Arg allele). The 506-Gly allelic frequency among 56 unrelated NPD Type A and B patients was 0.78.

Experimental Procedures

Genomic Library Screening

A human genomic library (average insert size ~10-15 kb) was constructed in the phage vector EMBL-3 (Promega) and kindly provided by Dr. Ruth Kornreich (Mount Sinai School of Medicine, NY). This library was screened at a density of ~10,000 plaques/150 mm petri dish using the full-length type 1 ASM cDNA, pASM-1FL (Schuchman *et al.*, 1991). Filter transfers and plaque hybridizations were performed by standard methods (Sambrook *et al.*, 1989). Random primed labeling of the cDNA probe was perfor[illegible]using [α-^{32}P] CTP (~3000 Ci/mmol;Amersham) and a random primed labeling kit according to the manufacturer's instructions (Amersham). Following three rounds of plaque purification, DNA was isolated from the positive clones by the plate lysate [illegible] (Sambrook *et al.*, 1989) and analyzed by Southern hybridization with oligonucleotides (17 mers) spanning the entire coding region of the full-length ASM cDNA. 5'- end labeling of the oligonucleotides was performed with T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P] ATP (>5000 Ci/mmol; Amersham). Oligonucleotide hybridizations were performed by standard methods (Sambrook *et al.*, 1989).

DNA Sequencing and Computer Analyses

Dideoxy sequencing was performed by the method of Sanger *et al.* (1977) using Sequenase kits (United States Biochemical). An ~8 kb *Sal*I/*Eco*RI restriction fragment

which contained the entire ASM coding region was isolated from the ASM genomic clone, pASMg-1, and digested with *Hinc*II (Promega) to generate four fragments of about 2.8, 2.0, 1.7 kb and 1.5 kb. The genomic restriction fragments were subcloned into Bluescript SK(+) (Stratagene) or pGEM7Z(+) (Promega) vectors, and sequenced in both orientations. Sequencing primers were synthesized on an Applied Biosystems DNA Synthesizer using phosphoramidite chemistry (Itakura *et al.*, 1984). Computer analyses were performed using the University of Wisconsin Genetics Computer Group DNA Sequence Analysis Software (version 7.0) and GeneBank (release 67) and Swiss-Prot (release 17) DNA and protein databases, respectively.

Detection of the 506 Polymorphism

To detect the 506 polymorphism, a 567 bp genomic fragment was amplified using sense (5′-AGTAGTCGACATGGGCAGGATGTGTGG-3′) and antisense (5′-AGTAGTGTCGACTTGCCTGGTTGAACCACAGC-3′) primers. Dot-blot hybridization was performed using allele-specific oligonucleotides for 506-Arg (5′-ACTACTCCAGGAGCTCT-3′) and for 506-Gly (5′-ACTACTCCGGGAGCTCT-3′), which were hybridized at 42° C and washed at 51 or 53° C, respectively. This polymorphism also can be detected by restriction enzyme analysis since the 506-Gly polymorphism creates a new *Msp*I restriction site. When the 567 bp PCR-amplified genomic fragment from a 506-Arg allele is digested with *Msp*I, two fragments of 395 and

159 bp are detected on a 1% agarose gel. In contrast, amplification of the 506-Gly allele results in the constant fragment of 395 bp, but the 159 bp fragment is digested into two fragments of 95 and 64 bp.

Results

Isolation of Human ASM Genomic Clones

Genomic clones encoding human ASM were isolated from a human genomic library using the full-length ASM cDNA, pASM-1FL (Schuchman *et al.*, 1991), as a probe. Nine putative positive plaques were purified and characterized by restriction enzyme and Southern hybridization analyses. The inserts ranged in size from about 12 to 20 kb and contained overlapping regions of the ASM gene. One clone, pASMg-1, contained an ~8 kb *Sal*I/*Eco*RI fragment which hybridized with oligonucleotides constructed from both the 5′ and 3′ ends of the full-length ASM cDNA. This genomic fragment was isolated and sequenced as described above.

Organization of the Human ASM Gene

Figure 2 shows the entire ASM genomic sequence (4,681 nt) including 210 bp upstream from the first in-frame ATG. This sequence was determined from the 1.5, 2.0 and 2.8 kb *Hinc*II fragments of pASMg-1. The 5′ end of exon 1 (nt -87 in *Figure 2*) was defined as the first nucleotide of the full-length ASM cDNA. An in-frame stop codon (TAA) was identified 15 bp upstream from this nucleotide. As shown schematically in

Figure 3, the ASM gene was composed of six exons ranging in size from 77 to 773 bp and five introns ranging in size from 153 to 1059 bp. Exon 2 was unusually large (773 bp) and encoded 258 amino acids, or ~44% of the mature ASM polypeptide. Note that exon 3 encoded the type 1 specific region, which is alternatively spliced in ~10% of the placental and fibroblast ASM transcripts. The type 2 specific sequence was found at the 5′ end of intron 2 (*Figure 2*).

When the nucleotide sequence of the full-length ASM cDNA and the genomic clone were compared, three differences were found within the coding region; Two of these, within codons 322 and 506 (e.g., a T or C at nucleotide 965 which predicted an Ile to Thr, and an A or G at nucleotide position 1516 which predicted an Arg or Gly) were previously found in the sequenced cDNA clones. The functional integrity of the full-length pASM-1FL sequence from placenta (e.g. 322-Ile and 506-Arg) was demonstrated by the transient expression of active enzyme in *COS-1* cells. The genomic sequence was similar to cDNA clone pASM-1FL in codon 322 and to cDNAa clone pASM-2 in codon 506. The third difference was a 6 bp deletion (CTGGTG) near the 5′ end of the ASM genomic sequence (nt 103 to 108 in the full-length ASM cDNA), which predicted a deletion of two amino acids (leucine and valine) within the putative ASM signal peptide. In addition, a difference in intron 2 was found between the sequence determined from the ASM genomic clone and a partial genomic sequence previously determined by PCR amplification and sequencing (Schuchman *et al.*, 1991). The length of the poly T tract in the genomic clone (beginning at nt 1881) was 30 nt, whereas in the sequence determined by PCR amplification the poly T tract was 23 nt. Other features of

this 1052 bp intron also deserve note. There are five potential lariat branch point sequences which fit the consensus sequence YNYURAY (Padgett *et al.*, 1986) located near the 3′ end of this intron. Only one of these potential branch points is followed by a polypyrimidine tract, however this branch point is 336 bp upstream from the 3′ acceptor site. In this intron, there was a 291 bp sequence which had more than 90% nucleotide identity with the *Alu* 1 consensus sequence (Britten *et al.*, 1988) (*Figure 4*); this *Alu* element was inserted in the reverse orientation relative to the ASM coding sequence. The 210 bp region upstream from the first in-frame ATG did not contain CAAT or TATA motifs, however the sequence was GC rich (~66%) and contained one SP-1 binding site (nt 150 to 155).

Intron/Exon Junctions

Each of the intron/exon junctions within the ASM gene had the GT/AG consensus donor/acceptor splice site sequences (Table II). However, within donor splice site D3 there was an A at position -1 (underlined in Table II), which is conserved as G in >90% of mammalian donor splice sites (Mount, 1982). Notably, this donor splice site was at the end of exon 3, which encoded the 172 bp type 1 specific region. This exon was alternatively spliced in about 10% of the ASM transcripts. At the 3′ end of the type 2 specific region within intron 2 (see *Figure 2*) there was a sequence, *aag gtgaat*, which may serve as a cryptic donor splice site. Thus, the occurrence of the type 2 specific transcripts may be explained by alternative splicing of exon 3 (due to the weak D3 donor splice site) and use of the cryptic splice site adjacent to the type 2 specific sequence.

Analysis of the ASM Promoter Region

Since no TATA or CAAT promoter elements were identified in the 210 bp upstream of the first in-frame ASM initiation codon, further DNA sequencing of the pASMg-1 genomic insert was performed. *Figure 5* shows an additional 906 nt of upstream sequence. Putative promoter elements were identified, including four SP-1 binding sites (nt 256 to 261, 267 to 273, 285 to 290 and 715 to 720), two TATA boxes (nt 863 to 868 and 894 to 898), two CAAT boxes (nt 704 to 709 and 867 to 872), one AP1 site (nts 604 to 609) and two NF-1 sites (nts 461 to 466 and 583 to 587). Overall, this region was GC rich (~63%), suggesting that it was a component of an HTF island.

ORFs in the Human ASM Gene

In addition to sequences coding for the ASM polypeptide, three other open reading frames (ORFs) were identified in this genomic sequence which may encode functional proteins (*Figure 6*). The predicted polypeptides contained 101, 104 and amino acid residues, respectively. The transcriptional orientations of ORF 1 (*Figure 6*, nt 176 to 485) and ORF 2 (nt 753 to 1067) were opposite those of ASM and the predicted proteins shared no homology with ASM or any other proteins in the Swiss-Prot protein database. In contrast, ORF 3 (nt 2517 to 2998) was in the same transcriptional orientation and coding phase as the ASM gene. This ORF began within intron 2, overlapped ASM exon 3, and extended into intron 3.

MspI Polymorphism

Analysis of genomic DNA from 110 unrelated individuals for codon 506 revealed that the base differences in codon 506 occurred in the Caucasian population as polymorphism (0.84 for the 506-Gly allele). The 506-Gly allelic frequency among 56 unrelated NPD Type A and B patients was 0.78.

Discussion

The genomic organization and complete nucleotide sequence of the gene encoding human ASM are described in this chapter. This housekeeping gene is small (i.e. ~ 5 kb) and the coding region is divided into six exons. Analysis of the genomic sequence documented the occurrence of alternative splicing at the ASM locus and further clarified the molecular mechanisms underlying these alternative transcripts. The type 1 specific 172 bp sequence was encoded by exon 3, whereas the type 2 specific 40 bp sequence was located at the 5′ end of intron 2, followed by a potential cryptic donor splice site. Furthermore, there was a poor donor splice site (D3; AAA gtgagg) at the exon 3/intron 3 junction. Thus, the occurrence of the type 2 and 3 ASM transcripts (see Chapter Two) resulted from the fact that in about 10% of the ASM transcripts the donor site D3 was not functional and splicing proceeded to either the cryptic donor splice site (indicated by the overline in *Figure 7*), or to donor site D2. The G to A transition of the nucleotide immediately adjacent to the invariant GT consensus dinucleotide in D3 (underlined in Table II) may cause these alternative splicing events, since this alteration previously was

shown to be the cause of abnormal splicing in the proα1(I) collagen gene, resulting in Ehlers-Danlos syndrome type VII (Weil *et al.*, 1989). In fact, the cryptic splice site D3 more closely matches the consensus sequence than either of the two identified donor splice sites D1 or D2. The presence of two poly (T) tracts at the 5′ end of intron 2 may cause abnormal RNA secondary structure, perhaps favorably positioning the cryptic splice site, D3. Only one of the potential branch points, found in intron 2, is followed by a polypyrimidine tract, however this branch point is relatively far from the 3′ acceptor site. It is generally assumed that the branch point should be within 20 to 50 nt of the 3′ acceptor and not closer than 70 nt to the 5′ donor site. Therefore, perhaps after cleavage at the D3 cryptic donor splice site, a lariat cannot efficiently form at branch points within this intron and thus the lariat occurs at the next available branch point, which is located in intron 3. Alternative splicing also occurs in the transcripts for two other human lysosomal enzymes, β-glucuronidase (Oshima *et al.*, 1987) and βgalactosidase (Morreau *et al.*, 1989). β-glucuronidase cDNAs, which had an internal 153 bp deletion, were identified in human fibroblast and placenta cDNA libraries. The shorter cDNA had lost an entire exon due to alternative splicing and expressed an enzyme protein that was not catalytically active in *COS-1* cells. For β-galactosidase, two distinct groups of cDNA clones were isolated from human fibroblast cDNA libraries. The shorter cDNAs were missing two noncontiguous protein coding regions present in the full-length cDNAs and produced truncated proteins which were targeted to the perinuclear region in *COS-1* cells. Interestingly, the exclusion of the first region introduced a frame shift which was only restored by the subsequent exclusion of the second region. Genomic sequencing revealed

that these two types of cDNAs were derived from a single gene by alternative splicing.

A single *Alu* 1 element was found within the ASM genomic region. This *Alu* 1 element may be placed into the "a branch" according to the classification of Jurka and Smith (1988), indicating the ancestral nature of the ASM gene. In addition, three other long ORFs were identified within the ASM genomic region. Although it is not known whether these genomic sequences are transcribed into functional RNAs, there is some precedent for overlapping transcriptional units within lysosomal enzyme genes. For example, within the first intron of the murine β-glucuronidase structural gene there is an RNA polymerase II promoter motif which drives transcription of a ~2.2 kb liver transcript which snares little homology with β-glucuronidase (Wang *et al.*, 1988). To date, the function of this transcript remains unknown.

A number of putative regulatory elements were identified within the upstream ~1 kb of the ASM gene. This region was GC rich and contained five SP1 binding sites. In addition, TATA, CAAT, AP1 and NF-1 binding sites were identified. Since the precise site of transcription initiation has not been determined for ASM, no conclusions can be drawn about the functional relevance of these sequences. However, the fact that these regulatory sequences are within 1 kb of the ASM coding region suggests that they comprise all or part of the ASM promoter. This is supported by the fact that transgenic mice containing the human ASM genomic region, including about 1.5 kb of upstream sequences, express human ASM activity at high levels (Schuchman *et al.*, unpublished results). Clearly, further studies (e.g., *in vitro* mutagenesis and expression experiments) are required to definitively map the ASM control region and determine the significance

of these putative regulatory sequences.

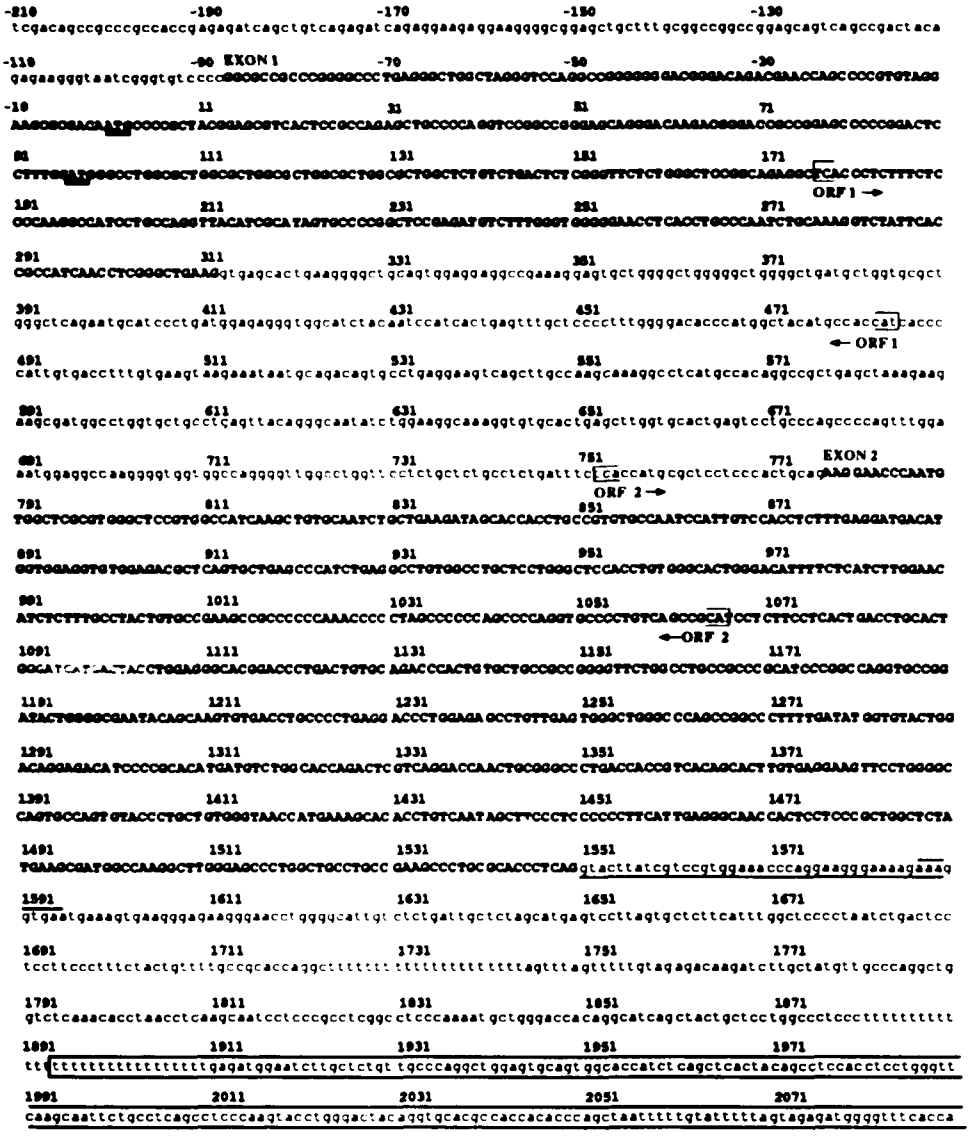
ASM is the eighth human lysosomal gene for which the genomic organization has been determined and the fourth to be completely sequenced (Table III). In addition to ASM, sequences encoding α-*N*-acetylgalactosaminidase (Wang and Desnick, 1991), acid phosphatase (Geier *et al.*, 1989), α-galactosidase A (Kornreich *et al.*, 1989), β-glucosidase (Reiner *et al.*, 1988; Horowitz *et al.*, 1989), β-glucuronidase (Miller *et al.*, 1990; Shipley *et al.*, 1991), and the α and β chains of β-hexosaminidase (Proia and Soravia, 1987; Proia, 1988), also have been reported. Of these, nucleotide sequences of the promoter regions are available for seven (Table III). Aside from the fact that all of these upstream sequences are GC rich, indicating that they may be components of HTF islands, analysis of these regions has not provided any consensus sequence for a lysosomal gene-specific promoter element. Although the promoter regions of many housekeeping genes are GC rich and lack TATA and/or CAAT motifs, the genes encoding α-galactosidase A, β-hexosaminidase α-chain, β-glucosidase and ASM contained these consensus sequences.

To date, mutagenesis and expression studies have been performed for three lysosomal gene promoter regions, β-glucosidase (Horowitz *et al.*, 1989), β-glucuronidase (Shipley *et al.*, 1991) and acid phosphatase (Geier *et al.*, 1989). In the human β-glucosidase gene, a 650 bp genomic fragment containing the putative control region (including 2 TATA and 2 CAAT motifs) was inserted upstream from the bacterial chloramphenicol acetyltransferase (CAT) gene and transfected into various human cells. The functional integrity of this regulatory region was demonstrated and, surprisingly, tissue-specific expression was observed. For human β-glucuronidase, deletion analysis

of minigene constructs demonstrated that the 200 bp of sequence upstream from the translation initiation site was sufficient for maximal expression in *COS* cells. This region was GC rich, but did not contain TATA or CAAT elements. For human acid phosphatase, a 590 bp upstream region which was GC rich and lacked a TATA element was shown to possess promoter activity by expression analysis of CAT constructs in *COS* cells.

Figure 2 **Nucleotide sequence of the gene encoding human ASM.**

Exonic sequences are shown in bold upper case letters. Intronic sequences are in lower case. The two putative ASM initiation codons and the first in-frame stop (TAG) codon are indicated by a double underline. A single *Alu 1* sequence within intron 2 is boxed and its transcriptional orientation is indicated by an arrow. A polyadenylation site is underlined. The putative initiation codons of the three additional ORFs (ORF 1, ORF 2, and ORF 3) also are underlined and their transcriptional directions are indicated by arrows. The type 1 specific sequence is encoded by exon 3. The type 2 specific sequence immediately follows exon 2 and is underlined. A cryptic donor splice site adjacent to the type 2-specific sequence is overlined.


EXON 1
ORF 1
ORF 1
EXON 2
ORF 2
ORF 2

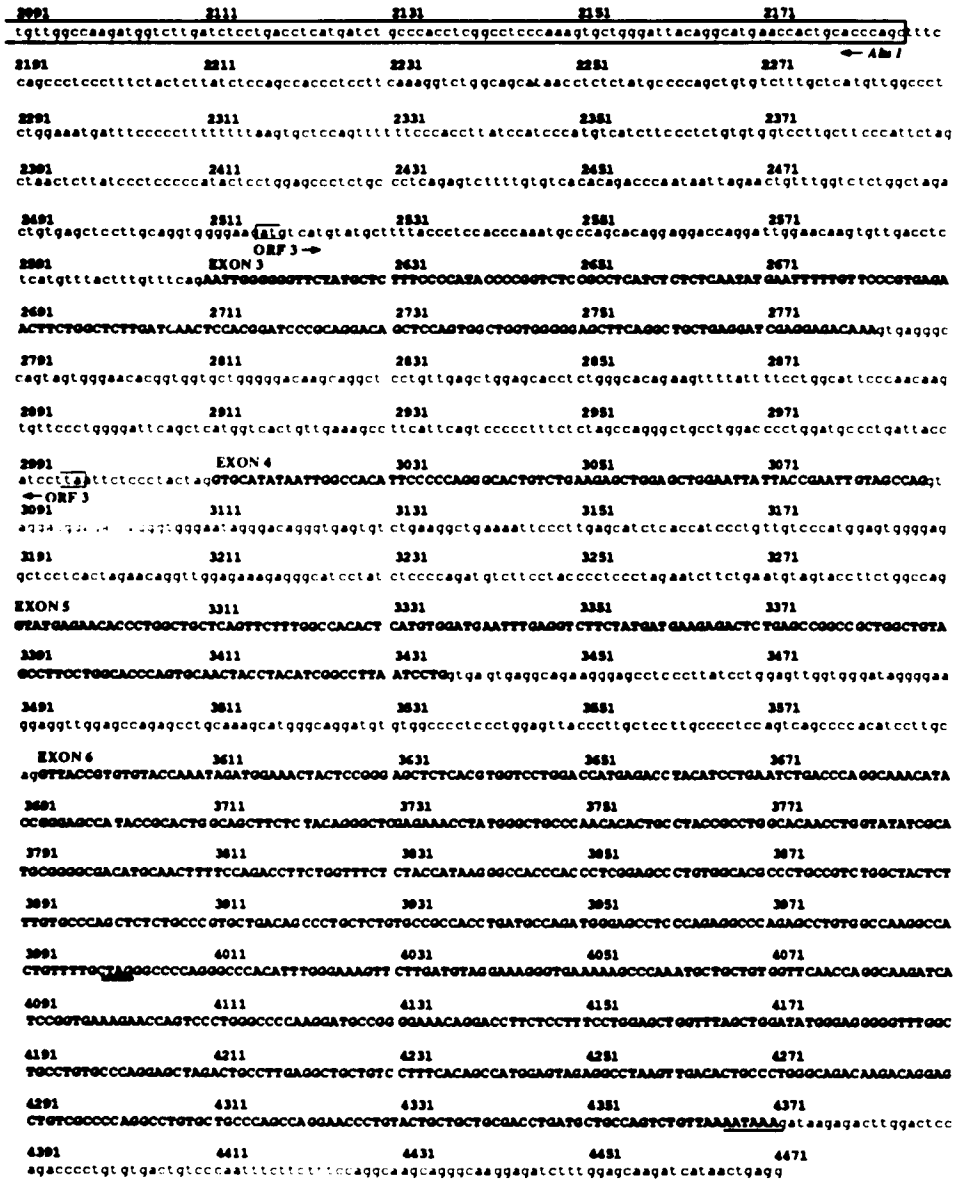

ORF 3
EXON 3
EXON 4
ORF 3
EXON 5
EXON 6

Table II Intron/exon junctions in the ASM gene

Exon Number	Exon Size (nt)	Codons	5' Donor Splice Site	Intron Size (nt)	3' Acceptor Splice Site
1	398	1-104	AG gtgagc (D1)	464	cag A (A1)
2	773	105-362	AG gtactt (D2)	1059	cag A (A2)
3	171	362-419	AA gtgagg (D3)	228	tag G (A3)
4	77	420-446	AG gtagga (D4)	201	cag G (A4)
5	145	446-494	TG gtgagt (D5)	153	cag G (A5)
6	778	495-630			

a The underlined residues represent divergences from the consensus sequences

Figure 3 **Schematic representation of the human ASM gene.**

The solid black boxes and straight lines represent the exons and introns, respectively. The two putative initiation codons and the first in-frame stop codons also are indicated, as is the location of the *Alu 1* element and its transcriptional orientation. The location of ORFs 1, 2, and 3 are indicated and their transcriptional orientations are shown by arrows.

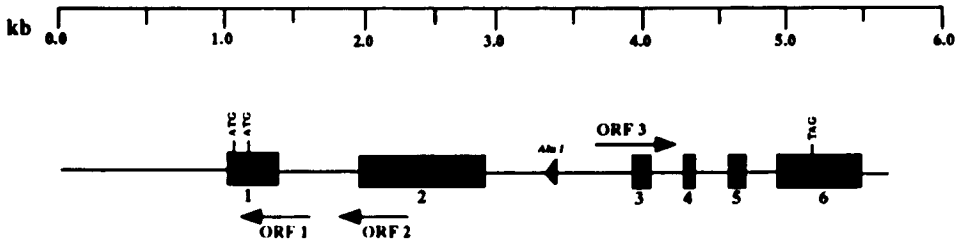

kb
0.0
1.0
2.0
3.0
4.0
5.0
6.0
ATG
ATG
TAG
ORF 3
1
2
3
4
5
6
ORF 1
ORF 2

Figure 4 ***Alu 1* sequence within intron 2 of the human ASM gene.**

Bestfit alignment of the *Alu 1* element within intron 2 of the ASM gene (indicated on the bottom in bold capital letters) and the *Alu 1* consensus sequence on top. Vertical lines between the two sequences represent nucleotide identity.

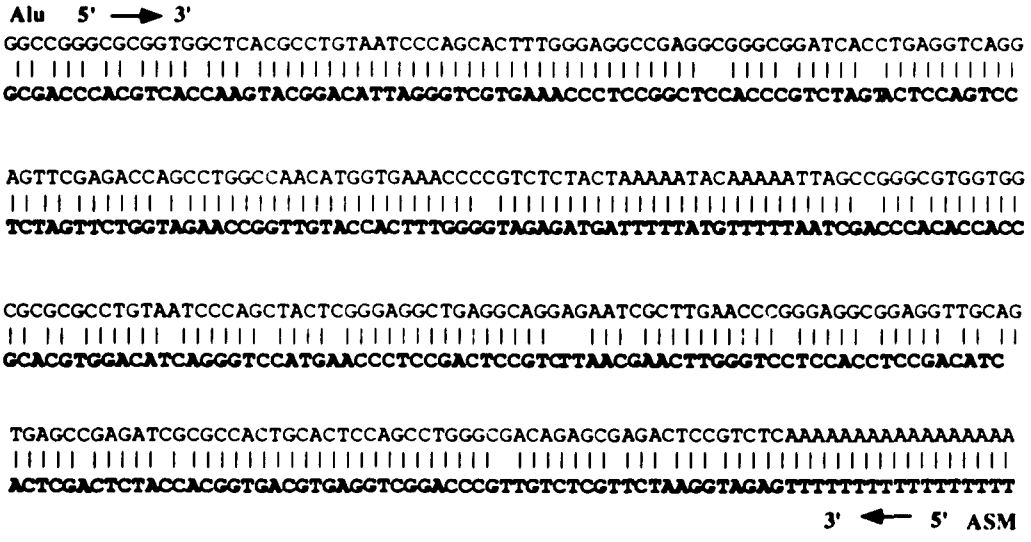

Alu 5' → 3'
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATCACCTGAGGTCAGG
GCGACCCACGTCACCAAGTACGGACATTAGGGTCGTGAAACCCTCCGGCTCCACCCGTCTAGTACTCCAGTCC
AGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGG
TCTAGTTCTGGTAGAACCGGTTGTACCACTTTGGGGTAGAGATGATTTTTATGTTTTTAATCGACCCACACCACC
CGCGCGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCGGAGGTTGCAG
GCACGTGGACATCAGGGTCCATGAACCCTCCGACTCCGTCTTAACGAACTTGGGTCCTCCACCTCCGACATC
TGAGCCGAGATCGCGCCACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAA
ACTCGACTCTACCACGGTGACGTGAGGTCGGACCCGTTGTCTCGTTCTAAGGTAGAGTTTTTTTTTTTTTTTTT
3' ← 5' ASM

Figure 5 **Putative promoter region of the human ASM gene.**

Sp1 (dashed arrow), TATA, CAAT, NF-1 and Ap-1 (underlined) binding sites are indicated. Where appropriate, the orientation of the consensus sequence is indicated by arrows.

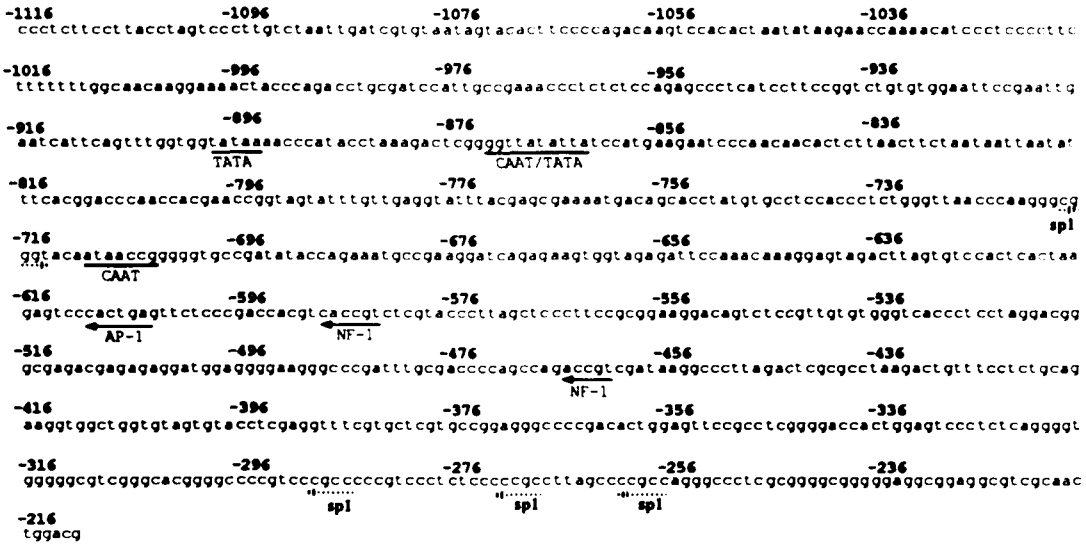
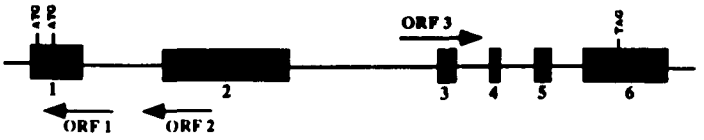

-1116
-1096
-1076
-1056
-1036
-1016
-996
-976
-956
-936
-916
-896
-876
-856
-836
TATA
CAAT/TATA
-816
-796
-776
-756
-736
Sp1
-716
-696
-676
-656
-636
CAAT
-616
-596
-576
-556
-536
AP-1
NF-1
-516
-496
-476
-456
-436
NF-1
-416
-396
-376
-356
-336
-316
-296
-276
-256
-236
Sp1
Sp1
Sp1
-216

Figure 6 **Other open reading frames in the ASM gene.**

The predicted amino acid sequence of the three ORFs also is shown; the numbers in parentheses indicate the length of the predicted polypeptide.



ORF 1 (101 residues):

MVACSHGCPQRGANSVMDCRCHPLHQGCILSPAHQHQPQPPAPALLSASTA

APSVLTFSPRLMAVNRPLQIGQVRFPHPKTSRSRGTMRCNLAGWPWGERG*

ORF 2 (104 residues):

MRLTGAPGAGGLGGLGGGFGTVGKEMFQDEKMSQCPQVEPRSRPQASDGLS

TERLHTSTMSSSKRWTMDWHTAGGAIFSRLHSLMATEPTRATLGSFCSGRS

AW*

ORF 3 (158 residues):

MYAFTLHPNAQHRRTRIGTSVDLSCLLCFRIGGFYALSPYPGLRLISLNMN

FCSRENFWLLINSTDPAGQIQWIVGELQAAEDRGDKVRASSGNTVVLGDKQ

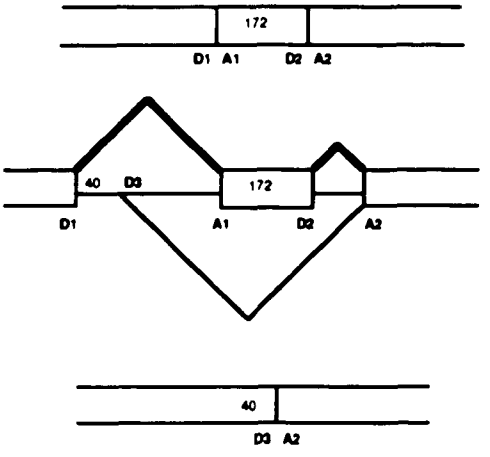
APVELEHLWAQKFYFPGIPNKCSLGIQLMVIVESLHSVPLSLARAAWTPGC

PDYHP*

Figure 7 **Scheme of alternative splicing of ASM transcripts**

The type 1 specific 172 bp sequence is encoded by exon 3, whereas the type 2 specific 40 bp sequence is located at the 5′ end of intron 2, followed by a potential cryptic donor splice site (D3). The occurrence of the type 2 specific transcripts may be explained by alternative splicing of exon 3 and use of the cryptic splice site D3.

Type 1



Type 2

Table III Putative promoter regions of genes encoding lysosomal enzymes

Gene	Transcription Start Site(s)	Nucleotides Analyzed Upstream From ATG	Percent GC	TATA	CAAT	Sp-1	Other	
α-N-acetylgalactosaminidase	-347[a]	1400	56	none	-678 -644	-364 -410	IR	-354 -445
acid phosphatase	~ -6 to -23	590	59	none	-161 -394 -456	-60 -326		
acid sphingomyelinase	nd	1117	63	-863 -894	-704 - 867	-150 -256 -267 -285 -715	AP1 NF1	-604 -461 -583
α-galactosidase A	-60	1179	59	-86 -93 -102 -129	-71 -104 -146 -178 -203	-63 -207	AP1 OCTA *c-fos* Enhancer Chorion Enhancer	-153 -835 -889 +70 +78 -274 -290 -307 -323
β-glucosidase	~ -657	1011		-680 -691	-754 -760	none		
β-glucuronidase	-30 -126	200	72	none	none	-248 -64 +26	AP2	-164
β-hexosaminidase α-chain	nd	480	66	-247	-287	nd		

[a] Transcriptional start site for the 2.2 kb transcript; start site for the 3.6 kb transcript unknown

Chapter Four

A FREQUENT MISSENSE MUTATION (R496L) IN THE ASM GENE OF ASHKENAZI JEWISH TYPE A AND B PATIENTS

Abstract

The recent isolation and characterization of cDNA and genomic sequences encoding ASM has facilitated investigation of the molecular lesions causing the NPD subtypes. Total RNA was reverse-transcribed and the ASM cDNA from an Ashkenazi Jewish Type A patient was specifically amplified by the polymerase chain reaction (PCR). Molecular analysis of the PCR products revealed a G to T transversion of nucleotide (nt) 1487 which occurred at a CpG dinucleotide and predicted an arginine to leucine substitution in residue 496 (i.e., R496L). Hybridization of PCR-amplified genomic DNA with allele-specific oligonucleotides (ASOs) indicated that the proband was homoallelic for the R496L mutation and that both parents and several other relatives were heterozygous. This mutation was detected in 32% (11 of 34) of the Ashkenazi Jewish NPD Type A alleles studied and occurred in only 5% (2 of 43) of ASM alleles from non-Jewish Type A patients. Of interest, the R496L mutation occurred in one of the ASM alleles from the two Ashkenazi Jewish NPD Type B patients studied and in none of the ASM alleles of 15 non-Jewish Type B patients. In contrast, the mutation was not present in the 180 ASM alleles from normal individuals of Ashkenazi Jewish descent. These findings identify a frequent missense mutation among NPD patients of Ashkenazi Jewish ancestry that results in neuronopathic Type A disease when homoallelic, or when heteroallelic, can result in the non-neuronopathic Type B phenotype.

Introduction

Mutations found in the human genome range from point mutations to large rearrangements. In the case of ASM gene the vast majority of the mutations are point mutations since no gross alterations have been found on Southern analyses of genomic DNA from NPD patients (unpublished results), and all of the patients studied to date were found to produce mRNA (Ferlinz *et al.*, 1991 and this work). In addition, most patients have residual ASM activity and some were shown to have immunologically normal enzyme (see Chapter One). Since the ASM gene was found to be relatively small (see Chapter Three) it seemed practical to use the PCR-based sequencing method for mutation detection. In most cases, PCR was performed on reverse-transcribed total RNA and/or genomic DNA, and the PCR products were subcloned into a sequencing vector. One of the concerns of this method is the accuracy of the DNA polymerase, and in order to overcome this problem a few independent amplifications were performed, and multiple clones were sequenced to exclude experimental artifacts.

This chapter describes the identification of the first mutation in the gene encoding ASM causing NPD.

Experimental Procedures

Cell Lines

Primary cultures of fibroblasts and lymphoblasts were established from skin biopsies and peripheral blood samples obtained from NPD patients and family members and from normal individuals, with informed consent. NPD cell lines GM00112A, GM00165, GM00370, GM00406, GM00559, GM02895, and GM03252 were obtained from the NIGMS Human Genetic Mutant Cell Repository Institute for Medical Research (Camden, NJ). Cell lines 444X.F01, 534R.F03, 556X.F01, 888V.F01, 2789X.F01, 4293Q.E02, 4774Z.F01, 5113C.L01, 5115E.F01 and 6791M.F01 were obtained from the Service de Biochimie, Hospice de Lyon (Lyon, France). Cell lines DMN 83.126, DMN 84.135, DMN 86.40, DMN 86.49, DMN 87.71, DMN 87.99, DMN 88.9, DMN 83.133, GJO, and RNS were provided by Dr. Peter Penchev, Developmental and Metabolic Neurology Branch, National Institute of Neurological and Communicative Disorders and Stroke. Cell line SL was provided by Dr. Gideon Bach (Hebrew University-Hadassah Medical School, Jerusalem). The cells were grown in RPMI 1640 media supplemented with 10% fetal bovine serum, 1% penicillin and 1 mg/ml streptomycin by standard procedures. The diagnosis of Types A and B NPD was based on clinical criteria (e.g., age at onset, presence of neurologic involvement, etc.) and by the demonstration of markedly deficient ASM activity in cultured cells (Klar *et al.*, 1988).

Enzyme and Protein Assays

The *in vitro* ASM activities were determined in cultured fibroblasts obtained from NPD patients and normal individuals using the fluorescent natural substrate, *N*-[11-(1-pyrene) sulphonyl-aminoundecanyl] sphingomyelin (PSA11-SPM) as previously described (Klar *et al.*, 1988). One unit (U) of activity equals that amount of enzyme that hydrolyzes one nanomole of substrate per hour. Protein determinations were performed by a modified fluorescamine assay (Bishop and Desnick, 1981).

cDNA and Genomic Amplification and Sequencing of the Mutant Allele

Total RNA and genomic DNA were isolated from cultured skin fibroblasts by standard procedures (Sambrook *et al.*, 1989). First-strand cDNA was reverse transcribed from ~5 μg of total RNA using a cDNA synthesis kit according to the manufacturer's instructions (Boheringer Mannheim). The cDNA (~10% of the total reaction) or genomic DNA (~0.5 μg) was amplified by the polymerase chain reaction (PCR) with *Thermus aquaticus (Taq)* polymerase (Cetus) essentially as described by Saiki *et al.* (1988) with the following conditions and modifications: PCR was performed for 30 to 40 cycles and consisted of denaturation for 1 min at 94° C and annealing and extension for 4 min at 66°C. In order to improve the specificity of the PCR amplification, a "PCR boost" procedure was used. In this procedure the buffer contains 67 mM Tris-HCl, 16.6 mM $(NH_4)SO_4$, 3 mM $MgCl_2$, 10 mM β-mercaptoethanol, 200 μM dNTP's and 170 μg/ml BSA. The concentrations of the primers and *Taq* polymerase were 0.1 μM and 0.5 U/100 μl, respectively for the first 15 cycles. Then each primer was added to a final

concentration of 0.5 uM and an additional 2 U of *Taq* polymerase were added to the reaction mixture. PCR-amplification was then performed for an additional 15-25 cycles. Pairs of sense and antisense oligonucleotide primers were synthesized on an Applied Biosystems Model 380B DNA Synthesizer (Itakura *et al.*, 1984) and used to specifically amplify the entire coding region of the reverse-transcribed type 1 ASM transcript in three overlapping cDNA fragments, the 1665 bp genomic region containing the alternatively spliced sequences in the type 1 and 2 ASM cDNAs, and a genomic region which included the point mutation for confirmation of the candidate missense mutation. To amplify a 984 bp fragment from the 5′ end of the ASM cDNA, the 29-mer sense primer, P1, (5′-AGTAGTCTCGAGACGGGACAGACGAACCA-3′) corresponded to ASM nucleotide 39 to 23 with an additional 12 nt which included an *Xho*I restriction site, and the 31-mer antisense primer, P2, (5′-AGTAGTCTGCAGAGCAGGGTACATGGCACTG-3′) corresponded to ASM nt 926 to 945 with an additional 12 nt containing an *Eco*RI restriction site. To amplify an internal 383 bp fragment of the ASM cDNA, the 29-mer sense primer, P3, (5′-ATCATCAAGCTTGGGTAACCATGAAAGCA-3′), corresponded to ASM nt 947 to 964 with an additional 12 nt containing a *Hind*III restriction site, and the antisense 32-mer primer, P4, (5′-ATCATCGAATTCTACAATTCGGTAATAATTCC -3′) corresponded to ASM nt 1310 to 1330 with an additional 12 nt containing an *Eco*RI restriction site. To amplify a 789 bp 3′ fragment from the ASM cDNA, a 19-mer sense primer, P5, (5′-CTCCACGGATCCCGCAGGA-3′) corresponded to ASM nt 1185 to 1203 and containing an internal *Bam*HI restriction site, and an antisense 32-mer primer, P6, (5′-AGTAGTGTCGACTTGCCTGGTTGAACCACAGC-3′) corresponded to ASM nt 1955

to 1974 with an additional 12 nt containing a *SalI* restriction site. Primers P3 and P4 also were used to amplify the 1665 bp internal genomic region which contains the alternatively spliced type 1 and 2 cDNA sequences. In order to confirm the candidate mutation by genomic sequencing and dot-blot analysis (see below), a 27-mer sense primer, P7, (5'-AGTAGTCGACATGGGCAGGATGTGTGG-3') was used with antisense primer P6 to amplify a 567 bp genomic fragment containing the G to T transversion. Following PCR amplification, the PCR products were isolated from agarose gels and subcloned into either Bluescript KS (+) (Stratagene) or pGEM 9Zf (-) (Promega) vectors. For each amplified product, from four to six independent subclones were sequenced in both orientations by the dideoxy method (Sanger *et al.*, 1977).

Dot-Blot Analysis

Any nucleotide change which occurred in more than two subcloned PCR products was considered a candidate mutation and analyzed by dot-blot hybridization with allele-specific oligonucleotides (ASOs). In addition, ASOs were constructed and used as probes to analyze amplified genomic DNA from normal individuals and NPD patients and family members. For these studies genomic DNA was rapidly isolated from either whole blood or cultured cells by the following procedure. About 0.5 ml of whole blood and 0.5 ml of lysis buffer (10 mM Tris/HCl buffer, pH 7.5, containing 5 mM $MgCl_2$, 0.32 M sucrose and 1% Triton X-100) were mixed at room temperature. Following centrifugation at 13,000 × g, the supernatant was removed and 0.5 ml of PCR buffer (10 mM Tris/HCl buffer, pH 8.3, containing 50 mM KCl, 2.5 mM $MgCl_2$, 0.1 mg/ml gelatin, 0.45% Nonidet

P40, 0.45% Tween 20, and 0.1 mg/ml Proteinase K) was added. For cultured cells, the lysis step was omitted and the washed cell pellets were resuspended directly in PCR buffer (~5 x 10^6 cells/ml). The samples were then incubated at 65° C for 1 hr, boiled for 10 min to inactivate the protease and 10-25 µl was removed for PCR amplification.

For the analysis of the R496L mutation, the 567 bp PCR product amplified from genomic DNA was analyzed by dot-blot hybridization (Sambrook *et al.*, 1989) using Zetabind nylon membranes (AMF-Cuno, Meriden, CT) and a Biorad dot-blot apparatus. Hybridizations were performed for at least 3 hr at 39° C. Following hybridization, the blots were washed at room temperature for 15 min in 6×SSC containing 0.1% SDS, and then for 1 hr in the same solution at either 53°C for the normal (5′-CTATTTGGTACACACGG-3′) or 48°C for the mutation-specific (5′-CTATTTGGTACACAAGG-3′) oligonucleotides.

Results

Identification of the R496L Mutation in an Ashkenazi Jewish Type A NPD Patient

To determine the molecular lesions in an Ashkenazi Jewish Type A NPD patient (proband 1), who had about 1% of normal ASM activity in cultured fibroblasts, total RNA was isolated from cultured lymphoblasts, reverse-transcribed into cDNA and the entire coding region was PCR-amplified. Nucleotide sequencing of the subcloned PCR products revealed a single point mutation in a CpG dinucleotide, a G to T transversion of nt 1487 of the full-length cDNA (*Figure 8*). This transversion predicted an arginine to leucine

substitution in residue 496 of the ASM polypeptide. All other base substitutions occurred in only one or two of the subcloned PCR products analyzed.

To confirm the authenticity of this candidate mutation, a 567 bp region of genomic DNA from the proband, various family members and 90 normal Ashkenazi Jewish individuals was PCR-amplified and then hybridized with normal and R496L-specific radiolabeled oligonucleotides. As shown in *Figure 9*, the PCR-amplified genomic DNA from proband 1 hybridized to the mutation-specific, but not to the normal ASO, confirming the authenticity of the transversion and indicating that proband 1 was homoallelic for the R496L mutation. Of the nine other family members studied, both parents, the paternal grandfather and a paternal aunt were heterozygous for the R496L mutation. The maternal grandmother did not have the mutation, suggesting that the maternal grandfather was heterozygous for this mutation. There was no known consanguinity between the maternal or paternal grandfathers whose ancestors were from different European countries. Moreover, the R496L mutation was not found in 180 ASM alleles studied from a group of normal Ashkenazi Jewish individuals, indicating that the G to T transversion was not a common polymorphism.

Occurrence of the R496L Mutation in Ashkenazi Jewish and Non-Jewish NPD Type A Families

The occurrence of the R496L mutation in other NPD families was determined by dot-blot analysis of genomic DNA from 17 unrelated Ashkenazi Jewish and 22 unrelated non-Jewish Type A NPD families. As indicated in Table IV, of the 34 Ashkenazi Jewish

NPD alleles studied (i.e., from 6 unrelated patients and 22 unrelated obligate heterozygotes), 32% had the R496L mutation. Proband 1 was homoallelic for the mutation, as well as another Type A patient, while seven obligate heterozygotes from unrelated families, in which material from patients was unavailable, were heterozygous for the R496L mutation. In contrast, only 2 of 43 (4.7%) alleles had the R496L mutation in non-Jewish NPD Type A patient, an American of German ancestry who was homoallelic.

Occurrence of the R496L Mutation in Ashkenazi Jewish and Non-Jewish NPD Type B Families

Analysis of genomic DNA from two unrelated Ashkenazi Jewish NPD Type B patients revealed the presence of one R496L allele in one patient (designated proband 3). In contrast, the R496L allele was not found in genomic DNAs from 16 non-Jewish NPD Type B patients (Table IV).

Discussion

Insights into the molecular nature of the remarkably distinct Type A and B NPD phenotypes have been gained by the identification of the first mutation in the ASM gene causing this lysosomal storage disease. The G to T transversion of coding nucleotide 1487 occurred at a CpG dinucleotide, a known hotspot for point mutations (Coulondre *et al.*, 1987), and predicted an arginine to leucine substitution in residue 496 of the ASM

polypeptide. Homoallelism for the R496L mutation resulted in the severe neuronopathic Type A phenotype, as evidenced by proband 1 who had ~1% of normal ASM activity. It is not known whether the substitution of the basic arginine for the more hydrophobic and neutral leucine residue altered the enzyme polypeptide's catalytic activity, stability, or both, since monospecific anti-human ASM antibodies useful for immunoblotting are not currently available.

Of the 17 unrelated Ashkenazi Jewish Type A families studied, nine were either homoallelic or heteroallelic for this lesion. In this sample, the frequency of the R496L allele was 32%, indicating that this lesion is an important mutation in Type A NPD among Ashkenazi Jewish patients. In contrast, analysis of 18 unrelated non-Jewish Type A patients revealed the presence of the R496L allele in only one (a frequency of 4.7%). The occurrence of the R496L allele in this individual may have resulted from an independent mutational event or the presence of Jewish ancestors in the non-Jewish family.

One of the two Ashkenazi Jewish Type B NPD patients was heteroallelic for the R496L mutation. The other allele in this Jewish Type B patient had a different ASM mutation which presumably resulted in the synthesis of a partially functional ASM polypeptide, as this patient had about 5% residual ASM activity in cultured fibroblasts (see Chapter Six). The fact that none of the 16 non-Jewish Type B patients had the R496L mutation suggests that this allele is extremely rare in Type B disease outside of the Ashkenazi population.

For the past three decades, the genetic mechanisms responsible for the high frequency of the mutations that cause Tay-Sachs disease, Gaucher disease and NPD in the Ashkenazi Jewish population (gene frequencies of about 0.02, 0.02 and 0.005, respectively) have been the subject of interest and debate (Knudson *et al.*, 1962; Chase and McKusick, 1972; Myrianthopoulos *et al.*, 1972a; Myrianthopoulos *et al.*, 1977 and Fraikor *et al.*, 1977). Intrigued by the fact that all three of these disorders are lysosomal diseases resulting from enzymatic defects in the sphingolipid degradative pathway, investigators suggested that there may have been a common selective pressure for their high gene frequencies in the Ashkenazi Jewish population. Others argued that the higher gene frequencies in Ashkenazi Jewish individuals could be due to higher mutation rates for these genes or founder effect and genetic drift. The recent identification of the mutations causing these three diseases in the Ashkenazi Jewish population has provided insight into this controversy. To date, three mutations in the β-hexosaminidase α-chain have been identified as the cause of Tay-Sachs disease in almost all Ashkenazi Jewish patients. Two of these mutations result in the infantile form, a four bp insertion (Myerowitz and Costigan, 1988) or a splice site mutation (Myerowitz, 1988), which account for ~80% and 20% of the mutant alleles, respectively. Affected Ashkenazi Jewish patients with the less frequent and milder chronic or adult-onset form all have been heteroallelic for a point mutation G269S (Navon and Proia, 1989) and one of the two infantile-onset alleles. Type 1 Gaucher disease among Ashkenazi Jewish patients results from multiple mutations in the acid β-glucosidase gene, the N370S allele occurring in about 75% of the mutant alleles (Tsuji *et al.*, 1987), while the other 25% include

several other lesions (Beutler *et al.*, 1991). It appears that in the Ashkenazi Jewish population each of these sphingolipidoses results from common mutations. The fact that two or more mutant alleles in each gene occur frequently in this population argues for selection, rather than for a higher mutation rate or founder effect and genetic drift, as the major mechanism responsible for their increased frequency. Although it is likely that the major mutation for each disease first became established in the Ashkenazi Jewish population by founder effect and genetic drift, the finding of two or more mutations in each of these genes supports a selective advantage. Since all three disorders involve defects in lysosomal enzymes that degrade sphingolipids, it is tempting to suggest that a common selective agent, such as resistance to an adverse situation (e.g., an infectious disease), could have increased the heterozygote frequency by differential survival (thus, increased fitness) for individuals heterozygous for each of these disorders. Alternatively, heterozygosity for these mutations may have been selected for by unrelated pressures in the past. Although several hypotheses have been advanced (Myrianthopoulos and Aronson, 1972b), the nature of the selective advantages for these mutations remains unknown. The recent finding that certain membrane sphingolipids are involved in the recognition and binding of various bacteria and bacterial toxins (Fishman, 1982; Hannun and Bell, 1989) may be relevant to such an hypothesis. It is tempting to speculate that the slight accumulation of these membrane lipids in heterozygous individuals and the slightly higher levels of these lipids in the circulation could lead to an increased ability to bind, endocytose and degrade these toxic agents, thereby providing a selective advantage for carriers of these mutant alleles. Alternatively, the slight accumulation of

these lipids may have offered a protective effect. Clearly, further insights into the nature of this selective advantage may be gained by the future delineation of the function of these substrates and enzymes in health and disease.

The identification of the R496L allele and other mutations in the ASM gene in Types A and B NPD may provide information for genotype-phenotype correlations and permit more accurate genetic counseling for newly diagnosed cases in families without a previously affected individual. Identification of other mutations, particularly those with residual activity that cause Type B disease, also may provide structure/function information and may facilitate delineation of the active site region. Previously, the enzymatic detection of heterozygotes for NPD Types A and B was not sufficiently reliable to permit mass voluntary screening in the Ashkenazi Jewish community. Thus, the identification of the R496L and other mutations in Types A and B NPD will permit accurate heterozygote identification in families with these lesions as well heterozygote screening and prevention of NPD in the general Ashkenazi Jewish population, as has been the prototypic experience with Tay-Sachs disease (Kabak, 1977). Using molecular techniques, the feasibility of molecular heterozygote screening for Tay-Sachs disease in the Ashkenazi Jewish population have been already demonstrated (Riggs-Raine *et al.*, 1990). The extension of such molecular screening to include the more common mutations causing Gaucher disease and NPD by the use of multiplex PCR should permit the simultaneous screening and prevention of all three sphingolipidoses in the Ashkenazi Jewish population.

Figure 8 **Partial sequence of the amplified ASM cDNA from proband 1.**

The arrows indicate the G to T transversion at nucleotide 1487 (right) which results in an arginine to leucine amino acid substitution of residue 496.

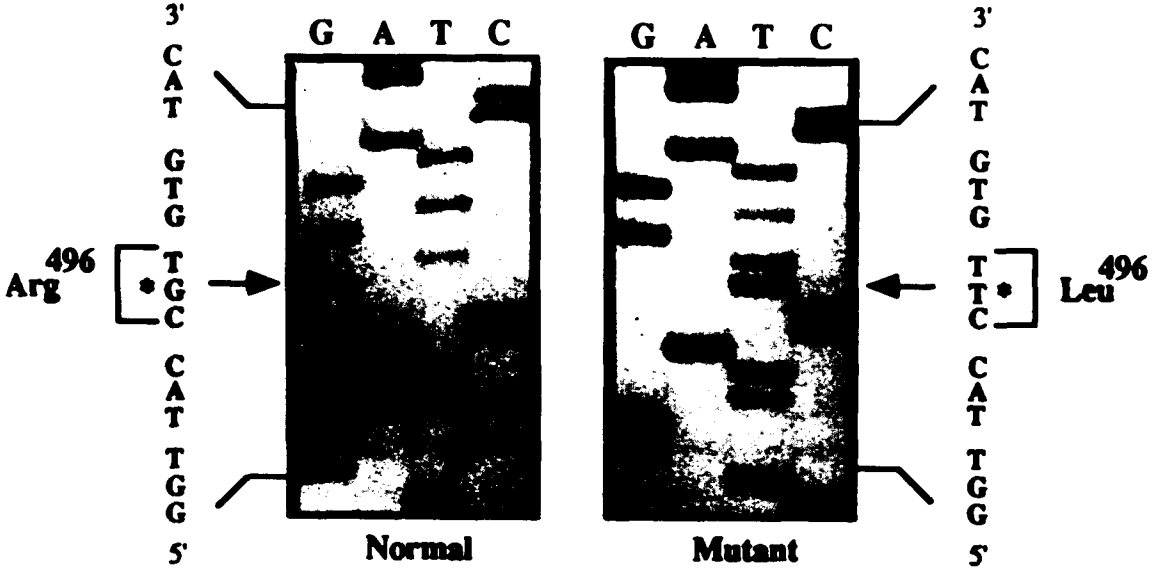

G A T C
G A T C
3'
CAT GTG TGC CAT TGG
5'
Arg 496
Normal
Mutant
3'
CAT GTG TTC CAT TGG
5'
Leu 496

Figure 9 **Dot-blot hybridization with allele-specific oligonucleotides.**

Identification of the R496L mutation in amplified genomic DNAs from the members of an Ashkenazi Jewish family with Type A NPD.

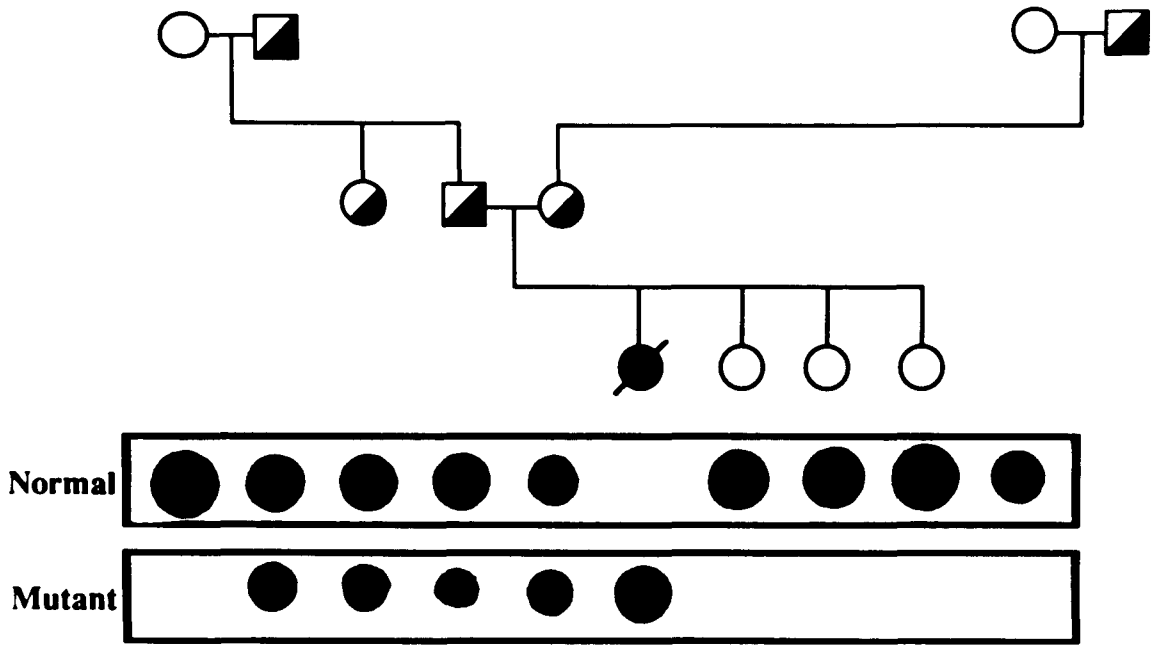

Normal
Mutant

Table IV **Frequency of the R496L alleles in Ashkenazi Jewish and Non-Jewish NPD families**

Source	Mutant alleles	R496L,%
Type A Ashkenazi Jewish	34	32.4
Type A Non-Jewish	43	4.7
Type B Ashkenazi Jewish	4	25.0
Type B Non-Jewish	32	-

Chapter Five

A SECOND COMMON MISSENSE MUTATION (L302P) IN THE ASM GENE OF ASHKENAZI-JEWISH TYPE A NPD PATIENTS

Abstract

A second common mutation, which resulted from a T to C transition at nucleotide 905 and predicted a leucine to proline substitution at ASM codon 302, was identified (designated L302P). The L302P mutation occurred in 24% (8 of 34) of the Ashkenazi Jewish Type A NPD alleles studied. In contrast, it was not found in any of the ASM alleles from non-Jewish Type A patients, in 36 alleles from Type B patients, or in 100 ASM alleles from normal Ashkenazi Jewish individuals. The identification of the L302P mutation should further facilitate molecular carrier detection for NPD in the Ashkenazi Jewish population, particularly since the L302P mutation can be easily detected using the restriction enzyme, *Alw*NI.

Introduction

The R496L mutation was found in about one-third of the Ashkenazi Jewish Type A NPD alleles studied. Some Type A patients were homoallelic for the R496L mutation, while others were heteroallelic, indicating the occurrence of other ASM mutations that caused the Type A phenotype. This chapter describes the identification of a second, common point mutation in the Ashkenazi Jewish Type A NPD patients.

Experimental Procedures

Cell Lines

Cell lines AR (proband 2; GM00112A) and SL (proband 4) were provided by Dr. Elaine Zakai (Childrens Hospital of Philadelphia) and Dr. Gideon Bach (Hebrew University - Hadassah Medical School, Jerusalem), respectively. The remaining information is described under *cell lines* in Chapter Four.

Identification of the L302P Mutation

To identify the ASM mutation(s) in an Ashkenazi Jewish Type A NPD patient, proband 2, PCR-amplification was performed on genomic DNA isolated from cultured skin fibroblasts (Sambrook *et al.*, 1989). The complete ASM coding region, including the intron/exon boundaries, was amplified in four fragments using the PCR primer pairs

shown in Table V (Page 93; primers P1-P8). PCR-amplifications were performed for 30 cycles with *Taq* polymerase (Promega, Madison, WI). Each amplification cycle consisted of denaturation for 1 min at 94° C, annealing for 1 to 2 min at temperatures varying from 60° C to 72° C, and extension for 1 to 2 min at 72° C. For PCR product 1, which included exon 1 and 632 bp of exon 2, a PCR "boost" procedure was required and a slightly modified PCR buffer was used (see Chapter Four). Following PCR-amplification, the PCR products were precipitated by isopropanol in the presence of 2.5 M ammonium acetate, and subcloned into the PCR-1000 vector (Invitrogen, San Diego, CA). For each ligation, from four to six independent subclones were sequenced.

Confirmation of the L302P Mutation by Dot-blot Hybridization and Restriction Enzyme Analyses

For detection of the L302P mutation in other NPD patients, their family members and normal individuals, dot-blot hybridization analyses were performed on a 570 bp genomic fragment amplified from cell lysates using PCR primers P3 and P4 (Table V). PCR-amplification was performed for 30 cycles, each consisting of denaturation for 1 min at 94° C, annealing for 1.5 min at 62° C, and extension for 1.5 min at 72° C. Dot-blot hybridization was performed using Zetabind nylon membranes (AMF-Cuno, Richmond, CA) and a Biorad dot-blot apparatus (Richmond, CA). Hybridizations were carried out by standard techniques using the normal and mutant allele-specific oligonucleotides (ASOs), (5′-GTCACAGCACTTGTGAG-3′) and (5′-GTCACAGCAC<u>C</u>TGTGAG-3′), respectively, with the mutated nucleotide underlined. Following hybridization, the blots

were washed at room temperature for 15 min in 6×SSC containing 0.1% SDS, and then in the same solution for 1 h at 50° C.

The L302P mutation also could be detected by restriction enzyme analysis since the T to C transition introduced an *Alw*NI site into the ASM gene. The 570 bp ASM genomic fragment was amplified using PCR primers P3 and P4. Following amplification, the fragment was digested with *Alw*NI and electrophoresed on a 1.5% agarose gel. Only the mutant L302P allele was digested, resulting in cleavage of the 570 PCR product into two fragments of 280 and 290 bp.

Results

Identification of the L302P Mutation in Proband 2

Since previous dot-blot hybridization studies had demonstrated that proband 2 did not have an R496L allele (Chapter Four), this Ashkenazi Jewish Type A patient was selected as a candidate for identifying new Type A NPD mutations. The ASM coding sequence was PCR-amplified from genomic DNA obtained from proband 2, and the four fragments containing the entire ASM coding region, including the intron/exon boundaries, were subcloned and sequenced. *Figure 10* shows the DNA sequence obtained from a region of exon 2; note that proband 2 had a C to T transition at nucleotide 902 in each of the six sequenced subclones. This single base change predicted a leucine to proline substitution in codon 302 of the ASM polypeptide (designated L302P).

To confirm the authenticity of the L302P missense mutation, dot-blot hybridization

analysis was performed on genomic DNA obtained from proband 2 and his family members (*Figure 11*). These studies demonstrated that proband 2 was homoallelic for the L302P mutation and that both of his parents, a phenotypically normal sister and brother, as well as the maternal grandfather were heterozygous. The L302P mutation was not identified in over 100 ASM alleles from normal Ashkenazi Jewish individuals, indicating that this nucleotide change was not a polymorphism in this population. Further confirmation of the authenticity of the L302P mutation was obtained by sequencing the PCR-amplified ASM coding region from an unrelated, Ashkenazi Jewish Type A NPD patient, proband 4. This patient also was homoallelic for the L302P mutation.

Occurrence of the L302P mutation in Types A and B NPD

The frequency of the L302P mutation in the ASM alleles of other NPD patients and obligate heterozygotes, was determined by Dot-blot or *Alw*NI restriction enzyme analysis of PCR-amplified genomic DNA. Notably, this mutation occurred in about 24% (8 of 34) of the ASM alleles studied from Ashkenazi Jewish Type A NPD patients, whereas it was not found in 43 alleles studied from non-Jewish Type A patients. In addition, the L302P mutation was not identified in 36 alleles studied from Jewish and non-Jewish Type B NPD patients.

Discussion

Previously, a missense mutation, R496L, was identified as the first lesion in the ASM gene causing Type A NPD. In this chapter, a second, common mutation (L302P) is described in Ashkenazi Jewish Type A NPD patients. This amino acid substitution may result in the production of unstable residual enzyme which is rapidly degraded within the cell or inefficiently targeted to the lysosomes. In fact, it is likely that the L302P substitution causes incorrect folding of the ASM polypeptide since it occurs within an α-helix region, and the introduction of proline residues often leads to the formation of β-turns (Klinkhamer *et al.*, 1989). Alternatively, the residual enzymes may be stable and properly targeted to the lysosomes, but remain catalytically inactive due to an altered active and/or substrate binding site. Further characterization of the proteins expressed from these mutant alleles will require immunologic studies using monospecific anti-human ASM antibodies.

The L302P and R496L mutations are responsible for over 55% of the alleles causing Type A NPD in the Ashkenazi Jewish population. Whether one additional major mutation or a series of less frequent mutations will account for the remainder of the lesions causing this disorder must await further analyses. However, evidence from three other disorders, which occur in the Ashkenazi Jewish population with a frequency of 1 in 2500 or greater, suggest that one or more common (15 to 45%) alleles is likely. Tay-Sachs disease and Gaucher disease were previously discussed (see Chapter Four). In addition, in cystic fibrosis, three mutations; ΔF508 (Kerem *et al.*, 1989), W1282X (Kerem

et al., 1991) and G542X (Cutting *et al.*, 1990) are responsible for about 95% of the disease alleles in Ashkenazi Jewish individuals. However, it should be noted that Type A NPD is 20 to 30 times less frequent than Tay-Sachs disease, Gaucher disease, or cystic fibrosis in this population, so the occurrence of multiple rare alleles causing the remainder of Type A NPD would not be surprising. Clearly, the continued elucidation of the mutations causing Type A NPD should provide additional insights into the molecular epidemiology of the diseases that occur frequently among the Ashkenazim. The new finding should facilitate carrier testing for this debilitating disease since analysis of this and the R496L mutation account for about 55% of the Type A NPD alleles among Ashkenazi Jewish individuals.

Figure 10 **Partial sequence of the amplified ASM cDNA from proband 2.**

A small area of the genomic sequence obtained from a normal individual (left) and proband 2 (right) is shown. The arrow indicates the T to C transition at nucleotide 905.

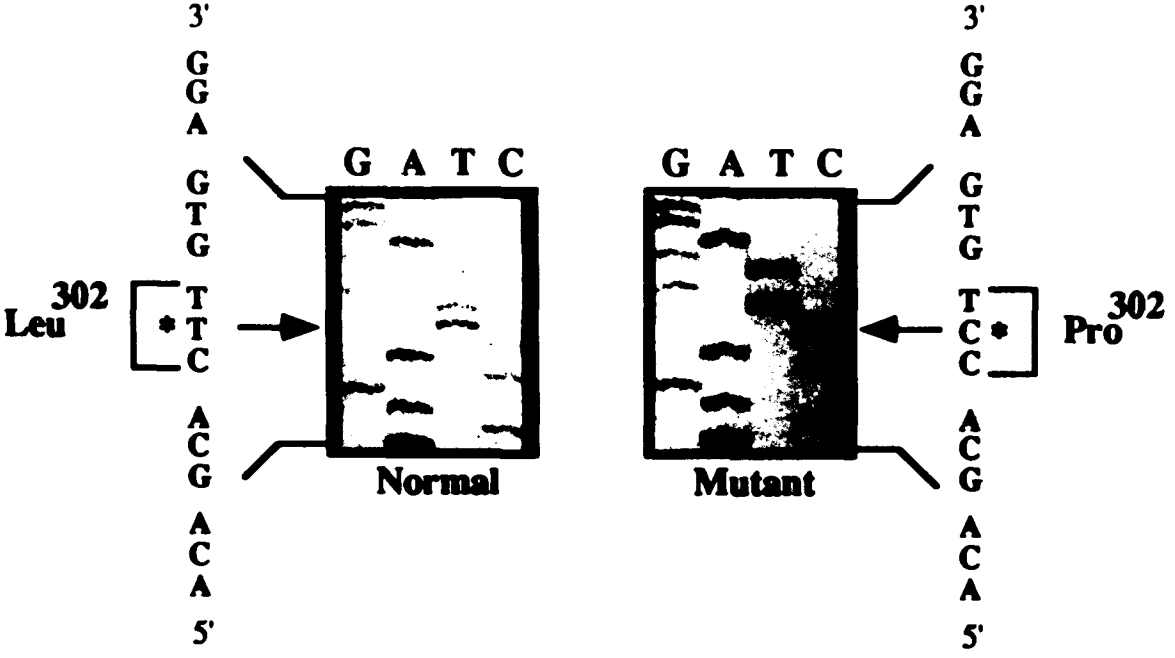

3'
GGA
GTG
TTC
ACG
ACA
5'
Leu 302
G A T C
Normal
G A T C
Mutant
3'
GGA
GTG
TCC
ACG
ACA
5'
Pro 302

Figure 11 **Identification of the L302P mutation in proband 2 and family members.**

Genotype analysis of proband 2 and family members by dot-blot hybridization. Note that proband 2 was homoallelic for the L302P mutation, and that his phenotypically normal brother, sister and maternal grandfather were heterozygous.

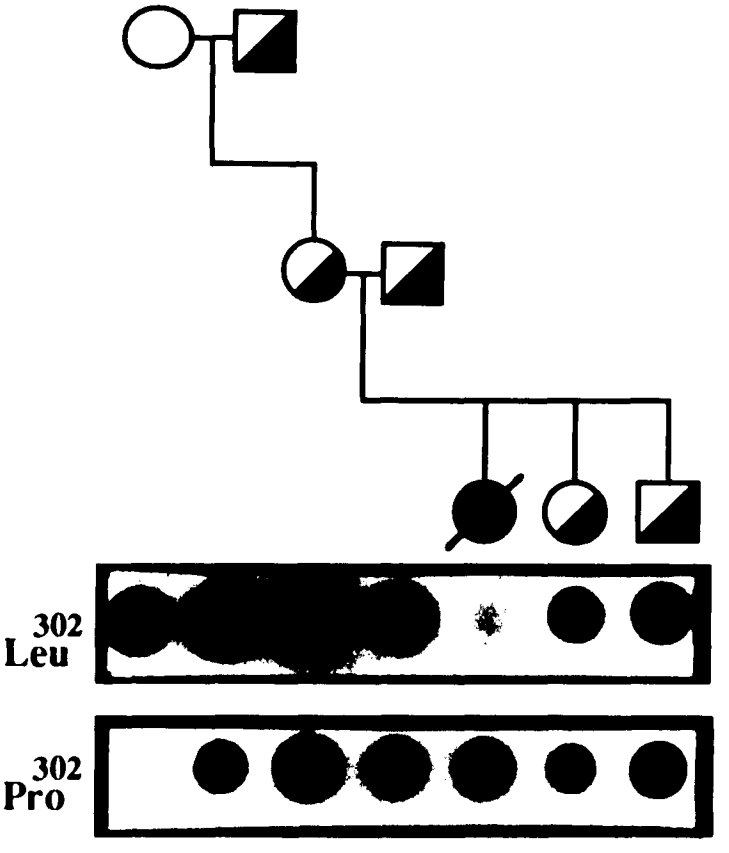

302
Leu
302
Pro

Table V **Synthetic oligonucleotides for PCR-amplifications and dot-blot hybridization**

Primer	Amplified Region	Fragment Size (bp)	Sequence (5'-3')
P1 (sense)	Exon 1 and 2	1443	ATCATCCTCGAGACG GGACAGACGAACCA
P2 (antisense)			AGTAGTCTGCAGAGC AGGGTACACTGGCAC TG
P3 (sense)	Exon 2	570	TCATCCTCGAGCACT GACCTGCACTGGG
P4 (antisense)			AGTAGTCGACTGCTA GAGCAATCAGAG
P5 (sense)	Exon 3-5	969	AGTAGTCGACCCAAA TGCCCAGCAC
P6 (antisense)			AGTAGTCGACAGGCT CTGGCTCCAAC
P7 (sense)	Exon 6	569	AGTAGTCGACATGGG CAGGATGTGTGG
P8 (antisense)			AGTAGTGTCGACTTG CCTGGTTGAACCACA GC

Chapter Six

TYPE B NPD: A SINGLE CODON DELETION (ΔR608) IN THE ASM GENE AND GENOTYPE/PHENOTYPE CORRELATIONS IN TYPE A AND B PATIENTS

Abstract

To investigate the molecular basis for the remarkable phenotypic heterogeneity of Types A and B NPD, the nature of a mutation causing Type B NPD in Ashkenazi Jewish patients was determined. The entire ASM coding region from an Ashkenazi Jewish Type B patient was PCR-amplified, subcloned and completely sequenced. A three base deletion was identified of nucleotides 1821-1823 in the cDNA which predicted the removal of an arginine residue from position 608 of the ASM polypeptide (ΔR608). The other cDNA clones from this patient had the R496L mutation. Both Ashkenazi Jewish Type B patients studied were heteroallelic for the ΔR608 mutation, whereas this allele was not present in 15 unrelated non-Jewish Type B patients, with the notable exception of one mildly affected patient of Arabic descent who was homoallelic for the ΔR608 mutation. These results indicate that the ΔR608 mutation predicts the Type B NPD phenotype, even in the presence of the R496L Type A allele, thereby providing the first genotype/phenotype correlation for this lysosomal storage disease.

Introduction

To date, biochemical investigations into the genetic heterogeneity underlying the neuronopathic Type A and non-neuronopathic Type B phenotypes have been limited since ASM is difficult to purify, the commonly used enzymatic assays are insensitive and labor intensive, and monospecific antibodies have been difficult to produce (Klar *et al.*, 1988; Gal and Fash, 1976; Levade *et al.*, 1986). In general, Type A NPD patients tend to have slightly lower residual activities in crude cell and tissue homogenates than Type B patients, however, the residual activities have not been well characterized (Schneider and Kennedy, 1967; Spence and Callahan, 1989; Besley and Elleder, 1986). Similarly, all of the Type A and B NPD patients studied to date have had cross-reactive immunologic material in crude cell homogenates, however accurate estimates of the amount of enzyme protein have not been determined (Rousson *et al.*, 1986; Jobb, 1987).

In contrast, investigations into the molecular nature of the genetic heterogeneity in NPD have been facilitated by the recent isolation and characterization of the full-length cDNA encoding human ASM. Interestingly, the R496L mutation was detected in one of the two Ashkenazi Jewish Type B patients studied. Thus, efforts were undertaken to determine the molecular lesion in the other ASM allele of the Ashkenazi Jewish Type B patient who was heteroallelic for the R496L mutation. This chapter describes the

identification of a three base deletion in the ASM coding region from this Type B NPD disease patient (proband 3) that results in the deletion of codon 608, which encoded an arginine residue (designated ΔR608).

Experimental Procedures

Cell Lines and Enzyme Assays

Primary cultures of fibroblasts and lymphoblasts were established as already described in Chapter Four. Clinical data on probands 1 and 3 have been published (Chapter Four; Crocker, 1961), and information on proband 5 (cell line 534R.F03) was provided by Dr. Vanier, M. T., Department of Biochemistry, Faculte de Medecine, Lyon, France. The list of cell lines and the methods are the same as described in Chapter Four.

cDNA and Genomic Amplification and Sequencing

For ASM cDNA amplification and sequencing, total RNA was isolated from cultured cells by standard procedures. First-strand cDNA was synthesized with reverse transcriptase from ~5 μg of total RNA using a cDNA synthesis kit according to the manufacturer's instructions (Boheringer Manheim, Indianapolis, IN). The ASM cDNA (~10% of the total reaction) was PCR-amplified with *Taq* polymerase (Perkin-Elmer Cetus, Norwalk, CT) using the oligonucleotide primers described in Chapter Four. PCR was performed for 30 cycles, each consisting of denaturation for 1 min at 94° C and

annealing and extension for 4 min at 72° C. Following amplification, the PCR products were isolated from agarose gels and subcloned into either Bluescript KS(+) (Stratagene, La Jolla, CA) or pGEM 9Zf(-) (Promega, Madison, WI) vectors. For each amplified product, from four to ten independent subclones were sequenced.

To confirm the candidate mutation, genomic DNA was isolated from the proband, and a 567 bp genomic fragment containing the mutation was PCR-amplified using primers P7 and P8 (Table V; Chapter Four; Page 93). The amplified genomic fragment was isolated, subcloned and sequenced as described above for the PCR-amplified ASM cDNAs.

Dot-Blot Analysis of The Mutant Alleles

For detection of the ΔR608 mutation in other NPD patients, their parents and relatives, as well as normal individuals, total genomic DNA was isolated, PCR-amplified as described above, and the 567 bp ASM product was analyzed by dot-blot hybridization. Hybridization of the PCR products with the normal allele-specific oligonucleotide (ASO) (5′-CTCTGTGCCGCCACCTG-3′) or with the ΔR608 ASO (5′-GCTCTGTGCCCCTGAT-3′) were performed for at least 3 hr at 39° C. ASOs labeling was performed as previously described. Following hybridization, the blots were washed at room temperature for 15 min in 6×SSC containing 0.1% SDS, and then for 1-2 hr in the same solution at either 54° C for the normal ASO or 50° C for the ΔR608 ASO. Dot-blot analysis of the R496L mutation was performed as previously described.

Results

Identification of The ΔR608 Mutation in an Ashkenazi Jewish Patient With Type B NPD

Previous studies of an Ashkenazi Jewish Type B NPD patient (proband 3; cell line MS 1271) indicated that he had about 5-8% residual ASM activity and that he was heteroallelic for the R496L mutation. To identify the molecular lesion in his other ASM allele, total RNA from the proband was reverse-transcribed, the ASM coding region was PCR-amplified and the PCR products were subcloned into plasmid vectors for DNA sequencing. A three base deletion (CCG) of nucleotides 1821-1823 in the full-length ASM cDNA was identified which predicted the removal of a single arginine residue at position 608 of the ASM polypeptide (designated ΔR608). The authenticity of this deletion was determined by genomic sequencing (*Figure 12*) and by dot-blot hybridization of PCR-amplified genomic DNA from proband 3 and other family members using an ASO specific for the ΔR608 mutation (*Figure 13*). In addition, Dot-blot hybridizations confirmed the ΔR608/R496L genotype of proband 3, and demonstrated that the ΔR608 and R496L mutations were transmitted from his father and mother, respectively. Proband 3's brother did not receive either mutant ASM allele. The ΔR608 mutation was not identified in over 100 ASM alleles from normal individuals, indicating that the deletion was not a polymorphism.

Occurrence of The ΔR608 Mutation in Types A and B NPD

Table VI (see Page 109) shows the occurrence of the ΔR608 mutation in the ASM alleles of other patients and obligate heterozygotes with Types A and B NPD, as determined by dot-blot analysis of PCR-amplified genomic DNA. Interestingly, a second, unrelated Ashkenazi Jewish Type B NPD patient was heteroallelic for the ΔR608 mutation and another, unknown mutant ASM allele. Of the 16 non-Jewish Type B patients studied, only one, an Arab patient from Algeria (proband 5), was homoallelic for this mutation. This 21 year old male has a mild Type B phenotype (Vanier, M. T., personal communication). Of the 77 ASM alleles from Type A NPD patients or obligate heterozygotes (Ashkenazi Jewish and non-Jewish), none had the ΔR608 mutation.

Comparison of The Residual ASM Activities in Type A and B NPD Patients

Table VII (see page 109) shows the ASM activities in cultured fibroblasts from probands 1, 3, and 5, which were determined using the fluorogenic natural substrate (PSA11-SPM). Normal individuals had a mean activity of 46.3 nmol cleaved/h/mg. In contrast, Type A proband 1, who was homoallelic for R496L, had less than 1% of normal activity. Type B proband 3, whose genotype was R496L/ ΔR608, had a residual activity of about 5% of normal, whereas proband 5 who was homoallelic for ΔR608 had about 13% of normal mean ASM activity, indicating that the ΔR608 allele expressed functional ASM activity in a dosage dependant manner.

Discussion

Since 1966, when Brady *et al.* reported that the primary enzymatic defect in Type A NPD was the deficient activity of ASM, and the demonstration that ASM activity also was markedly decreased in patients with the milder, visceral form of NPD now known as Type B disease (Schneider and Kennedy, 1967), investigators speculated that the remarkable clinical heterogeneity observed among Type A and B NPD patients was due to different mutations in the ASM gene which resulted in altered enzyme polypeptides that expressed varying amounts of residual activity (Rousson *et al.*, 1986; Jobb, 1987). somatic cell genetic studies demonstrated that the mutations causing Types A and B NPD were allelic (Besley *et al.*, 1980). However, efforts to reliably predict either the disease subtype or the severity of Type B patients by the amount of residual ASM activity have not been possible, in part due to the inability of assay procedures to accurately distinguish between patients with low levels of residual ASM activity and/or the presence of the neutral sphingomyelinase activity in cell homogenates (Chatterjee and Gosh, 1989).

The recent cloning and sequencing of the ASM cDNA, has permitted identification of the first mutations which result in Types A and B NPD. Proband 3, an Ashkenazi Jewish Type B NPD patient was found to be heteroallelic for the R496L mutation. In the present study, a three base deletion (ΔR608) in the ASM gene was identified as the other mutation in proband 3. As shown in *Figure 12B*, the three base deletion, CCG, removed the last nucleotide of codon Cys-607 (TG<u>C</u>) and the first two nucleotides of codon Arg-608 (<u>CG</u>C). The new codon 607 formed by this deletion, TGC, also encoded a

cysteine residue. Thus, the deletion resulted in the elimination of a single codon, 608, which encoded an arginine residue.

Detection of the R496L and ΔR608 mutations in patients with Types A and B NPD has permitted genotype/phenotype correlations and provided insights into the function of the altered enzymes encoded by these mutant alleles (Table VII). The homoallelic (R496L/R496L) Type A patient who expresses less than 1% of normal ASM activity in cultured fibroblasts indicates that the R496L mutation encodes an ASM polypeptide with little, if any, catalytic activity and/or markedly decreased stability, thereby resulting the neuronopathic phenotype. In contrast, both heteroallelic (R496L/ΔR608) and homoallelic (ΔR608/ΔR608) Type B patients express an enzyme with sufficient residual activity to prevent neurologic manifestations. These findings suggest that the substitution of an arginine for a leucine at position 496 was more damaging to the enzyme's activity and/or stability than the deletion of an arginine residue in position 608. It follows that Type B patients who are homoallelic for ΔR608 would have a milder disease course than Type B patients who are heteroallelic for ΔR608 and a Type A mutation. Notably, proband 5 had comparatively milder disease manifestations than proband 3 at the same age. Thus, these genotype/phenotype correlations indicate that the more residual ASM activity expressed by the mutant alleles, the milder the disease manifestations. The occurrence of genetic heterogeneity resulting in dramatically different phenotypes is a hallmark of the lysosomal storage diseases. A notable example in which the molecular lesions have been correlated with distinct phenotypes is Tay-Sachs disease (Neufeld, 1989; Navon and Proia, 1989). Mutations causing the classic, infantile form

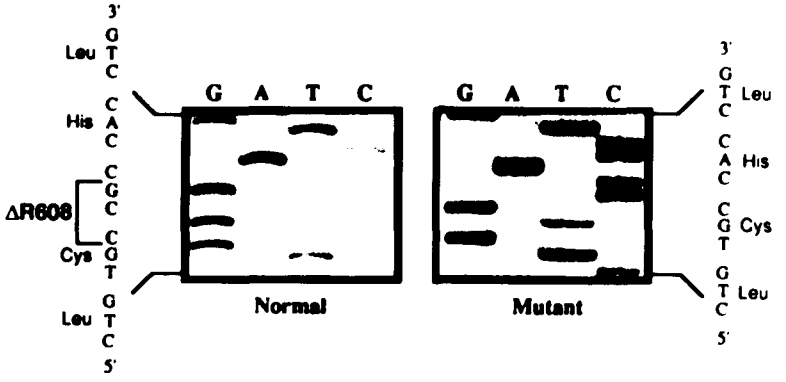
of Tay-Sachs disease, resulted in no detectable transcripts, whereas an exonic point mutation expresses sufficient enzymatic activity to result in the adult or chronic type of GM_2 gangliosidosis. Analogously, identification of the R496L and ΔR608 mutations has provided the first insights into the molecular lesions in the ASM gene underlying the remarkable phenotypic heterogeneity in NPD. Clearly, individuals who are homoallelic for the R496L mutation will have a severe neuronopathic phenotype and Type A NPD. The phenotype of patients heteroallelic for R496L will depend on the genetic lesion present on their other ASM allele. In contrast, patients who are either homo- or heteroallelic for the ΔR608 (e.g., probands 3 and 5) will most likely have Type B NPD. The identification of additional mutations causing Types A and B NPD should permit reliable genotype/phenotype correlations and provide further insights into the functional organization of the ASM polypeptide.

Figure 12 **Three base deletion in the ASM gene from proband 3.**

a) A small area of the genomic sequence obtained from a normal individual (left) and proband 3 (right) is shown.

b) A schematic representation of the ΔR608 mutation.

A.



B.

	606	607	608	609	610
	Leu	Cys	Arg	His	Leu
Normal : 5' -	CTG	TGC	CGC	CAC	CTG - 3'
ΔR608 : 5' -	CTG	TG	C	CAC	CTG - 3'
	Leu	Cys		His	Leu
	606	607		608	609

Figure 13 **Genotype analysis of proband 3 and family members by dot-blot hybridization.**

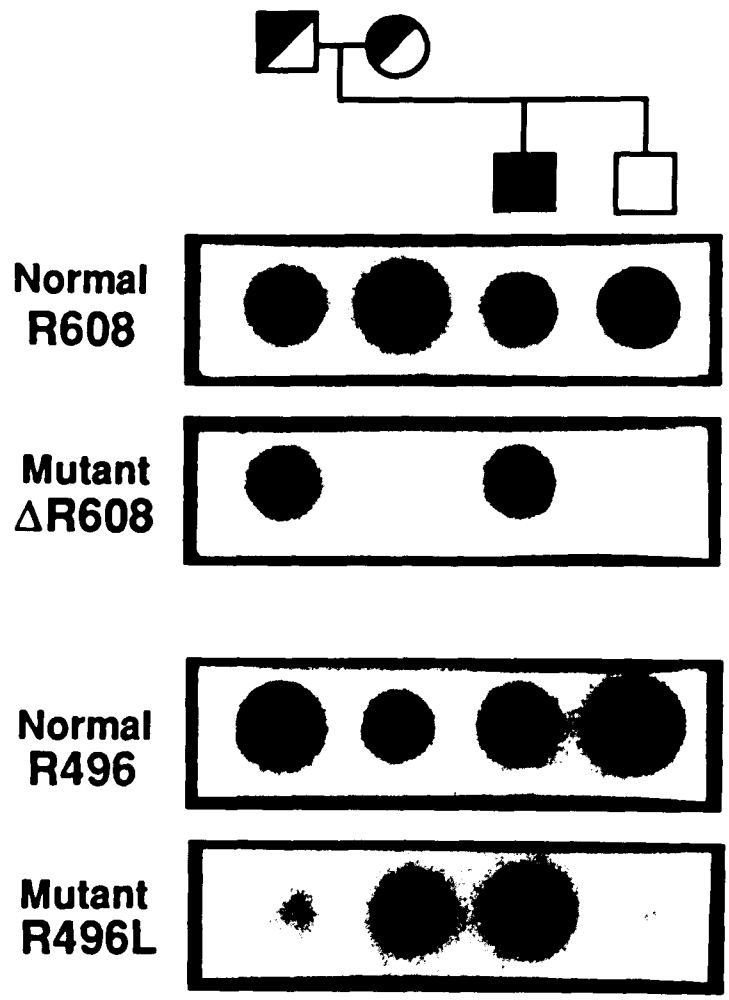

Normal
R608
Mutant
ΔR608
Normal
R496
Mutant
R496L

Table VI **Frequency of the ΔR608 mutation in Ashkenazi Jewish and Non-Jewish families with Types A and B NPD**

Source	Mutant alleles	ΔR608, %
Type A Ashkenazi Jewish	34	-
Type A Non-Jewish	43	-
Type B Ashkenazi Jewish	4	50.0
Type B Non-Jewish	32	6.25

Table VII **Genotype/phenotype correlation in Types A and B NPD**

Genotype	Phenotype	ASM activity	
		mean(range) nmol/h/mg	% of normal mean
R496L/R496L (proband 1)	Type A	0.33 (0.21-0.47)	0.7
R496L/ΔR608 (proband 3)	Type B	2.23 (1.8-2.3)	4.8
ΔR608/ΔR608 (proband 5)	Type B	5.95 (5.1-6.9)	12.8

Chapter Seven

TRANSIENT EXPRESSION OF THE THREE ASM MUTATIONS IN *COS-1* CELLS

Abstract

To confirm the authenticity of the R496L, L302P and ΔR608 mutations, the respective nucleotide changes were introduced into the full-length ASM cDNA by site-directed mutagenesis and transiently expressed in *COS-1* cells. Neither of the mutant constructs expressed catalytically active ASM (mean activity of 6.3-7.5 nmol/h/mg). Untransfected *COS-1* cells had a mean activity of 7.8 nmol/h/mg, and *COS-1* cells transfected with the full-length ASM cDNA, pASM-1FL, had a mean activity of 23.6 nmol/h/mg.

Introduction

Expression of Proteins From Cloned Genes

The scope of recombinant DNA research has broadened to include expression of cloned genes and analysis of the proteins they encode. Proteins synthesized in heterologous systems can be detected by assaying their biological activity *in vivo* or *in vitro*. However, such assays are limited in sensitivity and sometimes are not practical since the host cell expresses an endogenous protein that displays the same biological activity or interferes with it. To overcome these problems, immunological assays using specific antibodies have been developed (e.g. immunoprecipitation, western blotting etc.). Expression of proteins from cloned genes has been used for a number of different purposes, including confirmation of the functional integrity of a cloned gene, production of large amounts of protein, study of the biosynthesis of proteins, elucidation of structure-function relationships of normal and mutant proteins, and confirmation that changes in the DNA sequence of a cloned mutant gene causes the production of defective protein.

Although some mammalian proteins can be expressed in prokaryotic or lower eukaryotic hosts, many mammalian proteins require post-translational modifications that are either not performed or performed differently by bacterial cells, yeast or insect cells (Goedel *et al.*, 1990; Lukow and Summers, 1988). This problem is particularly severe when expression of human lysosomal enzymes is desired. The production of biologically active enzymes that are correctly targeted to the lysosomes requires glycosylation and specific phosphorylation. Some modifications are performed differently even in non-

human mammalian cells and it is important to consider the significance of these differences to the function of the enzyme studied.

Mammalian expression systems can be divided into transient and stable systems. The selection of an appropriate system is based on the questions that have to be answered and on the ease of generation the mutant proteins in an assayable form. Transient expression is the preferred route for surveying the properties of site-directed mutants created *in vitro*, but if a relatively large amount of protein is desirable, or more complicated studies are needed, it is worthwhile evaluating stable expression systems.

Mammalian expression vectors may contain both prokaryotic sequences and eukaryotic transcription units that consist of noncoding sequences and sequences coding for selectable markers. The prokaryotic elements often include a replicon, a gene encoding antibiotic resistance for selection and restriction sites for the insertion of eukaryotic sequences. Most vectors contain sequences from pBR322. The eukaryotic units are promotor elements, signals required for efficient polyadenylation, enhancers and introns with functional splice donor and acceptor sites. Elements that have a broader host range and are active in variety of tissues are the SV40 early gene promotor/enhancer, the SV40 polyadenylation signal and the enhancer/promotor elements from the LTR of the Rous sarcoma virus genome or the human CMV (Kaufman *et al.*, 1990). Vectors also may contain sequences that promote the extrachromosomal replication of the viral genome, as long as trans-acting factors are provided. For example, vectors containing the replicon of SV40 replicate to high copy number in cells that express the appropriate viral T antigen. They can only be used for transient expression because they cause cell death.

The most widely used of all eukaryotic systems is the Simian *COS* cell system. This line of cells were derived by transformation of *CV-1* cells with an origin-defective SV40 genome. They produce the SV40 large T antigen, and they require permissivity factors for replication of DNA containing the SV40 origin of replication. This system is not preferable when high levels of production of a protein are required or when the endogenous activity of the enzyme in interest is high, relative to the level that is achieved by the transient expression of the cloned gene. This system has facilitated analysis of mutations introduced into cloned cDNA.

Large quantities of foreign proteins can be obtained from cell lines carrying amplified chromosomal copies of the gene of interest. In these amplifiable vectors, the foreign gene is attached to a drug resistance marker. The cells are exposed to progressively increasing concentrations of the drug and sublines are selected in which the gene is amplified. The dihydrofolate reductase (*dhfr*) gene is the most extensively used as a marker that is resistant to methotrexate. In stable lines the genes are integrated into the chromosome and in unstable lines, the genes are present as extrachromosomal elements called double minute chromosomes. CHO cells generally contain stably amplified genes.

One of the most highly evolved of mammalian expression vectors is the p91023(B). It contains sequences derived from pBR322 that include the prokaryotic origin of replication and the tetracycline resistance. It also contains the eukaryotic regulatory elements from the SV40 and the adenovirus, as well as a hybrid intron consisting of a 5′ splice site from the adenovirus tripartite leader and a 3′ splice site from

a mouse immunoglobulin gene. It also contains the mouse *dhfr* gene. The inserted DNA is transcribed into a hybrid mRNA that contains the adenovirus tripartite leader which increases the translation efficiency, and the *dhfr* gene which enhance the stability and provides a marker for selection and amplification.

Expression of Genes Encoding Human Lysosomal Enzymes

Normal and mutant cDNAs encoding human lysosomal enzymes have been expressed in various systems. This has enabled the characterization of some naturally occurring and *in vitro* constructed mutant enzymes. The common expression system that has been used for the transient expression of β-glucuronidase (Oshima *et al.*, 1987), β-hexosaminidase A and B (Navon and Proia, 1989; Brown *et al.*, 1989; Lau and Neufeld, 1989; Sonderfeld and Proia, 1989; Neote *et al.*, 1990), α-GalNac (Wang *et al.*, 1990), α-Gal A (Bishop *et al.*, 1988), and β-glucosidase (Grabowski *et al.*, 1989; Reiner *et al.*, 1987; Wigderson *et al.*, 1989) consists of *COS* cells as a host and expression vectors that contain the *SV40* early gene promotor (e.g., p91023(B), pSVL, pSV-2 etc.). These cells usually express endogenous activity of lysosomal enzymes but in most cases the transfected cells showed increased levels after expression of normal human cDNAs. The level of activity of the mutant enzymes that was expressed from the mutant forms of the cDNA was generally close to the background level. In the case of β-glucuronidase, the endogenous *COS* cell enzyme could undergo heat inactivation that did not effect the human enzyme. In the case of β-glucosidase only 2-fold activity was achieved by this system and there are some indications that there is different post-translational glycosidic processing.

Different host cells also have been used, for example, expression studies of arylsulfatase B (Peters *et al.*, 1990) were performed in BHK cells under the control of the SV40 early promotor and with the co-transfection of vectors conferring antibiotic resistance in order to obtain stable expression. NIH-3T3 mouse cells have been used in order to obtain stable expression of β-glucosidase (Ohashi *et al.*, 1991) using a *neo* resistant vector.

Non-mammalian expression system have also been used: The baculovirus system (Lukow *et al.*, 1988), in which insect cells such as *Sf 9* are infected, has been used for stable and abundant expression of β-glucosidase (Grabowski *et al.*, 1989), however there is still uncertainty about the effect of the differences from human cells in some of the post-translational modifications on the biological activity of the protein and the influence of these differences on mutational analysis.

Another approach has been taken for the study of α-Gal A mutations (Koide *et al.*, 1990), when the host cells were Fabry patient fibroblasts with 0.6% residual activity of normal. The cells were transfected with the large T antigen of SV40 and an SV40 derived expression vector that carries the promotor and enhancer of the chick gene for β-actin. The transient expression that was achieved in this system for the normal cDNA was >100 time higher than the control. The mutated form gave 3.7% activity. A similar approach was taken for the expression studies of the acid α-glucosidase (GAA) polymorphic alleles (Martiniuk *et al.*, 1990). The host cells that were chosen were GAA deficient with no mRNA or enzyme activity. They were transformed with SV40 and transfected with SV40 derived vector.

In order to demonstrate if the base changes identified by sequencing the ASM mutant genes, result in the expression of biological inactive or partially active ASM, transient expression studies were performed. Expression studies of the normal full-length cDNA encoding ASM (pASM-1FL) have been conducted in our laboratory (Schuchman *et al.*, 1991; Chapter Two). The cDNA was constructed in the p91023(B) vector and expressed in *COS-1* cells. The endogenous level of ASM activity was 6-7 U/mg protein and the expressed ASM level was ~5 fold over the endogenous level. Based on this study, as well as the expression of other lysosomal enzymes, and taking into account the advantages and disadvantages of this expression system, it seemed practical to use the same system and methods for expression of the mutant ASM cDNAs, (e.g., R496L, L302P and ΔR608 mutations). The respective nucleotide changes were introduced into the full-length ASM cDNA by site-directed mutagenesis and transiently expressed in *COS-1* cells.

Experimental Procedures

Site-Directed Mutagenesis

To evaluate the effects of the mutations (e.g., R496L, L302P and ΔR608) on ASM catalytic activity, they were introduced into the full-length ASM cDNA using the PCR-overlap mutagenesis technique (Ho *et al.*, 1989; Horton *et al.*, 1989).

R496L Construct

As shown in *Figure 14*, two different PCR-amplifications were performed. In the first reaction, two overlapping fragments were independently amplified from the full-length ASM cDNA, pASM-1FL, using PCR primer sets P1 and P2, and P3 and P4 (Table VIII; Page 125), to amplify 356 and 496 bp products, respectively. Following PCR-amplification, the fragments were purified from agarose gels and combined for use as the template in a second PCR-amplification using sense and antisense PCR primers P1 and P4, respectively. Note that primers P1 and P4 overlapped unique *Bcl*I and *Sph*I sites in the pASM-1FL cDNA, respectively. PCR-amplifications were performed for 20 to 25 cycles, each consisting of denaturation for 1 min at 94° C, annealing for 1 to 2 min at temperatures varying from 45 to 65° C, and extension for 1 to 2 min at 72° C. Following amplification, the PCR product was subcloned into the PCR-1000 vector and sequenced in order to confirm the presence of the mutation and to determine that no other mutations had been introduced into the ASM cDNA fragment. To construct the full-length ASM cDNA containing the R496L mutation, the mutated *Bcl*I-*Sph*I fragment was isolated from the PCR-1000 subclone and inserted into the pASM-1FL cDNA by ligation exchange. The mutated full-length cDNA was then resequenced to ensure that the only mutation was R496L. To perform the methylation sensitive *Bcl*I digest, the PCR-1000 subclone containing the mutated fragment, and a Bluescript (Stratagene, La Jolla, CA) subclone containing the pASM-1FL cDNA were grown in the *dam*⁻ *E.coli* strain GM2163.

L302P Construct

The full-length ASM cDNA containing the L302P mutation was constructed by essentially the same method with the following modifications. As shown in *Figure 15*, For the first PCR-amplification primer sets P5 and P6, and P7 and P8 (Table VIII) were used. Primers P5 and P8, which overlapped unique *Sfi*I and *Bcl*I sites in the pASM-1FL cDNA, respectively, were then used for the second PCR-amplification. PCR-amplification conditions are essentially as described for the R496L construct. Following amplification, the 734 bp PCR product was subcloned into the PCR-1000 vector and sequenced. To construct the full-length ASM cDNA containing the L302P mutation, the mutated *Sfi*I-*Bcl*I fragment was isolated from the PCR-1000 subclone and inserted into pASM-1FL cDNA by ligation exchange. The mutated full-length cDNA was then resequenced to ensure that the only mutation was L302P.

ΔR608 Construct

The ΔR608 construct was built essentially as described for the other constructs. As is shown in *Figure 16*, the two sets of primers: P9 and P10, and P11 and P12 (Table VIII) were used for the first PCR-amplifications, and the external primers P9 and P12 that overlapped unique *Sph*I and *Nco*I sites respectively, were used for the second PCR-amplification. There was no need for transforming *dam*⁻ cells for this construct. The *Sph*I-*Nco*I fragment was introduced into pASM-1FL cDNA.

Transient Expression of the Mutant Constructs in COS-1 Cells

To evaluate the effect of the mutations on ASM catalytic activity, the mutated full-length ASM cDNAs were inserted into the eukarytoic expression vector p91023(B) (Kaufman and Sharp, 1982; Wong *et al.*, 1985). DNA (15 to 20 μg) from each construct, as well as the wild type full-length pASM-1FL cDNA, were introduced into *COS-1* cells by electroporation using a Biorad electroporation apparatus. As a negative control for the transient expression studies, *COS-1* cells were electroporated in the presence of buffer alone. The transfected cells were harvested after 72 h and ASM activities were determined using (PSA11-SPM), as previously described.

Results

Transient Expression of the Mutant Constructs

Table IX shows the ASM activities expressed by the three constructs in *COS-1* cells, using the fluorogenic natural substrate (PSA11-SPM). Neither of the mutant constructs expressed catalytically active ASM (mean activity of 6.3-7.5 nmol/h/mg). Untransfected *COS-1* cells had a mean activity of 7.8 nmol/h/mg. In comparison, *COS-1* cells transfected with the full-length ASM cDNA, pASM-1FL, had a mean activity of 23.6 nmol/h/mg.

Discussion

Transient expression in *COS-1* cells demonstrated that the R496L and L302P amino acid substitutions altered the ASM polypeptide such that the residual enzyme had no detectable catalytic activity, consistent with the fact that homoallelism for either of these mutations leads to the severe Type A NPD phenotype and less than 1% residual ASM activity in cultured cells (Beaudet *et al.* 1980, and this work). These amino acid substitutions may result in the production of unstable residual enzymes which are rapidly degraded within the cell or inefficiently targeted to the lysosomes. Alternatively, the residual enzymes may be stable and properly targeted, but remain catalytically inactive due to an altered active and/or substrate binding site. Further characterization of the proteins expressed from these mutant alleles will require immunologic studies using monospecific anti-human ASM antibodies, which are not currently available.

Expression of the in-frame deletion, ΔR608, resulted in no detectable catalytic activity, inconsistent with the fact that Type B patient homoallelic for the ΔR608 mutation had more then 12% residual ASM activity in cultured cells (see Chapter Six), and had a relatively mild form of NPD. In order to explain these contradicting results, one has to take into account the following; The *COS-1* system is limited and could not completely represent the human body, such that some kinds of mutation would behave differently (see Introduction). Another limitation of this system, in this particular study, was the high ASM background level and the relatively low expression level of the normal cDNA, what together made it theoretically impossible to detect differences between no enzymatic

activity (e.g., background level) and low enzymatic activity. Interestingly, similar results were obtained for the expression of another Type B mutation, (Schuchman, unpublished results), so it is possible that the mutant enzymes in these cases, are not active in the *COS-1* expression system, but behave differently in tissue cultures or *in vivo*.

Table VIII **Synthetic oligonucleotides for cDNA amplifications of the mutant constructs**

Primer	Amplified Region	Fragment Size	Sequence (5'-3')
P1 (sense)	nt 611-913	302 bp	TCATCCTCGAGCACTGAC CTGCACTGGG
P2 (antisense)			TCCTCACAGGTGCTGTGA CGGTGGT
P3 (sense)	nt 898-1330	432 bp	ACAGCACCTGTGAGGAAG TTCCTGG
P4 (antisense)			GAATTCCTACAATTCGGT AATAATTCC
P5 (sense)	nt 1139-1495	356 bp	CAATATGAATTTTTGTT
P6 (antisense)			GGTACACAAGGTAACCAG GATTAAG
P7 (sense)	nt 1478-1974	496 bp	TGGTTACCTTGTGTACCA AATAGAT
P8 (antisense)			AGTAGTGTCGACTTGCCT GGTTGAACCACAGC
P9 (sense)	nt 3802-4149	347 bp	TGGTTACCTTGTGTACCA AATAGAT
P10 (antisense)			CAGGTGG*CACAGAGCAG GGCTGTCA
P11 (sense)	nt 4132-4470	338 bp	GCTCTGTG*CCACCTGATG CCAGATG
P12 (antisense)			TCAACTTAGGCCTCTACTC C

The underlines indicate mutated nucleotide.
The * indicates deleted region.

Figure 14 **Construction of the p91023(B) expression vector containing the R496L mutation.**

The positions of the *Eco*RI, *Sph*I, and *Bcl*I restriction sites are indicated, as are the locations and orientations of the PCR primers (P1-P4).

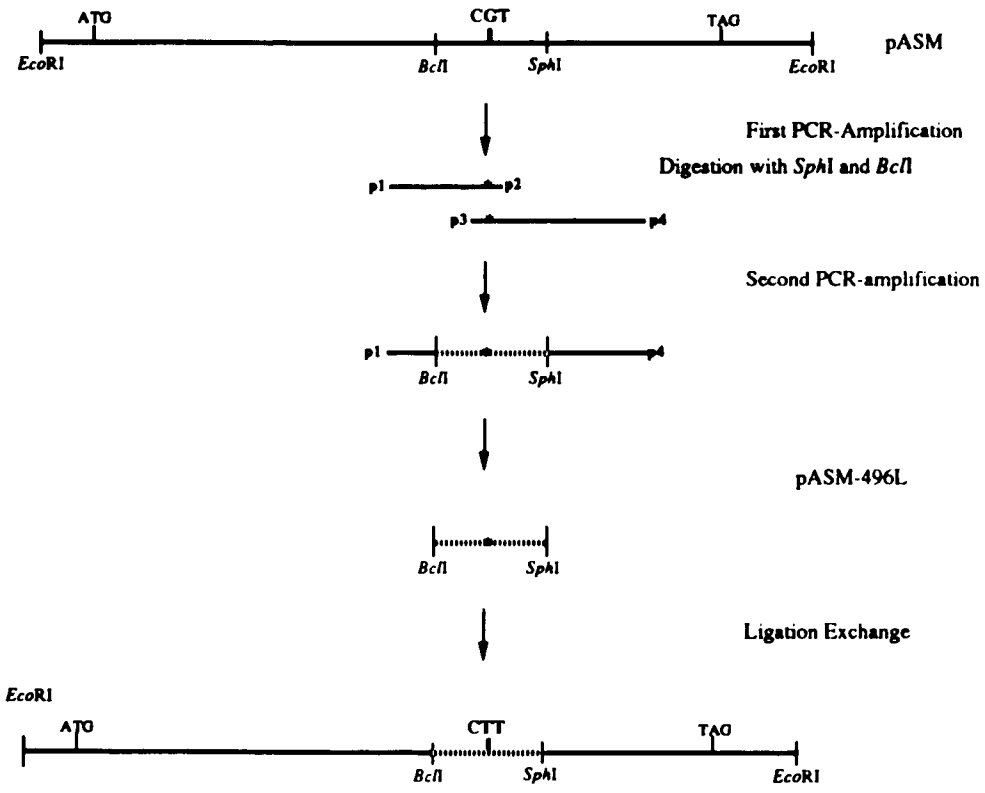

ATG
CGT
TAG
pASM
EcoRI
BclI
SphI
EcoRI
First PCR-Amplification
Digestion with SphI and BclI
p1
p2
p3
p4
Second PCR-amplification
p1
p4
BclI
SphI
pASM-496L
BclI
SphI
Ligation Exchange
EcoRI
ATG
CTT
TAG
BclI
SphI
EcoRI

Figure 15 **Construction of the p91023(B) expression vector containing the L302P mutation.**

The positions of the *Eco*RI, *Sfi*I, and *Bcl*I restriction sites are indicated, as are the locations and orientations of the PCR primers (P5-P8).

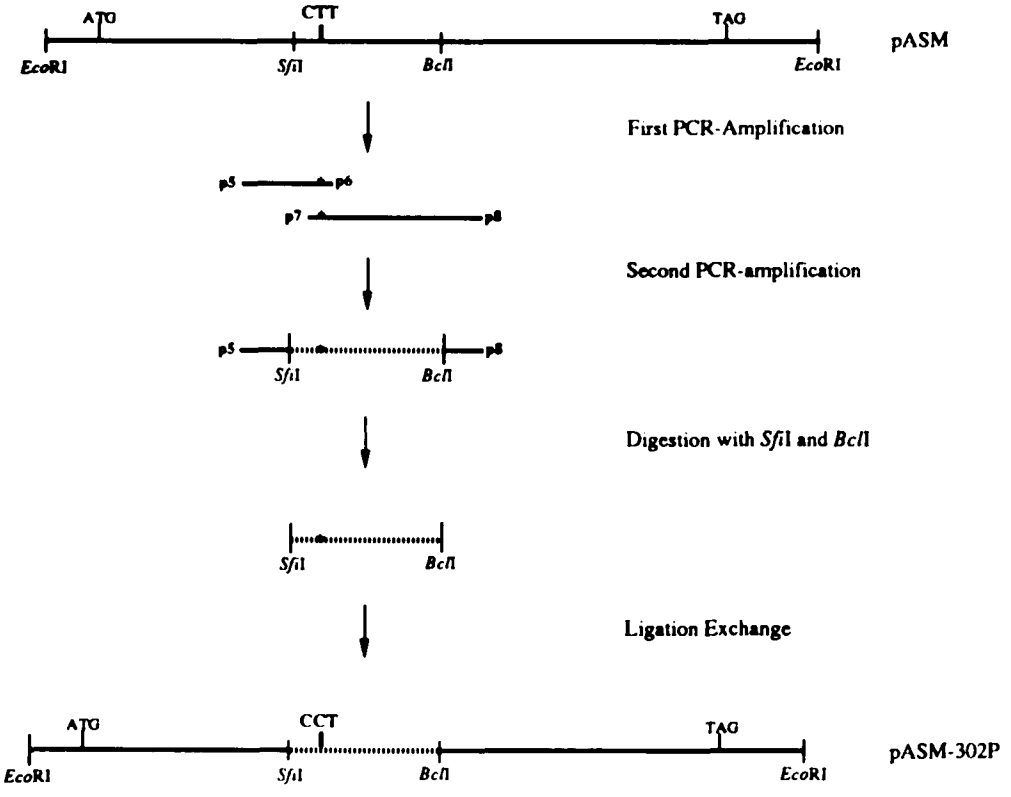

ATG
CTT
TAG
EcoRI
SfiI
BclI
EcoRI
pASM
First PCR-Amplification
p5
p6
p7
p8
Second PCR-amplification
p5
p8
SfiI
BclI
Digestion with SfiI and BclI
SfiI
BclI
Ligation Exchange
ATG
CCT
TAG
EcoRI
SfiI
BclI
EcoRI
pASM-302P

Figure 16 **Construction of the p91023(B) expression vector containing the ΔR608 mutation.**

The positions of the *Eco*RI, *Sph*I, and *Nco*I restriction sites are indicated, as are the locations and orientations of the PCR primers (P9-P12).

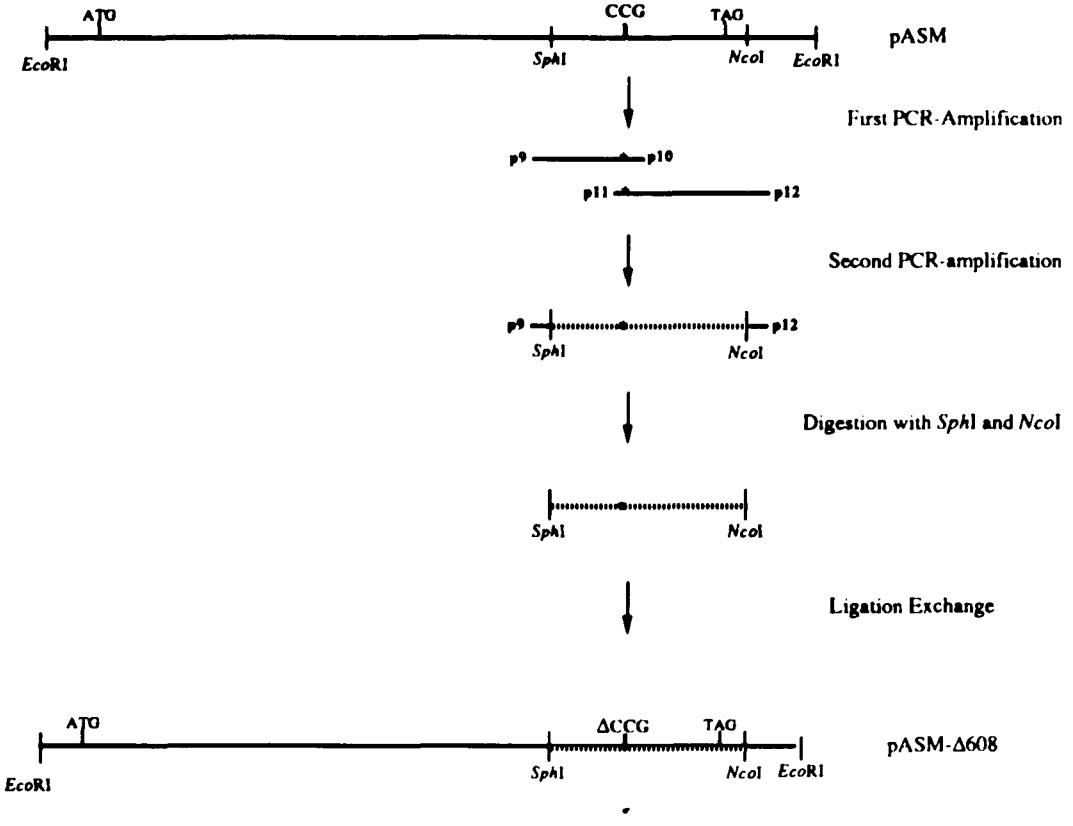

ATG
CCG
TAG
pASM
EcoR1
SphI
NcoI
EcoR1
First PCR-Amplification
p9
p10
p11
p12
Second PCR-amplification
p9
p12
SphI
NcoI
Digestion with SphI and NcoI
SphI
NcoI
Ligation Exchange
ATG
ΔCCG
TAG
pASM-Δ608
EcoR1
SphI
NcoI
EcoR1

Table VIX **Transient expression of ASM mutations in *COS-1* cells**

Enzyme source	Intracellular ASM Activity (nmol/h/mg)
COS-1 Cells	7.8 ∓ 3.42
ASM Constructs:	
pASM-1FL (sense)	23.6 ∓ 8.12
R496L	6.6 ∓ 2.29
L302P	7.5 ∓ 2.44
ΔR608	6.3 ∓ 1.77

Values represent the means of four independent transfections, ∓1 standard deviation

Bibliography

Agmon, V., Dinur, T., Cherbu, S., Dagan, A. and Gatt, S. (1991) *Exper. Cell Res.* 196, 151-157.

Banzai, A., Lowden, J. A. and Charleton, K. M. (1979) *Vet. Pathol.* 16, 530-538.

Barton, N. W., Furbish, F. S., Murray, G. J., Garfield, M. and Brady, R. O. (1990) *Proc. Natl. Acad. Sci.* USA 87(5), 4519-4523.

Beaudet, A. L., Hamptor, M. S., Patel, K. and Sparrow, J. T. (1980) *Clin. Chim. Acta.* 108, 403-414.

Besley, G. T. N., Hoogeboom, A. J. M., Hoogeveen, A., Kleijer, W. J. and Galjaard, H. (1980) *Hum. Genet.* 54, 409-412.

Beutler, E., Gelbart, T., Kuhl, W., Sorge, J. and West. C. (1991) *Proc. Natl. Acad. Sci. USA* 88, 10544-10547.

Beyreuther, K., Bieseler, B., Bovens, J., Dildrop, R., Neifer, K., Stuber, K., Zaiss, S. and Ehring, R. (1983). In: Modern Methods in Protein Chemistry. Walter de Gruyter, Berlin, pp. 303-325.

Bishop, D. F. and Desnick, R. J. (1981) *J. Biol. Chem.* 256, 1307-1316.

Bishop, D. F., Kornriech, R., Eng, C. M., Ioannou, Y. A., Fitzmaurice, T. F. and Desnick, R. J. (1988) In: Lipid Storage Disorders. Salvayre, R., Douste-Blazy, L. and Gatt, S. (eds). Plenum Publ. Corp. pp. 809-822.

Blanchette-Mackie, E. G., Dwyer, N. K., Amenda, L. M., Kruth, H. S., Butler, J. D., Sokol, J., Comly, M. E., Vanier, M. T., August, J. T., Brady, R. O. and Pentchev, P. G. (1988) *Proc. Natl. Acad. Sci.* USA 85, 8022-8026.

Brady, R. O., Kanfer, J. N., Mock, M. B., and Fredrickson, D. S. (1966) *Proc. Natl. Acad. Sci. USA* 55, 366-370.

Brady, R. O. (1983) In: The Metabolic Basis of Inherited Diseases. Stanbury, J. B., Wyngaarden, J. B., Fredrickson, D. S., Goldstein, J. L. and Brown, M. S. (eds). McGraw-Hill, N.Y., pp. 831-884.

Brown, C. A., Neote, K., Leung, A., Gravel, R. A. and Mahuran, D. J. (1989) *J. Biol. Chem.* 264, 21705-21710.

Callahan, J. W., Jones, C. S., Davidson, D. J. and Shankaran, P. (1983) *J. Neurosci. Res.* 10, 151-157.

Caruthers, M. H. (1985) *Science*. 230, 281-285.

Chase, G. A. and McKusick, V. A. (1972) *Am. J. Hum. Genet.* 24, 339-340.

Chatterjee, S. and Ghosh, N. (1989) *J. Biol. Chem.* 264, 12554-12561.

Christomanou, H., Aignesberger, A. and Linke, R. (1986) *Biol. Chem. Hoppe-Seyler* 367, 879-890.

Coulondre, C., Miller, J. H., Farabaugh, P. J., and Gilbert, W. (1978) *Nature* 274, 775-780.

Croker, A. C. and Farber, S. (1958) *Medicine* 37, 1-96.

Crocker, A. C. (1961). *J. Neurochem.* 7, 69-78.

Cutting, G. R., Kasch, L. M., Rosentein, B. J., Zielenski, J.,Tsui, L-C., Antonarakis, S. E. and Kazazian, H. (1990) *Nature* 346, 366-369.

Dahn, N. M., Lobel, P., Kornfeld, S. (1989) *J. Bio. Chem.* 25, 12115-12118.

Daloze, P., Delvin, E. E., Glorieux, F. H., Corman, J. L., Bettez, P. and Toussi, T. (1977) *Am. J. Med. Genet.* 1, 229-233.

Delvin, E., Glorieux, F., Dalozoe, P., Gorman, J. and Block, P. (1974) *Am. J. Hum. Genet.* 26, 25A.

Farquhar, M. G. (1985) *Ann. Rev. Cell Biol.* 1, 447-488.

Ferlinz, K., Hurwitz, R. and Sandhoff, K. (1991) *bioch. Bioph. Res. Commun.* 179, 1187-1191

Fishman, P. (1982) *J. Mem. Biol.* 69, 85-97.

Fraikor, A. L. (1977) *Soc. Biol.* 24, 117-134.

Gal, A. E. and Fash. F. J. (1976) *Chem. Phys. Lipids* 16, 71-79.

Gatt, S. (1970) *Chem. Phys. Lipid.* 5, 235-249.

Gatt, S., Dinur, T. and Kopolovic, J. (1978) *J. Neurochem.* 31, 547-551.

Geier, C., von Figura, K. and Pohlmann, R. (1989) *FEBS* 13, 611-616.

Gilboa, E., Eglitis, M. A., Kantoff, P. W. and Anderson, W. F. (1986) *Biotechniques* 4, 504-512.

Gilmor, M. S., Cruz-Rodz, A. L., Leimeister-Wachter, M., kreft, J. and Goebel, W. (1989) *J. Bacter.* 171, 744-753.

Goedel, D. V. (1990) In: Methods in Enzymology. Academic Press Inc. Vol. 85 pp. 3-7.

Goodman, R. M. (1979) Genetic Disorders Among The Jewish People (The Johns Hopkins University Press, Baltimore), pp. 96-100.

Grabowski, G. A., William. W. R. and Grace, M. G. (1989) *Enzyme* 41, 131-142.

Hannun, Y. A. and Bell. R. M. (1989) *Science* 243, 500-507.

Ho, N. S., Hunt, H. D., Hortur, R. M., Pullen, J. K. and Pease, L. R. (1989) *Gene* 77, 51-59.

Horowitz, M., Wilder, S., Horowitz, Z., Reiner, O., Gelbart, T. and Beutler, E. (1989) *Genomics* 4, 87-96.

Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. and Pease, L. R. (1989) *Gene* 77, 61-68.

Itakura, K., Rossi, J. J. and Wallace, R. B. (1984) *Annu. Rev. Biochcem.* 53, 323-356.

Jobb, E. (1987) University of Toronto. Thesis.

Jones, C. S., Shankaran, P. and Callahan, J. W. (1981) *Biochem. J.* 195, 373-382.

Jurka, J. and Smith, T. (1988) *Proc. Natl. Acad. Sci. USA* 85, 4775-4778.

Kaback, M. M. (1977) *Prog. Clin. Biol. Res.* 18, 1-7.

Kaufmann, R. J. and Sharp, P. A. (1982) *Mol. Cell Biol.* 2, 1304-1319.

Kaufman, R. J. (1990) In: Methods in Enzymology. Academic Press Inc. Vol. 185 pp. 487-511; 577-595.

Kerem, B., Rommens, J. M., Buchanan, J. A., Markiewich, D., Cox, T. K., Chakravarti, A., Buchwald, M. and Tsui. L-T. (1989) *Science* 245, 1073-1079.

Kitagawa, T. (1987) *Jpn. J. Hum. Genet.* 32, 55-69.

Klar, R., Levade, T. and Gatt, S. (1988) *Clin. Chim. Acta.* 176, 259-268.

Klinkhamer, M. P., Groen, N. A., van der Zon, R., Lindhout, D., Sandkuyl, L. A., Krans, H. M. J., Moller, W. and Maassen. J. A. (1989) *EMBO J.* 8, 2503-2507.

Knudson, A. G. and Kaplan, W. D. (1962). In: Cerebral Sphingolipidoses. Aronson, S. M. and Volk, B. W. (eds). Academic Press, New York. pp. 395-411.

Koide, T., Ishiura, M., Iwai, K., Inoue, M., Kanaeda, Y., Okada, Y. and Uchida, T. (1990) *FEB-Lett* 295, 353-356.

Kornfeld, S. and Mellman, I. (1989) *Ann. Rev. Cell Biol.* 5, 483-525.

Kornreich, R., Desnick, R. J. and Bishop, D. F. (1989) *Nucl. Acids Res.* 17, 3301-3302.

Koval, M. and Papano, R. E. (1991) *Biochem. Biophys. Acta.* 1082, 113-125.

Kurth, J. and Stoffel, W. (1991) *Biol. Chem. Hopp-Seyler* 372, 215-223.

Lau, M. M. H. and Neufeld, E. F. (1989) *J. Bio. Chem.* 264, 21376-21380.

Levade, T., Salvayre, R. and Douste- Blazy, L. (1986) *J. Clin. Chem. Clin. Biochem.* 24, 205-220.

Lukow, V. A. and Summers, M. D. (1988) *BioTech.* 6, 47-55.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982). In: Molecular Cloning: a Laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Martiniuk, F., Bodkin, M., Tzall, S. and Hirshhorn, R. (1990) *Am. J. Hum. Genet.* 47, 440-445.

Maruyama, E. N. and Arima, M. (1989) *J. Neurochem.* 52, 611-618.

Maziere, J. C., Maziere, C. and Hosli, P. (1978) *Mongr. Hum. Genet.* 9, 198-205.

Merril, A. H. Jr. and Jones, D. D. (1990) *Bioch. et Bioph. Acta.* 1044, 1-123.

Messing, J., Crea, R. and Seeburg, P. H. (1981) *Nucl. Acid. Res.* 9, 309-321.

Miller, R. D., Hoffmann, J. W., Powell, P. P., Kyle, J. W., Shipley, J. M., Bachinsky, D. R. and Sly, W. S. (1990) *Genomics* 7, 280-283.

Miyawaki, S., Yoshida, H., Mitsuoka, S., Enomoto, H. and Ikehara, S. (1986) *J. hered.* 77, 379-384

Morimoto, S., Martin, B. M., Kishimoto, Y. and O'Brian, J. S. (1988) *Biochem. Biophy. Res. Con.* 156, 403-410.

Morreo, H., Galjart, N. J., Gillemans, N., Willemsen, R., van der Horst, G. T. J. and d'Azzo, A. (1989) *J. Biol. Chem.* 264, 20655-20663.

Mount, S. M. (1982) *Nucl. Acids. Res.* 10, 459-472.

Myerowitz, R. and Costigan, F. C. (1988) *J. Biol. Chem.* 263, 18587-18589.

Myerowitz, R. (1988) *Proc. Natl. Acad. Sci. USA* 85, 3955-3959.

Myrianthopoulos, N. C. and Aronson, S. M. (1972). In: Advances in Experimental Medicine and Biology. Volk, B. W. and Aronson, S. M. (eds). Plenum Press, New York. pp. 561-570.

Myrianthopoulos, N. C., Naylor, A. F. and Aronson, S. M. (1972) *Am. J. Hum. Genet.* 24, 341-342.

Myrianthopoulos, N. C. and Melnick, M. (1977) *Prog. Clin. Biol. Res.* 18, 95-106.

Navon, R. and Proia, R. (1989) *Science* 243, 1471-1474.

Neote, K., Brown, C, A., Mahuran, D. J. and Gravel, R. A. (1990) *J. Biol. Chem.* 265, 20799-20806.

Neufeld, E. F. (1989) *J. Biol. Chem.* 264, 10927-10930.

Niemann, A. (1914) *Fahrb. Kinderheikd* 79, 1-6.

O'Brien, J. S. and Kishimoto, Y. (1991) *FASEB. J.* 5, 301-308.

Ohashi, T., Hong, C. M., Wieler, S., Tomich, J. M., Aerts, J. M., Tager, J. M. and Barranger, J. A. (1991) *J. Biol. Chem.* 266, 3661-3667.

Okayama, H. and Berg, P. (1983) *Mol. Cell. Biol.* 3, 280-289.

Oshima, A., Kyle, J. W., miller, R. D., Hoffman, J. W., Powell, P. P., Grubb, J. H., Sly, W., Tropak, M., Guise, K. S. and Gravel, R. A. (1987) *Proc. Natl. Acad. Sci. USA* 84, 685-689.

Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S. and Sharp, P. A. (1986) *Ann. Rev. Biochem.* 55, 1119-1150.

Parker, R. I., Barton, N. W., Read, E. J. and Brady, R. O. (1991) *Am. J. Hematol.* 38, 130-137.

Patrick, A. D., Young, E., Kleijer, W. J., Niermeijer, M. F. (1977) *lancet* 2, 144.

Pavone, L., Fiumara, A. and Larosa, M. (1986) *J. Inher. metab. Dis.* 9, 73-78.

Penchev, P. G., Brady, R. O., Gal, A. E. and Hibbert, S. R. (1977) *Bioch. Bioph. Acta.* 488, 312-321.

Peters, C., Bernard, S., Winfried, R., Rupp, K., Zuhlsdorf, M., Vingron, M., Meyer, H. E., Pohlmann, R. and Von Figura, K. (1990) *J. Biol. Chem.* 265, 3374-3381.

Pick, L. (1927) *Med. klin.* 23, 1483.

Proia, R. L. and Soravia, E. (1987) *J. Biol. Chem.* 262, 5677-5681.

Proia, R. L. (1988) *Proc. Natl. Acad. Sci. USA.* 85, 1883-1887.

Quintern, L. E., Weitz, G., Nehrkorn, H., Tager, J. M., Schram, A. W. and Sandhoff, K. (1987) *Biochem. Biophys. Acta.* 922, 323-336.

Quintern, L. E., Schuchman, E. H., Levran, O., Suchi, M., Ferlinz, K., Reinke, H., Sandhoff, K. and Desnick,R.J. (1989) *EMBO Journal.* 9, 2469-2473.

Reiner, O., Givol, W. D. and Horowitz, M. (1987) *DNA* 6, 101-108.

Reiner, O., Wigderson, M. and Horowitz, M. (1988) *DNA* 7, 107-116.

Riggs-Raine, B. L., Feigenbaum, A. S. J., Natowicz, M., Skomorowski, M., Schuster, S. M., Clarke, J. T. R., Mahuran, D. J., Kolodny, E. H. and Gravel, R. A. (1990) *New. Eng. J. Med.* 323, 6-12.

Rousson, R., Vanier, M. T. and Louisot. P. (1986). In: Enzymes of Lipid Metabolism II. (Freysz, L., Dreyfus, H., Massarelli, R. and Gatt, S., eds.). N.Y., Plenum Publishing Corporation, N.Y. pp. 273-283.

Rousson, R., Bonnet, J., Louisot. P. and Vanier, M. T. (1987) *Biochem. Biophys. Acta* 924, 502-508.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. (1988) *Science* 239, 487-491.

Sakiyama, T., Owada, M., Kitagawa, T., Miyawaki, S., Shinagawa, T. and Tadokoro, M. (1986) *J. Inher. Metab. Dis.* 9, 305-308.

Sakuragawa, N. (1982) *J. Biochem.* 42, 637-649.

Sambrook, J. Fritsch, E. F. and Maniatis, T. A. (1989). In: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nickleson, J. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.

Scaggiante, B., Pinwschi, A., Sustersich, M., Andolina, M., Agosti, E. and Romeo, D. (1987) *Transplantation* 44, 59-61.

Schneider, P. B. and Kennedy, E. (1967) *J. Lipid. Res.* 8, 202-209.

Schuchman, E. H., Suchi, M., Takahashi, T., Sandhoff, K. and Desnick, R. J. (1991) *J. Bio. Chem.* 261, 8531-8539.

Shipley, J. M., Miller, R. D., Wu, B. M., Grubb, J. H., Christensen, S. G., Kyle, J. K. and Sly, W. S. (1991) *Genomics*, 10, 1009-1018.

Shoshani, T., Augarten, A., Gazit, E., Bashan, N., Yahav, Y., Rivlin, Y., Tal, A., Seret, H., Yaar, L., Kerem, E. and Kerem, B. (1992) *Am. J. Hum. Genet.* 50, 222-228.

Sonderfeld-Fresko, S. and Proia, R. L. (1989) *J. Biol. Chem.* 264, 7692-7697.

Spence, M. W. and Callahan, J. W. (1989) In: The Metabolic Basis of Inherited Diseases. Schriver, C. R., Beaudet, A. L., Sly, W. S. and Valle, D. (eds). McGraw-Hill, N.Y., 8th ed. pp. 1655-1767.

Stoffel, W. and Melzner, J. (1980). *Hoppe-Seyler's Z. Physiol. Chem.* 361, 755.

Stone, K. and Williams, K. (1986) *J. Chromatogr.*, 359, 203-212.

Sribney, M and Kennedy, E. P. (1958) *J. Biol. Chem. 233, 1315.*
Vanier, M. T., Rousson, R., Garcia, I., Bailloud, G., Juge, M., Revol, A. and Louisot, P. (1985) *Clin. Genet.* 27, 20.

Tsuji, S., Choudary, P. V., Martin, B. M., Stubblefield, B. K., Mayor, J. A., Barranger, J. A. and Ginns, E. I. (1987) *New Engl. J. Med.* 316, 570-575.

von Figura, K. R., Hasilik, A. L. and Steckel, F. (1984) In Molecular Basis of Lysosomal Storage Disorders. Barranger, J. A., Brady, O. R. ed. Acad Press N.Y., pp. 133-146.

Vellodi, A., Hobbs, J. R., O'donnell, N. M., Coulter, B. S., Hugh-Jones, K. (1987) *Bri. Med. J.* 295, 1375-1376.

Wang, A. M., Schindler, D. and Desnick, R. J. (1990) *J. Clin. Invest.* 86, 1752-1756.

Wang, A. and Desnick, R. J. (1991) *Genomics* 10, 133-142.

Wang, B., Korfhagen, T. R., Gallagher, P. M., D'Amore, M. A., McNeish, J., Potter, S. S. and Ganschow, R. E. (1988) *J. Biol. Chem.* 263, 15841-15844.

Weil, D., D'Alessio, M., Ramirez, F., de Wet, W., Cole, W. G., Chan, D. and Bateman, J. F. (1989) *EMBO J.* 8, 1705-1710.

Wenger, D. A., Sattler, M., Kudoh, T., Snyder, S. P. and Kingston, R. S. (1980) *Science* 208, 1471

Wigderson, M., Firon, N., Horowitz, A., Wilder, S., Frischberg, Y., Reiner, O. and Horowitz, M. (1989) *Am J.Hum.Genet* 44, 365-377.

Wong, G. G., Witek, J. S., Temple, P. A., Wilkens, K. M., Leary, A. C., Luxenberg, D. P., Jones, S. S., Brown, E. L., Kay, R. M., Ore, E. C., Shoemaker, C., Golde, D. W., Kaufman, R. J., Hewick, R. M., Wang, E. A. and Clark, S. C. (1985) *science* 228, 810-815.

Yamada, A., Tsukagoshi, N., Udaka, S., Sasaki, T., Makino, S., Nakamura, S., Little, C., Tomita, M. and Ikezawa, H. (1988) *Eur. J. Biol. Chem.* 175, 213-220.