# Investigating Sandhoff Disease in Saskatchewan

A Thesis

Submitted to the Faculty of Graduate Studies and Research

In partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

in

Biology

University of Regina

Submitted by

Braden Bernard Fitterer

Regina, Saskatchewan

September, 2013

**UNIVERSITY OF REGINA**

**FACULTY OF GRADUATE STUDIES AND RESEARCH**

**SUPERVISORY AND EXAMINING COMMITTEE**

Braden Bernard Fitterer, candidate for the degree of Doctor of Philosophy in Biology**,** has presented a thesis titled, ***Investigating Sandhoff Disease in Saskatchewan,*** in an oral examination held on August 27, 2013. The following committee members have found the thesis acceptable in form and content, and that the candidate demonstrated satisfactory knowledge of the subject material.

| | |
|---|---|
| External Examiner: | *Dr. Don J. Mahuran, University of Toronto |
| Co-Supervisor: | Dr. Christopher Yost, Department of Biology |
| Co-Supervisor: | Dr. Nick Anthony Antonishyn, Adjunct |
| Committee Member: | Dr. John Stavrinides, Department of Biology |
| Committee Member: | Dr. Christopher Somers, Department of Biology |
| Committee Member: | Dr. Tanya Dahms, Department of Chemistry/Biochemistry |
| Chair of Defense: | Dr. Warren Wessel, Faculty of Education |

*participated via SKYPE

## I. Abstract

Sandhoff disease is an autosomal recessive lysosomal storage disease caused by mutations in the *HEXB* gene detrimentally affecting the enzyme β-hexosaminidase. A high incidence of Sandhoff disease has been reported in northern Saskatchewan. The variant of the disease present in the province causes the death of infants before 4 years of age. In order to ameliorate the impact of this disease on the families and communities where it occurs the following steps were taken.

Initially, genetic analysis of the *HEXB* gene from affected patients revealed a common variant shared among 4 individuals. That genetic information was used to develop a diagnostic molecular assay (Chapter 2). A novel synthetic substrate specific for the hydrolytic activity of β-hexosaminidase was obtained and used to develop a biochemical assay for measuring enzyme activity (Chapter 3). Finally, a retrospective study was designed using residual dried blood spots from the Saskatchewan Newborn Screening Program. The two assays were used to screen the dried blood spots for Sandhoff disease carriers in the northern Saskatchewan communities and the frequency of disease-causing alleles was estimated (Chapter 4).

A high carrier frequency for Sandhoff disease causing mutations was found during the retrospective analysis coinciding with the high incidence of Sandhoff disease previously reported. As such the assays developed throughout this project may serve as the basis for a preventative carrier screening program for Sandhoff disease in Saskatchewan.

## II. Acknowledgements

I would like to thank the following individuals for their support and aid in this project; Trisha Hall for her involvement with the biochemical assay, Michael Gelb for supplying the synthetic substrates used in the biochemical assay, Nick Antonishyn for his oversight and guidance throughout the project, Denis Lehotay for making the project possible and supporting conference attendance, Joyce LePage for her technical expertise concerning newborn screening cards, Jeff Eichhorst and Michelle Etter for their knowledge of mass spectrometry, and Greg Horsman for supporting this project at the Saskatchewan Disease Control Laboratory. Finally, I would like to thank my supervisor Chris Yost as well as Chris Somers, John Stavrinides, and Tanya Dahms for their guidance while serving on my graduate committee.

## III. Table of Contents

## IV. List of Tables

## V. List of Figures

## VI. List of Appendices

## VII. List of Abbreviations, Symbols, and Nomenclature

∆Rn: change in normalized reporter signal

%CV: percent coefficient of variation

%HexA: the amount of β-hexosaminidase A activity as a percent of total β-hexosaminidase activity

%HexB: the amount of β-hexosaminidase B activity as a percent of total β-hexosaminidase activity

ACMG: American College of Medical Geneticists

BLAST: basic local alignment search tool

C2: carbon 2

CAD: collision gas

CCDS: consensus coding sequence

CE: collision energy

COS-7 cells: cell line derived from monkey kidney tissue

Ct: threshold cycle

CUR: curtain gas

CXP: collision cell exit potential

DBS: dried blood spot

DNA: deoxyribonucleic acid

DP: declustering potential

EC: enzyme commission number

ELISA: enzyme-linked immuno sorbent assay

EP: entrance potential

FP: focusing potential

G: gravitational force

GalNAc: N-acetyl-galactose or 2-(Acetylamino)-2-deoxy-D-galactose

GA2: N-acetyl-β-D-galactopyranose-1,4-β-D-galactopyranose-1,4-β-D-glucopyranose-1,1-cerimide

GM1: β-D-galactopyranose-N-acetyl-β-D-galactopyranose-1,4-[N-acetyl-α-neuraminidate-2,3-]-β-D-galactopyranose-1,4-β-D-glucopyranose-1,1-cerimide

GM2: N-acetyl-β-D-galactopyranose-1,4-[N-acetyl-α-neuraminidate-2,3-]-β-D-galactopyranose-1,4-β-D-glucopyranose-1,1-cerimide

*GM2A*: the GM2 activator protein gene

GS1 & GS2: ion source gasses (both nitrogen)

H: wild type allele

h: mutant allele

HexA: the β-hexosaminidase A enzyme

*HEXA*: the β-hexosaminidase A gene

HexB: the β-hexosaminidase B enzyme

*HEXB*: the β-hexosaminidase B gene

HexS: the β-hexosaminidase S enzyme

HPLC: high-performance liquid chromatography

IS: ion spray voltage

kb: kilobase

kDa: kilo Dalton

mRNA: messenger ribonucleic acid

MS/MS: tandem mass spectrometry

MUG: 4-methylumbelliferyl-2-acetamido-2-deoxy-β-D-glucopyranoside

MUGS: 4-methylumbelliferyl-6-sulfo-2-acetamido-2-deoxy-β-D-glucopyranoside

NB-DNJ, miglustat, Zavesca®: N-butyldeoxynojirimycin

NB-DGJ: N-butyldeoxygalactonojirimycin

NCBI: National Center for Biotechnology Information

no.: number

nsSNPs: non-synonymous single nucleotide polymorphisms

Rn: normalized reporter signal

RPM: revolutions per minute

Sandhoff mice: (*HEXB* (-/-) mice), mice homozygous for a defective *HEXB* gene

SNP: single nucleotide polymorphism

Tay-Sachs mice: (*HEXA* (-/-) mice), mice homozygous for a defective *HEXA* gene

TE: tris-ethylenediaminetetraacetic acid

TEM: temperature

Total Hex: the total amount of β-hexosaminidase activity

Wt: wild type

## Chapter 1: Introduction

The field of medical genetics is a rapidly expanding area of science devoted to the diagnosis, management, understanding, and treatment of genetic disease [1]. Genetic diseases are deleterious conditions of human health caused by irregularities in a person's DNA. These irregularities can be as small as single nucleotide mutations within a gene or as large as whole chromosome duplications. Some genetic diseases are heritable while others are acquired [1].

A large group of heritable genetic disorders are known as inborn errors of metabolism or may also be referred to as inherited metabolic diseases [2]. The majority of inherited metabolic diseases are due to defects in single genes which encode enzymes responsible for catalyzing the conversion of specific substrates to their respective products [3]. Often the lack of a metabolic enzyme leads to the accumulation of toxic substances in the body but can also result in the lack of an essential compound. Historically, inherited metabolic diseases were classified into disorders of carbohydrate metabolism, amino acid metabolism, organic acid metabolism, or lysosomal storage diseases [3]. As the understanding of cellular metabolism continues to expand more and more inherited metabolic disorders are defined and the classification of such disorders has become more intricate.

In 1955 the lysosome was discovered and shortly after its metabolic role as a cellular recycling center was elucidated [4]. Macromolecules and membrane components are directed to the lysosome for degradation by lysosomal enzymes which function at acidic pH values (~ pH 4.5 - 5) [4]. The monomers of these macromolecules namely amino acids, fatty acids, and monosaccharides and other monomers end up in the cytosol

to be reused in various biosynthetic pathways or catabolised for energy production. Often a condition resulting in the improper activity of a lysosomal enzyme results in the build up of metabolic intermediates. Consequently, disorders resulting in compromised lysosomal function have been coined lysosomal storage diseases. Lysosomal storage diseases are a group of over 50 genetically distinct yet biochemically similar disorders. Disruption of lysosome function is typically due to defects in a gene encoding a lysosomal enzyme, a protein cofactor, lysosomal membrane-transport protein, or a protein involved in the post-translational modification or transport of a lysosomal protein [5].

Sandhoff disease (Online Mendelian Inheritance in Man no. 628800) also known as GM2 gangliosidosis type II, is one such lysosomal storage disorder. The term "ganglioside" originated in the 1930's to 1940's with the discovery that a group of acid glycosphingolipids were the stored compound in the brains of Tay-Sachs disease patients and that these compounds were also present in normal ganglion cells [6]. Sandhoff disease is clinically indistinguishable from Tay-Sachs disease and biochemically very similar which led to misclassification of Sandhoff disease cases as Tay-Sachs disease for many years.

Sandhoff disease was first described in the 1960's while Konrad Sandhoff was investigating lipid storage in post-mortem brain tissue from patients with Tay-Sachs disease [7]. At that time the basis for gangliosidoses were not well understood and the distinction of biochemical differences between Sandhoff disease and Tay-Sachs disease were not yet established. The primary compound (GM2) stored in neuronal cells of Tay-Sachs patients had been identified [8-10] followed by the observation that neuronal cells

from these patients contained storage granules resembling altered lysosomes [11-13]. The discovery of Sandhoff disease came when a suspected case of Tay-Sachs disease not only presented with lysosomal storage of GM2 and GA2 gangliosides but also globoside [7]. The observation was made that all of these storage molecules had terminal β-N-acetylgalactosamine residues and it was further determined that β-N-acetylglucosidases could also hydrolyze the galactosyl moiety shared by these three storage compounds [14]. Two major forms of this enzyme were described, β-hexosaminidase A (HexA; E.C. 3.2.1.52) and β-hexosaminidase B (HexB; E.C. 3.2.1.52) [15, 16]. The case initially thought to be Tay-Sachs disease buth aving the additional storage of globoside was later found to be biochemically distinct from Tay-Sachs disease based on the additional deficiency of β-hexosaminidase B.

### **1.1** *β-hexosaminidases and the GM2 activator protein*

There are three forms of β-hexosaminidase, two are considered major components and one minor. The two major components are β-hexosaminidase A, a heterodimer of α-β subunits and β-hexosaminidase B which is a β-β subunit homodimer (Figure 1.1) [17]. β-hexosaminidase A and β-hexosaminidase B can be separated by ion exchange chromatography [15] or isoelectric focusing [16] which allowed for the initial characterization of these two enzymes. Further insights have come from the X-ray crystallographic structure analysis of β-hexosaminidase A [18] and β-hexosaminidase B [19, 20]. The third and minor form, β-hexosaminidase S exists as an α-α homodimer and may be responsible for minor residual enzyme activity in Sandhoff disease patients [21].

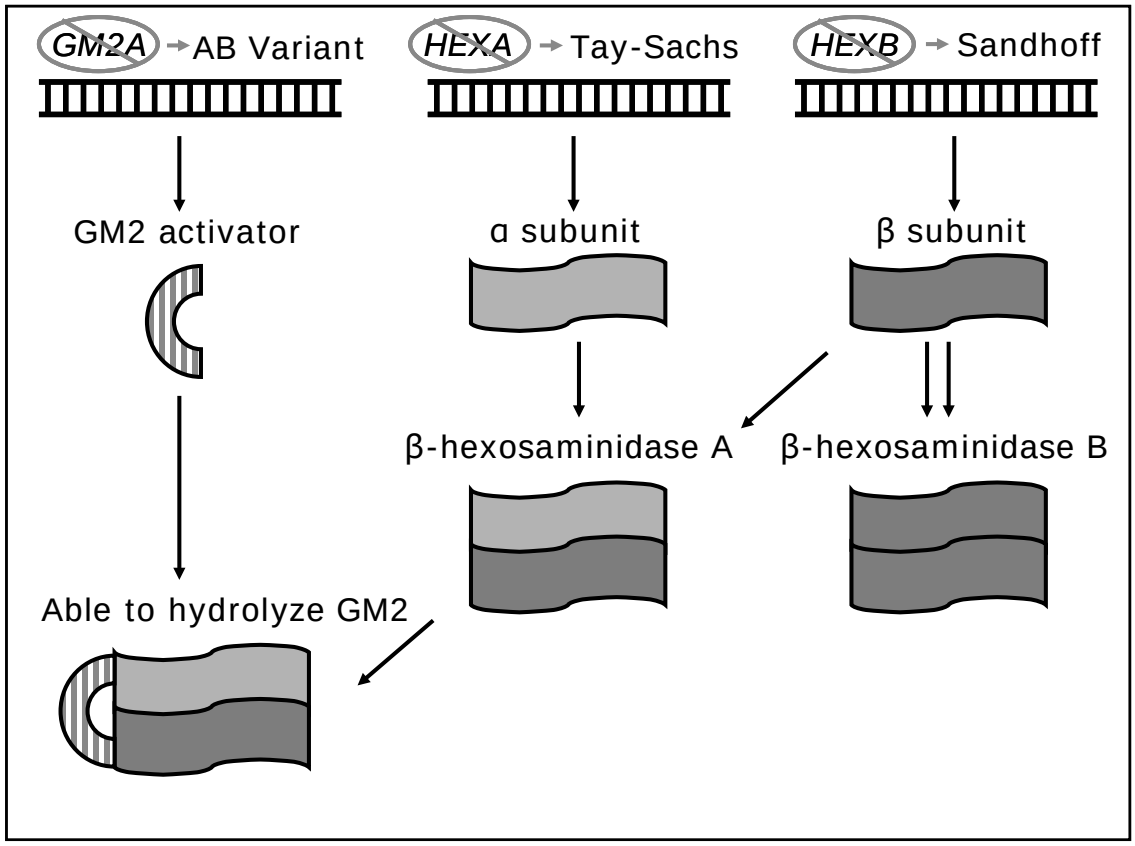



**Figure 1.1**: Association of the *GM2A*, *HEXA*, and *HEXB* genes with the GM2 activator protein, β-hexosaminidase A, and β-hexosaminidase B. Mutations in the *HEXA* gene cause Tay-Sachs disease resulting diminished activity of β-hexosaminidase A. Mutations in the *HEXB* gene cause Sandhoff disease resulting in diminished activity of both β-hexosaminidase A and B. Mutations in the *GM2A* gene cause the AB variant by affecting the GM2 activator protein, preventing it from presenting GM2 to β-hexosaminidase A for hydrolysis.

The α and β subunits of β-hexosaminidase A as well as the GM2 activator protein are transcribed in the endoplasmic reticulum [22] where they are glycosylated at subunit-specific asparagine residues [23]. An N-terminal signal peptide directs the molecules through the endoplasmic reticulum [24]. Vesicles then transport the glycoproteins to the Golgi apparatus where a two-step process results in the transfer of a phosphate to a mannose residue to produce the mannose-6-phosphate moiety [22, 25], a lysosome specific targeting signal which facilitates transport of the enzymes and other lysosomal proteins to the lysosome. Mannose-6-phosphates bind to vesicle-bound mannose-6-phosphate receptors that transfer of the glycoproteins to a pre-lysosomal staging area where they are packaged into lysosomes [22, 25, 26]. Once in the lysosome, the pro-β-hexosaminidase subunits undergo proteolytic and glycosidic modification to produce the active enzyme [27].

The fully processed α subunit includes 2 polypeptide chains ($\alpha_p$: $Leu_{23}$ - $Gly_{73}$, $\alpha_m$: $Thr_{88}$ - $Thr_{528}$) [28-31] covalently linked by 3 disulfide bonds ($Cys_{58}$ – $Cys_{104}$, $Cys_{277}$ – $Cys_{328}$, $Cys_{505}$ – $Cys_{522}$) [32, 33] as well as 3 Asn-linked glycosylation sites ($Asn_{114}$, $Asn_{156}$, $Asn_{294}$) (Figure 1.2 A) [32]. The fully processed β subunit consists of 3 polypeptides ($\beta_p$: $Ala_{50}$ - $Gly_{107}$, $\beta_b$: $Thr_{122}$ - $Ser_{311}$, $\beta_a$: $Lys_{316}$ - $Met_{556}$) [28, 30, 31, 34] joined together by 3 disulphide bonds ($Cys_{93}$ – $Cys_{137}$, $Cys_{309}$-$Cys_{360}$, and $Cys_{534}$-$Cys_{551}$) (Figure 1.2 B) [20, 33]. A total of 4 Asn-linked glycosylation sites have also been identified on β-subunit residues ($Asn_{84}$, $Asn_{142}$, $Asn_{190}$, and $Asn_{327}$) [20, 26, 32, 33].

Furthermore, β-hexosaminidase A was found to be heat labile compared to β-hexosaminidase B [35-38]. Analysis of the β-hexosaminidase A and B crystal structures suggests that the thermo instability of β-hexosaminidase A relative to β-hexosaminidase

B may be due to different amino acid residues at the dimer interface between the two subunits [20].

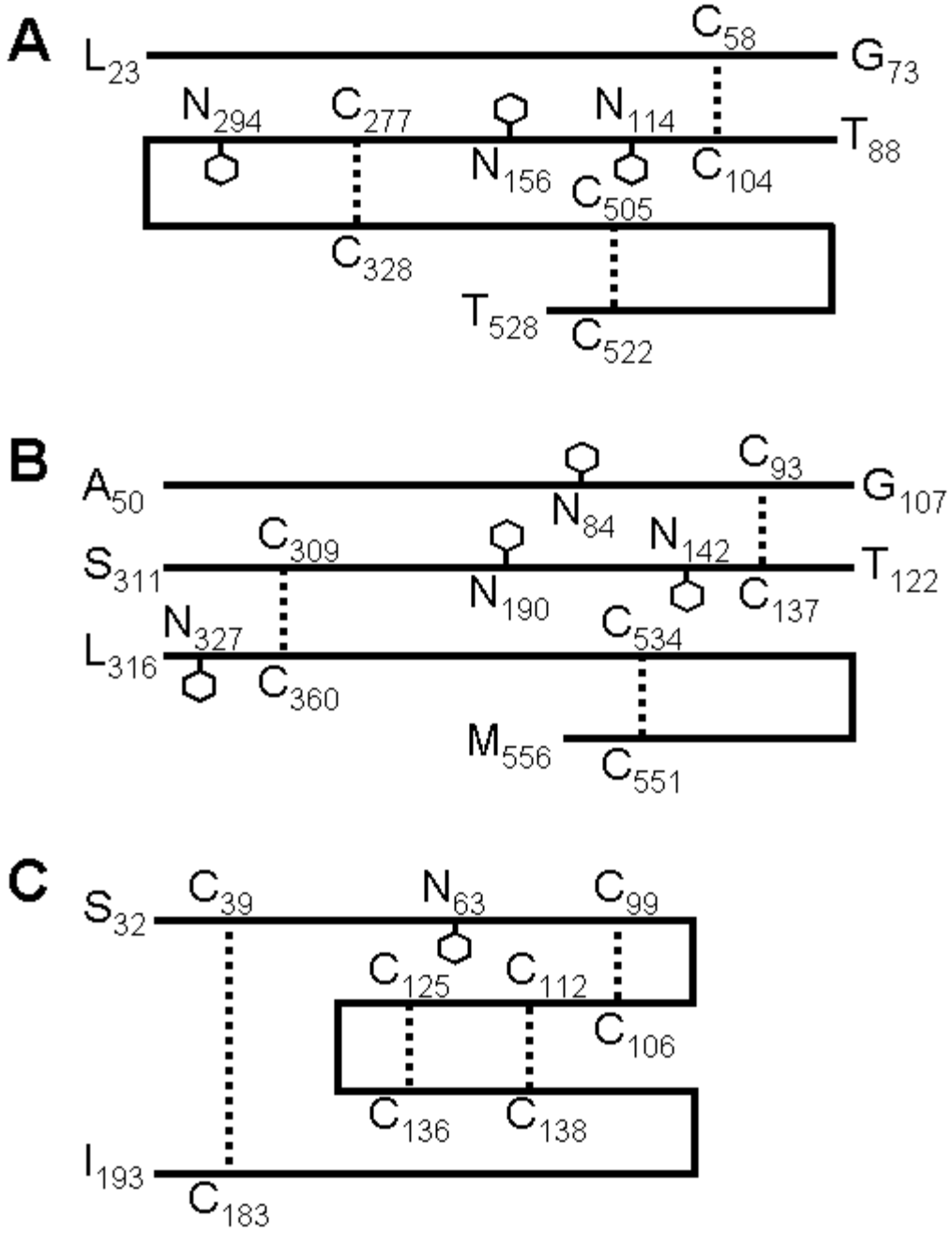



**Figure 1.2** Polypeptide structure for β-hexosaminidase A α-subunit (A), β-hexosaminidase A and B β-subunit (B), and the GM2 activator protein (C). The solid line depicts the polypeptide backbone. Dashed lines depict disulfide bonds. Hexagons show glycosylated asparagine residues.

The key active site residues directly involved in enzymatic catalysis are Glu 323 and Asp 322 from the α subunit and Glu 355 and Asp 354 from the β subunit [18, 20]. Other residues at the active site have also been identified as being important for their role as enzyme-substrate conformational stabilizers [18, 20]. β-Hexosaminidase enzymes hydrolyze terminal β-N-acetylglucosamine and β-N-acetylgalactosamine residues from glycolipids and glycoproteins in which glutamate acts as a general base and aspartate orients the C2 acetamido group into position for nucleophilic attack as well as stabilizing the ionic intermediate [18, 20].

Although each subunit possesses an active site, dimerisation is required for catalytic activity. Several amino acid residues from one subunit complete and stabilize the active site on the adjacent subunit via hydrogen bonding and van der Waals forces [20]. Competition experiments using different substrates have shown that β-hexosaminidase A has two different active sites of which one can hydrolyze neutral and negatively charged substrates and one that can only hydrolyze neutral substrates whereas β-hexosaminidase B can only hydrolyze neutral substrates [39]. This substrate specificity is important to keep in mind when considering the storage compounds in patients with GM2 gangliosidosis. Both β-hexosaminidase A and β-hexosaminidase B are able to hydrolyse the neutral GA2 and globoside substrates stored in neural cells of patients with β-hexosaminidase deficiency. However, only β-hexosaminidase A is able to degrade the negatively charged primary storage molecule GM2 (Figure 1.3) [14, 40] which binds to the positively charged Arg residue located at position 424 of the α subunit [41]. In order to hydrolyze GM2 *in vivo* β-hexosaminidase A requires a third peptide, the GM2 activator protein, which acts as a protein cofactor [42].

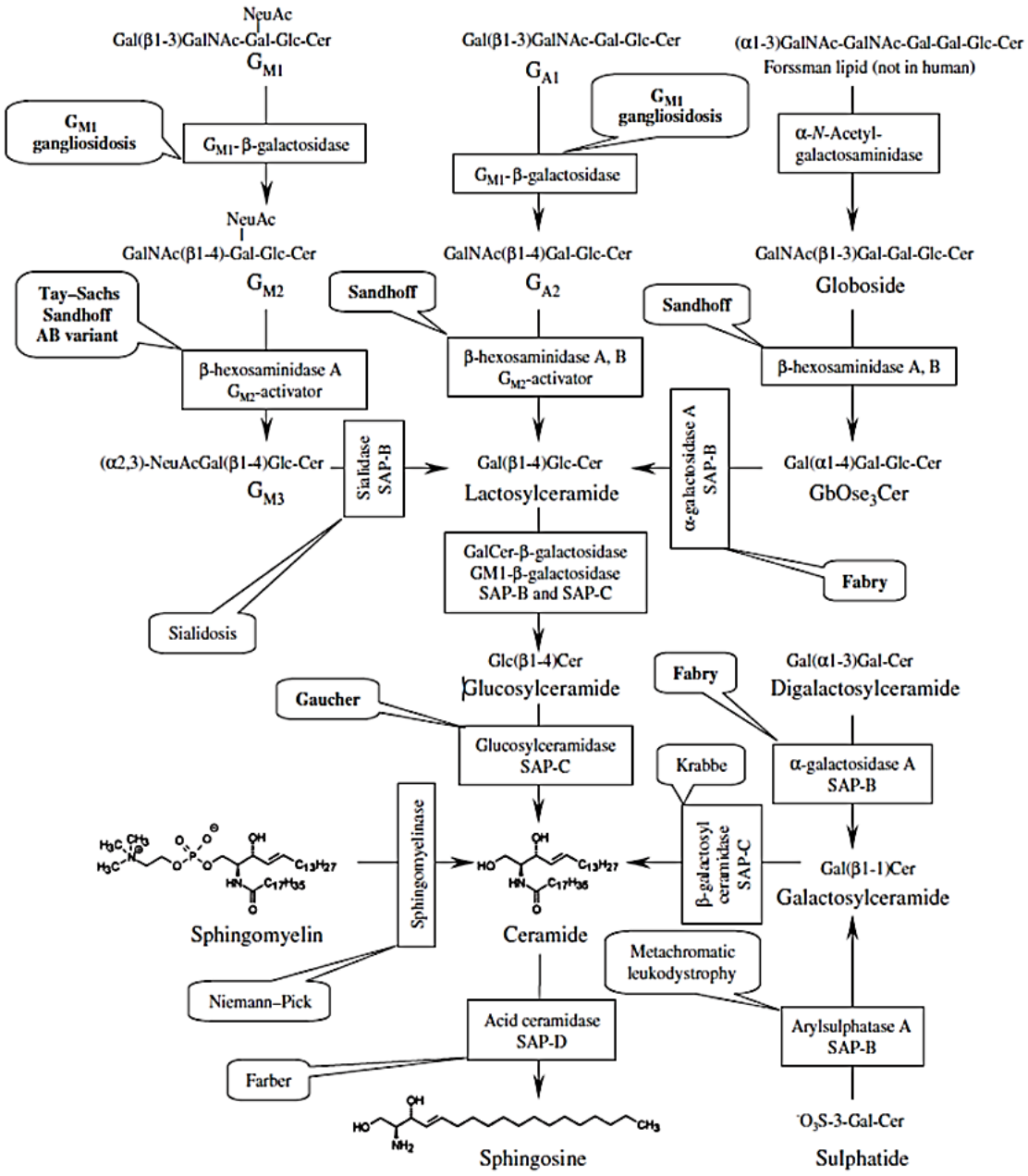



**Figure 1.3**: The role of β-hexosaminidase A and β-hexosaminidase B in lysosomal metabolism and the lysosomal storage diseases caused by deficient enzymes in those pathways. Only β-hexosaminidase A is able to hydrolyze GM2 with the aid of the GM2 activator protein. GA2 and globoside can be hydrolyzed by β-hexosaminidase A or β-hexosaminidase B. Taken with permission from: [43].

The GM2 activator protein is a small (~20kDa) glycolipid binding protein that can form 1:1 complexes with GM2 [44-46]. The mature GM2 activator protein is a single polypeptide ($Ser_{32}$ – $Ile_{193}$) [47] containing 4 disulphide bonds ($Cys_{39}$ – $Cys_{183}$, $Cys_{99}$ – $Cys_{106}$, $Cys_{112}$ – $Cys_{138}$, $Cys_{125}$ – $Cys_{136}$) [48] and one Asn-linked glycosylation site ($Asn_{63}$) (Figure 1.2 C) [49]. The GM2 activator protein crystal structure revealed a substrate binding site [47] in addition to a hydrophobic loop thought to aid the extraction GM2 from membranes at acidic pH [50]. A β-hexosaminidase A binding region has been described on the GM2 activator protein [51] and although the α subunit is required for the catalysis of negatively charged substrates such as GM2, the β-subunit is also required to bind the GM2 activator to β-hexosaminidase A [52].

**1.2** *Molecular basis of GM2 gangliosidosis*

The metabolic errors that lead to GM2 gangliosidosis are predominantly caused by mutations in the genes responsible for producing the α subunit (*HEXA* gene, Tay-Sachs disease, B-variant), the β subunit (*HEXB* gene, Sandhoff disease, 0-variant), and the GM2 activator protein (*GM2A* gene, AB-variant) (Figure 1.1). A rare form of GM2 gangliosidosis has also been described where β-hexosaminidase A retains the ability to hydrolyse neutral substrates but lacks the ability to hydrolyse the negatively charged substrates even in the presence of a functional GM2 activator protein (B1-variant) [39, 53].

The *HEXA* gene comprises 14 exons within a 32,742 base pair length of chromosome 15 (15q23-q24) and transcribes 2.0 and 2.3 kilobase (kb) mRNAs [54]. The *HEXB* gene is distributed across 82,625 base pairs of chromosome 5, contains 14 exons,

is located at position 5q13 and transcribes a 2.2 kb mRNA [31, 55]. Sequence similarity between the α and β subunits of 57% shows an evolutionary relationship between the *HEXA* and *HEXB* genes [54]. The *GM2A* gene is located on a 58,291 base pair section of chromosome 5 (5q33), is made up of 4 exons as well as associated introns and regulatory elements, and transcribes a 2.5 kb mRNA [56].

Due to the location of these genes on non-sex linked chromosomes, GM2 gangliosidosis including Sandhoff disease exhibits an autosomal recessive mode of inheritance. Both parents must carry a disease causing allele in the same gene in order to have an affected child. Subsequently, the chance of two carriers conceiving an affected child is 1 in 4.

**1.3** *Cellular dysfunction*

Several mechanisms have been implicated in the cellular dysfunction caused by lysosomal ganglioside storage. Accumulation of the toxic compound galactosylsphingosine has been described [57] as well as the mechanical distortion of cells due to storage granules such as altered branching of dendrites, nuclei pushed to the outer perimeter of the cell, and the formation of meganeurites (swollen neurons) which may disrupt normal synaptic connections [58, 59]. Demyelination of neural fibers has been observed [60] as well as inflammation of the central nervous system [61]. Neuronal cell death can occur potentially due to apoptosis [59]. Alterations to lipid rafts from ganglioside storage, compromised cell trafficking, and increased degradation of the B cell receptor possibly having immunological consequences have also been observed [62].

Despite the progress made towards understanding gangliosidosis induced cellular dysfunctions, the detrimental effect of GM2 gangliosidosis on neural cells is still not fully understood.

**1.4** *Infantile, juvenile and adult onset and symptoms*

The manifestation of symptoms for GM2 gangliosidosis has been classified into three forms based on the age of onset; infantile, juvenile, or adult onset, and is directly related to the amount of residual β-hexosaminidase activity present [63-65]. β-hexosaminidase A activity in the range of 10-20% [7] of normal is the minimum activity required to avoid symptoms and thus these three variants exist in the range of 0% (infantile variant), up to 0.5% (juvenile variant), and 0.5% up to the minimum threshold (adult onset) [66]. The concept of minimum β-hexosaminidase A activity required to avoid GM2 build up as been referred to as the "critical threshold hypothesis" [65].

Lipid storage in the infantile form of the disease is quite consistent showing very high amounts of ganglioside throughout neural tissue. Juvenile and adult onset forms show a more complicated storage pattern and equally complex aberrant biochemistry throughout different types of neuronal cells possibly due to the different level of tolerance for substrate storage in different cell types [67]. Late-onset forms of the disease have less pronounced lipid storage which is not evenly expressed throughout neuronal tissue but rather concentrated in certain regions of the brain [13, 67].

Symptoms of the infantile form of GM2 gangliosidoses begin to present between 3 to 6 months and continue until demise typically before 4 years of age. Infantile onset symptoms are largely homogeneous and include motor weakness, startled reaction to

sound, early blindness, progressive mental and motor deterioration, enlargement of the head, cherry-red spots in the eye, seizures, shock-like muscle contractions, and frequent respiratory infections [68].

The juvenile onset form of GM2 gangliosidosis typically has an age of onset of ~5 +/- 4 years of age with demise occurring between ~4-40 years of age. Maegawa et al. described 21 new juvenile onset GM2 gangliosidosis cases as well as a summary of 134 case reports from the literature [69]. Based on these 155 cases of the juvenile form of the disorder the more commonly reported symptoms include impaired speech and movement, difficulties with coordination, intellectual deficit, as well as behaviour and psychiatric complications. Less frequent symptoms include deformities of the feet, incontinence, feeding complications, inability to initiate movement or inability to remain motionless, poor weight gain, seizures, muscle wasting, and visual complications [69]. The mean age of onset for the various symptoms occurs between the age of 3 and 10 years with the earliest reported symptom typically being seizures manifesting at a mean age of 3.5 years of age and the latest being muscle wasting presenting at a mean age of 9.6 years of age [69].

The adult onset or chronic form of the disease has been suggested to have an age of onset of ~18 +/- 7 years for a sample of Tay-Sachs disease patients (n = 21) [70] or the 3$^{rd}$ to 4$^{th}$ decade of life for Sandhoff disease patients [71]. Although the late onset form of the disease has a high degree of variability of symptoms between patients, reported symptoms include clumsiness, lack of coordination, sharp eye movement, speech abnormalities, muscle spasms, muscle weakness, depression, muscle atrophy, difficulty swallowing, incontinence, and sensory loss in the extremities [70-80].

Interestingly, a recent study by Sargeant et al. has showed that the progressive neural degradation that occurs in Sandhoff disease may not be associated with the disruption of a developmental process but rather, due to the cumulative effect of ganglioside storage [81]. As such a therapy that removes the excess GM2 storage may be able to partially improve the condition for Sandhoff disease patients.

### **1.5** *Importance of incidence and carrier frequency estimates*

Much effort has been put into trying to determine the prevalence and incidence of genetic diseases in the general and discrete populations. This type of information is not only useful to clinical geneticists, healthcare authorities, and medical laboratories but it is also valuable to research groups looking to develop treatment options for genetic disease. Health care decision makers may also use the information to estimate the social and economic burden that these diseases inflict upon society.

Like other lysosomal storage disorders, Sandhoff disease is distributed globally and extremely rare making it difficult to accurately estimate its carrier frequency. The prevalence of the group of lysosomal storage diseases as a whole in the general population is even quite rare with estimates varying from 1 in 7700 live births in the general population of Australia [82] to 1 in 7100 live births in the Netherlands [83]. Reports of the incidence and carrier frequency for Sandhoff disease causing mutations are discussed fully in chapter 2 and 4.

**1.6** *Therapeutic research*

Research efforts into treatment options for gangliosidoses causing diseases such as Sandhoff disease focus on either restoring the catabolic activity of the dysfunctional β-hexosaminidase enzymes or disrupting the glycosphingolipid biosynthetic pathway to prevent toxic accumulation of the metabolic intermediates such as GM2. When considering treatment options for GM2 gangliosidosis it is important to remember that neuronal storage and degradation occurs, highlighting potential complications for achieving therapeutic effects on both sides of the blood-brain barrier. As such many of the following therapeutic strategies are not only examined alone but also in combination with a therapy that can access the neuronal areas of the body.

Disruption of the glycosphingolipid biosynthetic pathway is known as substrate reduction therapy. Other efforts such as those aiming to restore the catabolic activity of the non-functioning lysosomal enzymes include bone marrow transplantation, the introduction of mammalian artificial chromosomes, stem cell therapy, intravenous enzyme replacement, gene therapy [43], and the use of pharmacological chaperones [84]. Of note for strategies replacing enzyme activity is the ability of cells to internalize lysosomal enzymes and target them to the lysosome via mannose-6-phospate receptors [85]. Enzymes are introduced by various means to a limited number of sites such as the blood stream or cerebrospinal fluid and produce positive effects on the pathogenesis of disease in a broad range of cells and tissues.

The typical model for studying Sandhoff disease is *HEXB* (-/-) mice (Sandhoff mice) [86]. Sandhoff mice unlike the *HEXA* (-/-) (Tay-Sachs mouse) model develop similar ganglioside storage, neurological dysfunction, and disease progression to that

which occurs in humans. Conversely, Tay-Sachs mice show increased ganglioside storage in neurons compared to controls however, they do not develop neurological dysfunction due to a difference in the ganglioside metabolic pathway in mice compared to humans [86]. The reason for this is that the mouse sialidase enzyme is able to convert GM2 (the primary storage molecule in human β-hexosaminidase deficiency) into GA2 which can be degraded by β-hexosaminidase B [86]. Therefore, in the Sanhoff mice which lack both β-hexosaminidase A and B the mice develop symptoms similar to humans whereas in the Tay-Sachs mice sialidase and β-hexosaminidase B are able to bypass the defect associated with β-hexosaminidase A deficiency [86]. As such, Sandhoff mice are a more valid model than Tay-Sachs mice for studying β-hexosaminidase deficiency.

The primary example of substrate reduction therapy is N-butyldeoxynojirimycin (NB-DNJ, miglustat, Zavesca®) which is currently being evaluated for the treatment of late onset Tay-Sachs Disease [87] and N-butyldeoxygalactonojirimycin (NB-DGJ). These imino sugars work by blocking a cerimide glucosyltransferase which is required for the synthesis of GM1 and GM2 thereby reducing the amount of glycolipid synthesized and able to be stored in cells lacking β-hexosaminidase [61, 88-97]. Since Tay-Sachs and Sandhoff diseases are biochemically similar the results of this clinical trial may be broadly applicable to all GM2 gangliosidosis causing disorders including Sandhoff disease.

NB-DNJ has been shown to be effective in reducing the pathogenesis of Sandhoff disease in Sandhoff mice [43]. However, NB-DNJ has also been shown to cause adverse effects such as weight loss and lymphoid organ shrinkage [89, 98]. Trials in Sandhoff

mice treated with NB-DGJ, a close analogue to NB-DNJ, have shown significant reduction in the total levels of ganglioside and more specifically reduced levels of GM2 build up in the brains of Sandhoff mice without producing noticeable adverse effects [89]. Due to the lack of adverse effects and greater administrable dosage, NB-DGJ shows potential for being a more effective substrate reduction therapy for the treatment of Sandhoff disease than its glucose analogue NB-DNJ [89, 98]. Dietary restrictions have also been shown to enhance the action of these imino sugars for reducing lipid storage [99, 100].

Despite the success of substrate reduction therapy in animal models of gangliosidosis, preliminary human trials have had mixed results. Unfortunately, a 3 year trial using NB-DNJ showed little effect for treating a patient with adult onset Sandhoff disease [101] though the progression of some symptoms was halted by the drug in juvenile onset patients [102, 103]. Another report described no measurable gains in adult Tay-Sachs patients [104], while others have observed a reduction of only specific symptoms [105]. The results in adult Tay-Sachs patients are not to be unexpected since progressive accumulation of ganglioside and irreversible cellular disruption may already have taken place in these individuals prior to starting treatment. Further studies with larger sample sizes are needed to better assess the efficacy of these drugs for treating GM2 gangliosidosis.

Furthermore, bone marrow transplantation studies in Sandhoff mice indicate an increased life span of nearly two fold and slowed neural degradation in treated mice compared to control groups [106]. While these results are promising they have been enhanced by combining bone marrow transplantation therapy with substrate reduction

therapy. A combination therapy of bone marrow transplantation and NB-DNJ has led to a significant increase in the lifespan of Sandhoff mice over an untreated group as well as groups treated with bone marrow transplantation or NB-DNJ alone [107].

Intravenous enzyme replacement therapy as a means of restoring missing enzyme activity has been shown to be an effective treatment for several lysosomal storage disorders that primarily affect peripheral organs [108-115]. However, recombinant enzyme injected into the blood stream is unable to cross the blood-brain barrier making the treatment of neurologically compromising lysosomal storage disorders such as Sandhoff and Tay-Sachs disease more difficult. Attempts to inject recombinant enzyme into the cerebrospinal fluid of animal models of several lysosomal storage diseases have had varying success at reducing the levels of stored metabolites in neural tissues [116-121]. Thus the technique was applied to Sandhoff mice with the injection of recombinant human β-hexosaminidase A into the cerebrospinal fluid [122]. Upon receiving only a single injection the lifespan of Sandhoff mice was increased by as much as 12.9% relative to untreated Sandhoff mice [122]. Further animal studies are required however, this technique could provide a treatment for Sandhoff and Tay-Sachs disease requiring patients to receive regular cerebrospinal injections of recombinant enzyme.

Stem cell therapy for Sandhoff disease involves the implantation of neural stem cells into the intracranial space so that these cells may restore β-hexosaminidase activity in neuronal tissues. An obvious advantage of stem cell therapy over enzyme replacement therapy is that eventually once enough new stem cells are present and producing functional enzyme for the patient they may not need further injections, whereas a patient receiving enzyme replacement therapy requires repeated injections for the rest of their

life. Application of this therapy to infant and adult Sandhoff mice has shown that the cells will migrate to the site of neurodegradation, develop into neurons, and delay the onset of symptoms by restoring β-hexosaminidase activity and lowering the amount of stored GM2 and GA2 [123, 124]. The positive effects of this treatment were also accompanied by an increased lifespan of 40% for infantile mice (pre-symptomatic treatment) and 19% in adult mice (symptoms already present) showing proof-in-principle that this treatment may be a viable therapy for those individuals already presenting with Sandhoff disease symptoms at the time of diagnosis. Subsequent efforts into a combination therapy between stem cell therapy and substrate reduction therapy showed no improvement over substrate reduction therapy alone [125].

Preclinical gene therapy trials in mice have also shown promising results for treating Sandhoff disease and other lysosomal storage diseases. Several groups have had success using either adenoviral [126-129] or lentiviral vectors [130]. A novel method for restoring β-hexosaminidase activity in Sandhoff mice using a modified adenovirus vector to deliver the *HEXB* gene specifically to neural cells has shown that neuron specific *HEXB* delivery reverses gangliosidosis and improves the state of neural degradation [129]. Development of this type of therapeutic strategy not only has the potential to treat lysosomal storage diseases such as Sandhoff and Tay-Sachs diseases but also other neurological disorders caused by genetic abnormalities.

Adeno-associated viral vectors have also been used to deliver the human *HEXA* and *HEXB* genes to the brains of Sandhoff mice [126]. Initial work showed an increased lifespan of the Sandhoff mice from 4-5 months (no treatment) to greater than 1 year in treated mice [126]. Injection of adenoviral vector containing the β-hexosaminidase genes

into the cerebrospinal fluidic space rescued Sandhoff disease mice from neuronal degradation however, disease progression in peripheral tissues persisted [127]. As such Sandhoff patients would require injection on both sides of the blood-brain barrier in order to avoid GM2 storage throughout the body. Further studies refining dosage, gene expression, the number of injections and injection sites have increased the life span of Sandhoff mice to as much as 2 years, an increase of 500-600% [127].

Pharmacological chaperones have also been investigated for treating adult-onset GM2 gangliosidosis [84]. For a very select group of mutations in the *HEXA* and *HEXB* genes causing GM2 gangliosidosis the enzyme is produced with aberrant conformation such that a large percentage of the enzyme is retained and degraded in the endoplasmic reticulum leaving only a small amount of β-hexosaminidase able to reach the lysosome. The low levels of lysosomal enzyme obviously results in greatly diminished enzymatic activity. Pharmacological chaperones are essentially small molecules capable of traversing the blood-brain barrier that are able to stabilize the native conformation of β-hexosaminidase enzymes so that they may be transported from the endoplasmic reticulum to the lysosome thereby increasing β-hexosaminidase activity above the critical level. Evaluation of a panel of chaperones applied to Tay-Sachs and Sandhoff disease fibroblasts have shown rescue of β-hexosaminidase activity to as high as 35% of the activity found in normal cells. As such, further refinements to this method may prove it to be a valuable treatment for a specific set of mutations causing protein misfolding providing that the chaperones do not cause significant adverse affects in patients. Furthermore, initial reports from a Phase I/II clinical trial for the use of pyrimethamine as a pharmacological chaperone for adult onset Tay-Sachs and Sandhoff disease have shown

the drug to be well tolerated when administered at low dose however no reports as to the clinical efficacy have been observed in the short study duration [131].

Although there has been some success in treating the various onset forms of Sandhoff disease in the mouse model and regardless of what therapeutic strategy is ultimately used to treat Sandhoff disease in the future it will be important to diagnose patients early so that preventative treatment can occur prior to irreversible neural damage. This point becomes especially important for the infantile variant of Sandhoff disease which manifests very early in life.

### **1.7** *Laboratory diagnostic approaches*

The standard method for detecting GM2 gangliosidosis including Sandhoff disease and Tay-Sachs disease involves adding a synthetic substrate specific for β-hexosaminidase to a crude enzyme preparation and measuring fluorescence from the resulting product. Differentiation between β-hexosaminidase A and β-hexosaminidase B can be achieved either through the use of separate substrates (negatively charged substrate to measure only β-hexosaminidase A and neutral substrate to measure total β-hexosaminidase activity) [132] or by taking advantage of the thermal instability of β-hexosaminidase A relative to β-hexosaminidase B by heating an aliquot to abolish β-hexosaminidase A leaving only β-hexosaminidase B, and measuring total β-hexosaminidase in another aliquot [133, 134].

The common substrates used to assay β-hexosaminidase activity are 4-methylumbelliferyl-2-acetamido-2-deoxy-β-D-glucopyranoside (MUG, neutral substrate) and 4-methylumbelliferyl-6-sulfo-2-acetamido-2-deoxy- β-D-glucopyranoside (MUGS,

negatively charged substrate) [135-137]. In the case of MUG or MUGS the assay measures the enzymes activity towards hydrolyzing the terminal β-N-acetyl-hexosamine residues from the β-N-acetyl-hexosaminides. The measurement is made based on the fluorescence generated from the 4-methylumbelliferone product [138, 139]. The false positive rate for assaying hexosaminidase activity in non-Jewish populations has been estimated at 35% using this substrate due to pseudodeficiency alleles in the *HEXA* gene (R247W and R249W) which prevent hydrolysis of the synthetic substrate but allow for normal hydrolysis of the native GM2 substrate [140].

Sandhoff disease is characterized as presenting with low total β-hexosaminidase (Total Hex) activity and a high % β-hexosaminidase A (% HexA) relative to a normal range due to the involvement of the β subunit produced from the *HEXB* gene in both β-hexosaminidase A and B [133, 141, 142]. Tay-Sachs disease patients will have low β-hexosaminidase A activity and a high % β-hexosaminidase B [133, 141, 142].

Identifying the biochemical defect associated with the AB variant of gangliosidosis (GM2 activator protein deficiency) is more complicated. These patients will present with normal levels of β-hexosaminidase A and B when assayed with the standard synthetic substrates since the GM2 activator protein is not required for their hydrolysis. As such other methods are required to identify GM2 activator deficiency. One such method involves the use of radiolabelled GM2 ganglioside to measure β-hexosaminidase activity in the presence and absence of isolated GM2 activator protein from healthy controls [42]. Isolation of GM2 activator protein from healthy controls can be achieved by heating an enzyme preparation to 60°C thereby abolishing β-hexosaminidase and leaving the GM2 activator protein intact [42]. Samples that receive

the healthy GM2 activator show restored β-hexosaminidase activity indicating a *GM2A* defect.

Typical sample types for detecting Sandhoff disease and other GM2 gangliosidosis disorders include blood, plasma, serum, white blood cells, and cultured human skin cells [133, 143-147]. More recently detection of these disorders has been accomplished using dried blood spots on filter paper [148, 149] to enable retrospective analysis of these disorders using residual dried blood spots from newborns however, carrier detection using dried blood spots has proven to be unreliable [148]. Prenatal diagnosis can also be achieved using amniocentesis to obtain fetal cells [150, 151].

Other substrates for assaying β-hexosaminidase activity include; 4-nitrophenyl-N-acetyl-β-D-glucosaminide which is hydrolysed by β-hexosaminidase to produce the fluorescent product 4-nitrophenol [152-155], sodium-3,3'-dichlorophenylsulfonphthaleinyl-N-acetyl-β-D-glucosaminide which is hydrolyzed to produce a chromogenic (color producing) product [156-159], and radiolabelled GM2 ganglioside for which the radiolabelled N-acetyl-galactosyl product can be separated by HPLC and detected by liquid scintillation counting [40, 160, 161]. An important caveat to remember is that in order to assay β-hexosaminidase A activity using radiolabelled GM2, either recombinant GM2 activator protein or a detergent must be added in order for the enzyme to hydrolyze the negatively charged ganglioside.

Other less common techniques used for the detection of gangliosidoses include the ability to detect aberrant oligosaccharide levels in urine [162], lipid analysis combined with metabolic assays in cultured cells [163-166], the measurement of stored GM2 in tissues and cultured cells using tandem mass spectrometry [167], anti-GM2

antibodies [168], and ELISA [169]. Mutation analysis of the *HEXA* [170], *HEXB* [171, 172], and *GM2A* [173] genes by genetic sequencing of the coding and splicing regions has also been described.

### **1.8** *Sandhoff disease in Saskatchewan and ACMG screening guidelines*

Lowden et al. first described the presence of Sandhoff disease in Saskatchewan in 1978 [174]. Three cases were examined from isolated northern Saskatchewan communities along with extensive serum-based carrier screening in the consanguineous families [174]. The cases described by Lowden et al. and subsequent cases since then in Saskatchewan have been the infantile variant presenting with symptoms around 6 months of age and a life span of less than 4 years. A new Sandhoff disease case was typically being diagnosed roughly every 2 years from these small communities comprising several thousand individuals. As such we set out to investigate the risk of future newborns being fatally affected by Sandhoff disease.

With no intervention more newborns would likely be born with this fatal condition. In order to ameliorate the impact of this disease two approaches for managing it were considered, newborn screening and carrier screening. Newborn screening has been conducted for over 40 years, having started with a single test for phenylketonuria, an inherited genetic disorder which can cause mental retardation if not identified and treated in early childhood [175]. Over the past several decades the phenylketonuria example has served to guide the expansion of newborn screening programs. Currently these programs aim to screen children born with debilitating and sometimes fatal genetic

conditions for which early intervention can lessen the negative impact of the disorder. In many cases medical intervention can allow the patient to develop and live normally.

Recently the American College of Medical Geneticists has recommended that health authorities adhere to a newborn screening panel made up of 29 primary conditions and an additional 25 secondary conditions [176]. The 29 primary conditions were selected from a large list of genetic disorders on the basis of having an available screening test, adequate knowledge of the natural history of the condition, and an effective treatment. Most of the secondary conditions were selected due to their diagnosis being associated with the analysis of a primary condition. The number of disorders included in the panel continues to grow as new diagnostic tests are developed and additional treatment options become available.

It has been suggested that a likely addition to the screening panel assembled by the American College of Medical Geneticists may be the group of disorders known as lysosomal storage diseases [177]. Many of these lysosomal storage diseases can already be successfully treated with therapies such as enzyme replacement therapy [178] whereas no such therapy has been approved for Sandhoff disease.

Considering the American College of Medical Geneticists criteria for selecting a disorder to include in a newborn screening panel and applying those criteria to Sandhoff disease it is apparent that the disease currently falls short of those criteria. Despite the extensive research into treatment options for GM2 gangliosidosis as summarized here, at this time there is no approved method for treating individuals diagnosed with Sandhoff disease. Although there are methods for detecting aberrant β-hexosaminidase activity the

current methods lack high-throughput. In addition, sample collection and sample type are not ideal for screening.

Furthermore, the natural history of Sandhoff disease is well understood at the level of gene, gene product, enzyme, metabolite storage, and symptoms however, the mutations causing the disease in Saskatchewan and the carrier frequency for those mutations is unknown. Therefore newborn screening for Sandhoff disease is inappropriate until a treatment is available, the causative mutations in the province are known, and an appropriate screening assay is in place.

Alternatively, carrier screening for genetic disorders is intended to give genetic information about carrier status to individuals who are considering reproduction so that they may make informed reproductive decisions. The prevention-based technique can be applied prior to conception in order to avoid conceiving an affected child and it can be further supplemented by fetal screening to allow for optional termination of affected pregnancies. The quintessential example of such a carrier screening effort is its application to Tay-Sachs disease in the Ashkenazi Jewish population [179].

The Tay-Sachs disease prevention program was started in 1971 as the first large scale community-based genetic screening program. Either at the time of marriage or prior to conception Jewish couples participating in the program may elect to have biochemical and genetic analysis conducted to determine their Tay-Sachs disease carrier status. Couples who are both carriers are provided with genetic counselling as to the risk of having an affected child as well as being provided with information about the options available to aid in having an unaffected child. Between 1971 and 1999 over 1.4 million individuals elected for Tay-Sachs disease carrier testing which identified nearly 1400 at-

risk couples [180]. Furthermore, by monitoring over 3200 pregnancies 628 affected fetuses were identified resulting in 609 elective abortions [180]. As such the incidence of Tay-Sachs disease in the Ashkenazi Jewish population was reduced by 90% in Canada and the United states between 1971-1999 [180].

Due to the success of the Tay-Sachs disease prevention program a total of 16 recessively inherited genetic disorders are now screened for in the Jewish population with carrier frequencies ranging from 1:15 to 1:168 [181]. Furthermore the success of this program has sparked consideration of preventative screening programs for a number of other genetic diseases in specific populations including cystic fibrosis among Caucasians, sickle cell anemia in African Americans, and β-thalassemia in those of Mediterranean or Asian descent [182]. As such the Tay-Sachs disease prevention program seems to be a good model on which to base a Sandhoff disease prevention program on here in Saskatchewan.

However, prior to the work presented here several knowledge gaps existed precluding the possibility of offering a preventative screening program in Saskatchewan. Molecular and biochemical-based methods suited to the detection of Sandhoff disease in Saskatchewan needed to be established. In order to achieve a high degree of sensitivity and specificity the Tay-Sachs disease prevention program utilized both enzymatic molecular methods for determining an individual's genotype [180]. Since the results of these diagnostic tests play a large role in the decision making process for parents considering termination of an affected pregnancy, highly reliable diagnostic assays are required since incorrect results could lead to the termination of a healthy featus. In this

regard both biochemical and molecular based screening methods were investigated for detecting Sandhoff disease in Saskatchewan.

Although enzymatic assays were already established for the investigation of β-hexosaminidase deficiency, the established assays had several shortcomings rendering their use for screening in Saskatchewan impractical (further discussed in chapter 2 and 3). Furthermore, prior to the work presented here the mutations causing Sandhoff disease in Saskatchewan and the frequency of those mutations were unknown leading us to investigate the mutations present in the population and their frequency while establishing highly reliable molecular- and biochemical-based screening methods.

**1.9** *Study objectives and design*

In order to supply the information needed to propose a preventative screening program in Saskatchewan the following objectives were established:

1. Determine the Sandhoff disease causing mutations present in the northern Saskatchewan population
2. Estimate the carrier frequency for Sandhoff disease causing mutations present in the northern Saskatchewan communities
3. Develop highly reliable high-throughput enzymatic and molecular based screening methods for use in a Sandhoff disease prevention program for Saskatchewan

In order to address these objectives the following tasks were undertaken. Initially genetic analysis of the *HEXB* gene from affected patients revealed a common variant

shared among 4 individuals. That genetic information was used to develop a diagnostic molecular assay (chapter 2). A novel synthetic substrate specific for the hydrolytic activity of β-hexosaminidase was obtained and used to develop a biochemical assay for measuring enzyme activity (chapter 3). Finally, a retrospective study was designed using residual dried blood spots from the Saskatchewan Newborn Screening Program. The two assays were used to screen the dried blood spots for Sandhoff disease carriers in the northern Saskatchewan communities and the frequency of disease causing alleles was estimated (chapter 4).

## 1.10 References


1. Epstein, C.J., *Medical genetics in the genomic medicine of the 21st century.* Am J Hum Genet, 2006. **79**(3): p. 434-8.
2. Garrod, A., *Inborn Errors of Metabolism*, ed. H. Frowde. 1923, London: University of Oxford.
3. Saudubray, J.M., *Inborn errors of metabolism.* Semin Neonatol, 2002. **7**(1): p. 1.
4. de Duve, C., *The lysosome turns fifty.* Nat Cell Biol, 2005. **7**(9): p. 847-9.
5. Winchester, B., A. Vellodi, and E. Young, *The molecular basis of lysosomal storage diseases and their treatment.* Biochem Soc Trans, 2000. **28**(2): p. 150-4.
6. Sandhoff, K., *Tay-Sachs Disease*, R.J. Desnick and M.M. Kaback, Editors. 2001, Academic Press: San Diego CA.
7. Sandhoff, K., *My journey into the world of sphingolipids and sphingolipidoses.* Proc Jpn Acad Ser B Phys Biol Sci, 2012. **88**(10): p. 554-82.
8. Ledeen, R., et al., *Structure Comparison of the Major Monosialogangliosides from Brains of Normal Human, Gargoylism, and Late Infantile Systemic Lipidosis. I.* J Neuropathol Exp Neurol, 1965. **24**: p. 341-51.
9. Makita, A. and T. Yamakawa, *The Glycolipids of the Brain of Tay-Sachs' Disease.--the Chemical Structures of a Globoside and Main Ganglioside.* Jpn J Exp Med, 1963. **33**: p. 361-8.
10. Svennerholm, L., *The chemical structure of normal human brain and Tay-Sachs gangliosides.* Biochem Biophys Res Commun, 1962. **9**: p. 436-41.
11. Samuels, S., et al., *Studies in Tay-Sachs disease. IV. Membranous cytoplasmic bodies.* J Neuropathol Exp Neurol, 1963. **22**: p. 81-97.
12. Terry, R.D. and S.R. Korey, *Membranous cytoplasmic granules in infantile amaurotic idiocy.* Nature, 1960. **188**: p. 1000-2.
13. Terry, R.D. and M. Weiss, *Studies in Tay-Sachs disease. II. Ultrastructure of the cerebrum.* J Neuropathol Exp Neurol, 1963. **22**: p. 18-55.
14. Sandhoff, K. and W. Wassle, *[Enrichment and characterization of 2 forms of human N-acetyl-beta-D-hexosaminidase]*. Hoppe Seylers Z Physiol Chem, 1971. **352**(8): p. 1119-33.
15. Robinson, D. and J.L. Stirling, *N-Acetyl-beta-glucosaminidases in human spleen.* Biochem J, 1968. **107**(3): p. 321-7.
16. Sandhoff, K., *[Resolution of mammalian N-acetyl-beta-D-hexosaminidase into multiple forms by electrofocussing]*. Hoppe Seylers Z Physiol Chem, 1968. **349**(9): p. 1095-8.
17. Beutler, E., W. Kuhl, and D. Comings, *Hexosaminidase isozyme in type O Gm2 gangliosidosis (Sandhoff-Jatzkewitz disease).* Am J Hum Genet, 1975. **27**(5): p. 628-38.
18. Lemieux, M.J., et al., *Crystallographic structure of human beta-hexosaminidase A: interpretation of Tay-Sachs mutations and loss of GM2 ganglioside hydrolysis.* J Mol Biol, 2006. **359**(4): p. 913-29.
19. Maier, T., et al., *The X-ray crystal structure of human beta-hexosaminidase B provides new insights into Sandhoff disease.* J Mol Biol, 2003. **328**(3): p. 669-81.

20. Mark, B.L., et al., *Crystal structure of human beta-hexosaminidase B: understanding the molecular basis of Sandhoff and Tay-Sachs disease.* J Mol Biol, 2003. **327**(5): p. 1093-109.
21. Ikonne, J.U., M.C. Rattazzi, and R.J. Desnick, *Characterization of Hex S, the major residual beta hexosaminidase activity in type O Gm2 gangliosidosis (Sandhoff-Jatzkewitz disease).* Am J Hum Genet, 1975. **27**(5): p. 639-50.
22. Kornfeld, S., *Trafficking of lysosomal enzymes in normal and disease states.* J Clin Invest, 1986. **77**(1): p. 1-6.
23. Kornfeld, R. and S. Kornfeld, *Assembly of asparagine-linked oligosaccharides.* Annu Rev Biochem, 1985. **54**: p. 631-64.
24. Walter, P., R. Gilmore, and G. Blobel, *Protein translocation across the endoplasmic reticulum.* Cell, 1984. **38**(1): p. 5-8.
25. von Figura K Fau - Hasilik, A. and A. Hasilik, *Lysosomal enzymes and their receptors.* (0066-4154 (Print)).
26. Sonderfeld-Fresko, S. and R.L. Proia, *Analysis of the glycosylation and phosphorylation of the lysosomal enzyme, beta-hexosaminidase B, by site-directed mutagenesis.* J Biol Chem, 1989. **264**(13): p. 7692-7.
27. Sagherian, C., et al., *Proteolytic processing of the pro beta chain of beta-hexosaminidase occurs at basic residues contained within an exposed disulfide loop structure.* Biochem Cell Biol, 1993. **71**(7-8): p. 340-7.
28. Hubbes, M., et al., *The amino-terminal sequences in the pro-alpha and -beta polypeptides of human lysosomal beta-hexosaminidase A and B are retained in the mature isozymes.* FEBS Lett, 1989. **249**(2): p. 316-20.
29. Little, L.E., et al., *Proteolytic processing of the alpha-chain of the lysosomal enzyme, beta-hexosaminidase, in normal human fibroblasts.* J Biol Chem, 1988. **263**(9): p. 4288-92.
30. Mahuran, D.J., *Characterization of human placental beta-hexosaminidase I2. Proteolytic processing intermediates of hexosaminidase A.* J Biol Chem, 1990. **265**(12): p. 6794-9.
31. Mahuran, D.J., et al., *Proteolytic processing of pro-alpha and pro-beta precursors from human beta-hexosaminidase. Generation of the mature alpha and beta a beta b subunits.* J Biol Chem, 1988. **263**(10): p. 4612-8.
32. O'Dowd, B.F., et al., *Oligosaccharide structure and amino acid sequence of the major glycopeptides of mature human beta-hexosaminidase.* Biochemistry, 1988. **27**(14): p. 5216-26.
33. Schuette, C.G., J. Weisgerber, and K. Sandhoff, *Complete analysis of the glycosylation and disulfide bond pattern of human beta-hexosaminidase B by MALDI-MS.* Glycobiology, 2001. **11**(7): p. 549-56.
34. Quon, D.V., et al., *Proteolytic processing of the beta-subunit of the lysosomal enzyme, beta-hexosaminidase, in normal human fibroblasts.* J Biol Chem, 1989. **264**(6): p. 3380-4.
35. Carroll, M. and D. Robinson, *Immunological properties of N-acetyl-beta-D-glucosaminidase of normal human liver and of GM2-gangliosidosis liver.* Biochem J, 1973. **131**(1): p. 91-6.
36. Okada, S. and J.S. O'Brien, *Tay-Sachs disease: generalized absence of a beta-D-N-acetylhexosaminidase component.* Science, 1969. **165**(3894): p. 698-700.

37. Srivastava, S.K., et al., *Studies on human beta-D-N-acetylhexosaminidases. II. Kinetic and structural properties.* J Biol Chem, 1974. **249**(7): p. 2049-53.
38. Tallman, J.F., et al., *Isolation and relationship of human hexosaminidases.* J Biol Chem, 1974. **249**(11): p. 3489-99.
39. Kytzia, H.J. and K. Sandhoff, *Evidence for two different active sites on human beta-hexosaminidase A. Interaction of GM2 activator protein with beta-hexosaminidase A.* J Biol Chem, 1985. **260**(12): p. 7568-72.
40. Sandhoff, K., *The hydrolysis of Tay-Sachs ganglioside (TSG) by human N-acetyl-beta-D-hexosaminidase A.* FEBS Lett, 1970. **11**(5): p. 342-344.
41. Sharma, R., et al., *Identification of the 6-sulfate binding site unique to alpha-subunit-containing isozymes of human beta-hexosaminidase.* Biochemistry, 2001. **40**(18): p. 5440-6.
42. Conzelmann, E. and K. Sandhoff, *AB variant of infantile GM2 gangliosidosis: deficiency of a factor necessary for stimulation of hexosaminidase A-catalyzed degradation of ganglioside GM2 and glycolipid GA2.* Proc Natl Acad Sci U S A, 1978. **75**(8): p. 3979-83.
43. Jeyakumar, M., et al., *Glycosphingolipid lysosomal storage diseases: therapy and pathogenesis.* Neuropathol Appl Neurobiol, 2002. **28**(5): p. 343-57.
44. Conzelmann, E., et al., *Complexing of glycolipids and their transfer between membranes by the activator protein for degradation of lysosomal ganglioside GM2.* Eur J Biochem, 1982. **123**(2): p. 455-64.
45. Conzelmann, E. and K. Sandhoff, *Purification and characterization of an activator protein for the degradation of glycolipids GM2 and GA2 by hexosaminidase A.* Hoppe Seylers Z Physiol Chem, 1979. **360**(12): p. 1837-49.
46. Glombitza, G.J., et al., *Biosynthesis, processing, and intracellular transport of GM2 activator protein in human epidermal keratinocytes. The lysosomal targeting of the GM2 activator is independent of a mannose-6-phosphate signal.* J Biol Chem, 1997. **272**(8): p. 5199-207.
47. Wright, C.S., S.C. Li, and F. Rastinejad, *Crystal structure of human GM2-activator protein with a novel beta-cup topology.* J Mol Biol, 2000. **304**(3): p. 411-22.
48. Schutte, C.G., et al., *Complete localization of disulfide bonds in GM2 activator protein.* Protein Sci, 1998. **7**(4): p. 1039-45.
49. Furst, W., et al., *The complete amino-acid sequences of human ganglioside GM2 activator protein and cerebroside sulfate activator protein.* Eur J Biochem, 1990. **192**(3): p. 709-14.
50. Wendeler, M., et al., *Photoaffinity labelling of the human GM2-activator protein. Mechanistic insight into ganglioside GM2 degradation.* Eur J Biochem, 2004. **271**(3): p. 614-27.
51. Wendeler, M., et al., *The enzyme-binding region of human GM2-activator protein.* FEBS J, 2006. **273**(5): p. 982-91.
52. Meier, E.M., et al., *The human GM2 activator protein. A substrate specific cofactor of beta-hexosaminidase A.* J Biol Chem, 1991. **266**(3): p. 1879-87.
53. Kytzia, H.J., et al., *Variant of GM2-gangliosidosis with hexosaminidase A having a severely changed substrate specificity.* Embo J, 1983. **2**(7): p. 1201-5.

54. Korneluk, R.G., et al., *Isolation of cDNA clones coding for the alpha-subunit of human beta-hexosaminidase. Extensive homology between the alpha- and beta-subunits and studies on Tay-Sachs disease.* J Biol Chem, 1986. **261**(18): p. 8407-13.
55. O'Dowd, B.F., et al., *Isolation of cDNA clones coding for the beta subunit of human beta-hexosaminidase.* Proc Natl Acad Sci U S A, 1985. **82**(4): p. 1184-8.
56. Xie, B., et al., *Isolation and expression of a full-length cDNA encoding the human GM2 activator protein.* Biochem Biophys Res Commun, 1991. **177**(3): p. 1217-23.
57. Igisu, H. and K. Suzuki, *Progressive accumulation of toxic metabolite in a genetic leukodystrophy.* Science, 1984. **224**(4650): p. 753-5.
58. Purpura, D.P. and K. Suzuki, *Distortion of neuronal geometry and formation of aberrant synapses in neuronal storage disease.* Brain Res, 1976. **116**(1): p. 1-21.
59. Walkley, S.U., *Cellular pathology of lysosomal storage disorders.* Brain Pathol, 1998. **8**(1): p. 175-93.
60. Tifft, C.J. and R.L. Proia, *The beta-hexosaminidase deficiency disorders: development of a clinical paradigm in the mouse.* Ann Med, 1997. **29**(6): p. 557-61.
61. Jeyakumar, M., et al., *Central nervous system inflammation is a hallmark of pathogenesis in mouse models of GM1 and GM2 gangliosidosis.* Brain, 2003. **126**(Pt 4): p. 974-87.
62. te Vruchte, D., et al., *Glycosphingolipid storage leads to the enhanced degradation of the B cell receptor in Sandhoff disease mice.* J Inherit Metab Dis, 2010. **33**(3): p. 261-70.
63. Conzelmann, E. and K. Sandhoff, *Partial enzyme deficiencies: residual activities and the development of neurological disorders.* Dev Neurosci, 1983. **6**(1): p. 58-71.
64. Conzelmann, E. and K. Sandhoff, *Biochemical basis of late-onset neurolipidoses.* Dev Neurosci, 1991. **13**(4-5): p. 197-204.
65. Leinekugel, P., et al., *Quantitative correlation between the residual activity of beta-hexosaminidase A and arylsulfatase A and the severity of the resulting lysosomal storage disease.* Hum Genet, 1992. **88**(5): p. 513-23.
66. Dreyfus, J.C., et al., *Characterization of a variant of beta-hexosaminidase: "hexosaminidase Paris".* Am J Hum Genet, 1977. **29**(3): p. 287-93.
67. Escola, J., *[On the spreading of the amaurotic idiocy processes in the central nervous system at various ages and on features of a late form with regard to pigment atrophy]*. Arch Psychiatr Nervenkr Z Gesamte Neurol Psychiatr, 1961. **202**: p. 95-112.
68. *NINDS Sandhoff Disease Information Page*. Office of Communications and Public Liaison, National Institute of Neurological Disorders and Stroke. National Institutes of Health 2007 [cited 2009 August 19]; Available from: http://www.ninds.nih.gov/disorders/sandhoff/sandhoff.htm.
69. Maegawa, G.H., et al., *The natural history of juvenile or subacute GM2 gangliosidosis: 21 new cases and literature review of 134 previously reported.* Pediatrics, 2006. **118**(5): p. e1550-62.

70. Neudorfer, O., et al., *Late-onset Tay-Sachs disease: phenotypic characterization and genotypic correlations in 21 affected patients.* Genet Med, 2005. **7**(2): p. 119-23.
71. Delnooz Cc Fau - Lefeber, D.J., et al., *New cases of adult-onset Sandhoff disease with a cerebellar or lower motor neuron phenotype.* (1468-330X (Electronic)).
72. Bolhuis, P.A., et al., *Molecular basis of an adult form of Sandhoff disease: substitution of glutamine for arginine at position 505 of the beta-chain of beta-hexosaminidase results in a labile enzyme.* Biochim Biophys Acta, 1993. **1182**(2): p. 142-6.
73. Gomez-Brouchet, A., et al., *Adult Sandhoff disease presenting with chronic diarrhoea.* J Inherit Metab Dis, 1999. **22**(2): p. 187-8.
74. Gomez-Lira, M., et al., *A common beta hexosaminidase gene mutation in adult Sandhoff disease patients.* Hum Genet, 1995. **96**(4): p. 417-22.
75. Hara, A., et al., *Adult Sandhoff's disease: R505Q and I207V substitutions in the HEXB gene of the first Japanese case.* J Neurol Sci, 1998. **155**(1): p. 86-91.
76. McInnes, B., et al., *An unusual splicing mutation in the HEXB gene is associated with dramatically different phenotypes in patients from different racial backgrounds.* J Clin Invest, 1992. **90**(2): p. 306-14.
77. Modigliani, R., et al., *Diarrhea and autonomic dysfunction in a patient with hexosaminidase B deficiency (Sandhoff disease).* Gastroenterology, 1994. **106**(3): p. 775-81.
78. Oonk, J.G., H.J. van der Helm, and J.J. Martin, *Spinocerebellar degeneration: hexosaminidase A and B deficiency in two adult sisters.* Neurology, 1979. **29**(3): p. 380-4.
79. Redonnet-Vernhet, I., et al., *Significance of two point mutations present in each HEXB allele of patients with adult GM2 gangliosidosis (Sandhoff disease) homozygosity for the Ile207-->Val substitution is not associated with a clinical or biochemical phenotype.* Biochim Biophys Acta, 1996. **1317**(2): p. 127-33.
80. Schnorf, H., et al., *[Adult form of GM2-gangliosidosis: a man and 2 sisters with hexosaminidase-A and -B deficiency (Sandhoff disease) and literature review].* Schweiz Med Wochenschr, 1996. **126**(18): p. 757-64.
81. Sargeant, T.J., et al., *Characterization of inducible models of Tay-Sachs and related disease.* PLoS Genet, 2012. **8**(9): p. e1002943.
82. Meikle, P.J., et al., *Prevalence of lysosomal storage disorders.* JAMA, 1999. **281**(3): p. 249-54.
83. Poorthuis, B.J., et al., *The frequency of lysosomal storage diseases in The Netherlands.* Hum Genet, 1999. **105**(1-2): p. 151-6.
84. Tropak, M.B., et al., *Pharmacological enhancement of beta-hexosaminidase activity in fibroblasts from adult Tay-Sachs and Sandhoff Patients.* J Biol Chem, 2004. **279**(14): p. 13478-87.
85. Ghosh, P., N.M. Dahms, and S. Kornfeld, *Mannose 6-phosphate receptors: new twists in the tale.* Nat Rev Mol Cell Biol, 2003. **4**(3): p. 202-12.
86. Sango, K., et al., *Mouse models of Tay-Sachs and Sandhoff diseases differ in neurologic phenotype and ganglioside metabolism.* Nat Genet, 1995. **11**(2): p. 170-6.

87. Lachmann, R.H., *Miglustat. Oxford GlycoSciences/Actelion.* Curr Opin Investig Drugs, 2003. **4**(4): p. 472-9.
88. Andersson, U., et al., *Improved outcome of N-butyldeoxygalactonojirimycin-mediated substrate reduction therapy in a mouse model of Sandhoff disease.* Neurobiol Dis, 2004. **16**(3): p. 506-15.
89. Baek, R.C., et al., *N-butyldeoxygalactonojirimycin reduces brain ganglioside and GM2 content in neonatal Sandhoff disease mice.* Neurochem Int, 2008. **52**(6): p. 1125-33.
90. Elliot-Smith, E., et al., *Beneficial effects of substrate reduction therapy in a mouse model of GM1 gangliosidosis.* Mol Genet Metab, 2008. **94**(2): p. 204-11.
91. Jeyakumar, M., et al., *Delayed symptom onset and increased life expectancy in Sandhoff disease mice treated with N-butyldeoxynojirimycin.* Proc Natl Acad Sci U S A, 1999. **96**(11): p. 6388-93.
92. Kasperzyk, J.L., et al., *Substrate reduction reduces gangliosides in postnatal cerebrum-brainstem and cerebellum in GM1 gangliosidosis mice.* J Lipid Res, 2005. **46**(4): p. 744-51.
93. Platt, F.M. and M. Jeyakumar, *Substrate reduction therapy.* Acta Paediatr Suppl, 2008. **97**(457): p. 88-93.
94. Platt, F.M., et al., *New developments in treating glycosphingolipid storage diseases.* Adv Exp Med Biol, 2005. **564**: p. 117-26.
95. Platt, F.M., et al., *Substrate reduction therapy in mouse models of the glycosphingolipidoses.* Philos Trans R Soc Lond B Biol Sci, 2003. **358**(1433): p. 947-54.
96. Platt, F.M., et al., *Prevention of lysosomal storage in Tay-Sachs mice treated with N-butyldeoxynojirimycin.* Science, 1997. **276**(5311): p. 428-31.
97. Ranes, M.K., et al., *N -butyldeoxynojirimycin reduces growth and ganglioside content of experimental mouse brain tumours.* Br J Cancer, 2001. **84**(8): p. 1107-14.
98. Andersson, U., et al., *N-butyldeoxygalactonojirimycin: a more selective inhibitor of glycosphingolipid biosynthesis than N-butyldeoxynojirimycin, in vitro and in vivo.* Biochem Pharmacol, 2000. **59**(7): p. 821-9.
99. Denny, C.A., et al., *Restricted ketogenic diet enhances the therapeutic action of N-butyldeoxynojirimycin towards brain GM2 accumulation in adult Sandhoff disease mice.* J Neurochem, 2010. **113**(6): p. 1525-35.
100. Denny, C.A., et al., *Influence of caloric restriction on motor behavior, longevity, and brain lipid composition in Sandhoff disease mice.* J Neurosci Res, 2006. **83**(6): p. 1028-38.
101. Masciullo, M., et al., *Substrate reduction therapy with miglustat in chronic GM2 gangliosidosis type Sandhoff: results of a 3-year follow-up.* J Inherit Metab Dis, 2010.
102. Maegawa, G.H., et al., *Substrate reduction therapy in juvenile GM2 gangliosidosis.* Mol Genet Metab, 2009. **98**(1-2): p. 215-24.
103. Tallaksen, C.M. and J.E. Berg, *Miglustat therapy in juvenile Sandhoff disease.* J Inherit Metab Dis, 2009. **32 Suppl 1**: p. S289-93.

104. Shapiro, B.E., et al., *Miglustat in late-onset Tay-Sachs disease: a 12-month, randomized, controlled clinical study with 24 months of extended treatment.* Genet Med, 2009. **11**(6): p. 425-33.
105. Bembi, B., et al., *Substrate reduction therapy in the infantile form of Tay-Sachs disease.* Neurology, 2006. **66**(2): p. 278-80.
106. Norflus, F., et al., *Bone marrow transplantation prolongs life span and ameliorates neurologic manifestations in Sandhoff disease mice.* J Clin Invest, 1998. **101**(9): p. 1881-8.
107. Jeyakumar, M., et al., *Enhanced survival in Sandhoff disease mice receiving a combination of substrate deprivation therapy and bone marrow transplantation.* Blood, 2001. **97**(1): p. 327-9.
108. Barton, N.W., et al., *Replacement therapy for inherited enzyme deficiency--macrophage-targeted glucocerebrosidase for Gaucher's disease.* N Engl J Med, 1991. **324**(21): p. 1464-70.
109. Clarke, L.A., et al., *Long-term efficacy and safety of laronidase in the treatment of mucopolysaccharidosis I.* Pediatrics, 2009. **123**(1): p. 229-40.
110. Eng, C.M., et al., *Safety and efficacy of recombinant human alpha-galactosidase A--replacement therapy in Fabry's disease.* N Engl J Med, 2001. **345**(1): p. 9-16.
111. Grabowski, G.A., et al., *Dose-response relationships for enzyme replacement therapy with imiglucerase/alglucerase in patients with Gaucher disease type 1.* Genet Med, 2009. **11**(2): p. 92-100.
112. Klinge, L., et al., *Enzyme replacement therapy in classical infantile pompe disease: results of a ten-month follow-up study.* Neuropediatrics, 2005. **36**(1): p. 6-11.
113. Schiffmann, R., et al., *Infusion of alpha-galactosidase A reduces tissue globotriaosylceramide storage in patients with Fabry disease.* Proc Natl Acad Sci U S A, 2000. **97**(1): p. 365-70.
114. Van den Hout, J.M., et al., *Long-term intravenous treatment of Pompe disease with recombinant human alpha-glucosidase from milk.* Pediatrics, 2004. **113**(5): p. e448-57.
115. Wraith, J.E., et al., *Enzyme replacement therapy for mucopolysaccharidosis I: a randomized, double-blinded, placebo-controlled, multinational study of recombinant human alpha-L-iduronidase (laronidase).* J Pediatr, 2004. **144**(5): p. 581-8.
116. Chang, M., et al., *Intraventricular enzyme replacement improves disease phenotypes in a mouse model of late infantile neuronal ceroid lipofuscinosis.* Mol Ther, 2008. **16**(4): p. 649-56.
117. Dickson, P., et al., *Intrathecal enzyme replacement therapy: successful treatment of brain disease via the cerebrospinal fluid.* Mol Genet Metab, 2007. **91**(1): p. 61-8.
118. Dodge, J.C., et al., *Intracerebroventricular infusion of acid sphingomyelinase corrects CNS manifestations in a mouse model of Niemann-Pick A disease.* Exp Neurol, 2009. **215**(2): p. 349-57.
119. Kakkis, E., et al., *Intrathecal enzyme replacement therapy reduces lysosomal storage in the brain and meninges of the canine model of MPS I.* Mol Genet Metab, 2004. **83**(1-2): p. 163-74.

120. Lee, W.C., et al., *Single-dose intracerebroventricular administration of galactocerebrosidase improves survival in a mouse model of globoid cell leukodystrophy.* FASEB J, 2007. **21**(10): p. 2520-7.
121. Pastores, G.M., *Laronidase (Aldurazyme): enzyme replacement therapy for mucopolysaccharidosis type I.* Expert Opin Biol Ther, 2008. **8**(7): p. 1003-9.
122. Tsuji, D., et al., *Highly phosphomannosylated enzyme replacement therapy for GM2 gangliosidosis.* Ann Neurol, 2011. **69**(4): p. 691-701.
123. Jeyakumar, M., et al., *Neural stem cell transplantation benefits a monogenic neurometabolic disorder during the symptomatic phase of disease.* Stem Cells, 2009. **27**(9): p. 2362-70.
124. Lee, J.P., et al., *Stem cells act through multiple mechanisms to benefit mice with neurodegenerative metabolic disease.* Nat Med, 2007. **13**(4): p. 439-47.
125. Arthur, J.R., et al., *Therapeutic effects of stem cells and substrate reduction in juvenile Sandhoff mice.* Neurochem Res, 2012. **37**(6): p. 1335-43.
126. Cachon-Gonzalez, M.B., et al., *Effective gene therapy in an authentic model of Tay-Sachs-related diseases.* Proc Natl Acad Sci U S A, 2006. **103**(27): p. 10373-8.
127. Cachon-Gonzalez, M.B., et al., *Gene transfer corrects acute GM2 gangliosidosis--potential therapeutic contribution of perivascular enzyme flow.* Mol Ther, 2012. **20**(8): p. 1489-500.
128. Guidotti, J.E., et al., *Adenoviral gene therapy of the Tay-Sachs disease in hexosaminidase A-deficient knock-out mice.* Hum Mol Genet, 1999. **8**(5): p. 831-8.
129. Terashima, T., et al., *DRG-targeted helper-dependent adenoviruses mediate selective gene delivery for therapeutic rescue of sensory neuronopathies in mice.* J Clin Invest, 2009. **119**(7): p. 2100-112.
130. Kyrkanides, S., et al., *Intraperitoneal inoculation of Sandhoff mouse neonates with an HIV-1 based lentiviral vector exacerbates the attendant neuroinflammation and disease phenotype.* J Neuroimmunol, 2007. **188**(1-2): p. 39-47.
131. Clarke, J.T., et al., *An open-label Phase I/II clinical trial of pyrimethamine for the treatment of patients affected with chronic GM2 gangliosidosis (Tay-Sachs or Sandhoff variants).* Mol Genet Metab, 2011. **102**(1): p. 6-12.
132. Inui, K. and D.A. Wenger, *Usefulness of 4-methylumbelliferyl-6-sulfo-2-acetamido-2-deoxy-beta-D-glucopyrano sid e for the diagnosis of GM2 gangliosidoses in leukocytes.* Clin Genet, 1984. **26**(4): p. 318-21.
133. Hechtman, P. and F. Kaplan, *Tay-Sachs disease screening and diagnosis: evolving technologies.* DNA Cell Biol, 1993. **12**(8): p. 651-65.
134. Kaback, M.M., et al., *Automated thermal fractionation of serum hexosaminidase: effects of alteration in reaction variables and implications for Tay-Sachs disease heterozygote screening.* Prog Clin Biol Res, 1977. **18**: p. 197-212.
135. Kresse, H., et al., *Liberation of N-acetylglucosamine-6-sulfate by human beta-N-acetylhexosaminidase A.* J Biol Chem, 1981. **256**(24): p. 12926-32.
136. Leaback, D.H. and P.G. Walker, *Studies on glucosaminidase. 4. The fluorimetric assay of N-acetyl-beta-glucosaminidase.* Biochem J, 1961. **78**: p. 151-6.

137. Potier, M., et al., *Fluorometric assay of neuraminidase with a sodium (4-methylumbelliferyl-alpha-D-N-acetylneuraminate) substrate.* Anal Biochem, 1979. **94**(2): p. 287-96.
138. Sandhoff, K., E. Conzelmann, and H. Nehrkorn, *Specificity of human liver hexosaminidases A and B against glycosphingolipids GM2 and GA2. Purification of the enzymes by affinity chromatography employing specific elution.* Hoppe Seylers Z Physiol Chem, 1977. **358**(7): p. 779-87.
139. Srivastava, S.K., et al., *Studies on human beta-D-N-acetylhexosaminidases. I. Purification and properties.* J Biol Chem, 1974. **249**(7): p. 2043-8.
140. Sutton, V.R., *Tay-Sachs disease screening and counseling families at risk for metabolic disease.* Obstet Gynecol Clin North Am, 2002. **29**(2): p. 287-96.
141. Ben-Yoseph, Y., et al., *Diagnosis and carrier detection of Tay-Sachs disease: direct determination of hexosaminidase A using 4-methylumbelliferyl derivatives of beta-N-acetylglucosamine-6-sulfate and beta-N-acetylgalactosamine-6-sulfate.* Am J Hum Genet, 1985. **37**(4): p. 733-40.
142. Vallance, H., et al., *Common HEXB polymorphisms reduce serum HexA and HexB enzymatic activities, potentially masking Tay-Sachs disease carrier identification.* Mol Genet Metab, 2006. **87**(2): p. 122-7.
143. Isaksson, A., et al., *Enzyme immunoassay of beta-hexosaminidase A and B in serum: carrier detection of GM2-gangliosidoses, and equivalence of enzyme activity and enzyme protein reactivity.* Clin Chem, 1993. **39**(7): p. 1412-5.
144. Natowicz, M.R. and E.M. Prence, *Heterozygote screening for Tay-Sachs disease: past successes and future challenges.* Curr Opin Pediatr, 1996. **8**(6): p. 625-9.
145. O'Brien, J.S., et al., *Tay-Sachs Disease.* New England Journal of Medicine, 1970. **283**(1): p. 15-20.
146. Okada, S., et al., *Ganglioside GM2 storage diseases: hexosaminidase deficiencies in cultured fibroblasts.* Am J Hum Genet, 1971. **23**(1): p. 55-61.
147. Suzuki, K., *Enzymatic diagnosis of sphingolipidoses.* Methods Enzymol, 1987. **138**: p. 727-62.
148. Chamoles, N.A., et al., *Tay-Sachs and Sandhoff diseases: enzymatic diagnosis in dried blood spots on filter paper: retrospective diagnoses in newborn-screening cards.* Clin Chim Acta, 2002. **318**(1-2): p. 133-7.
149. Civallero, G., et al., *Twelve different enzyme assays on dried-blood filter paper samples for detection of patients with selected inherited lysosomal storage diseases.* Clin Chim Acta, 2006. **372**(1-2): p. 98-102.
150. Callahan, J.W., et al., *First trimester prenatal diagnosis of Tay-Sachs disease using the sulfated synthetic substrate for hexosaminidase A.* Clin Biochem, 1990. **23**(6): p. 533-6.
151. Grabowski, G.A., et al., *First-trimester prenatal diagnosis of Tay-Sachs disease.* Am J Hum Genet, 1984. **36**(6): p. 1369-78.
152. Chatterjee, S., L.F. Velicer, and C.C. Sweeley, *Glycosphingolipid glycosyl hydrolases and glycosidases of synchronized human KB cells.* J Biol Chem, 1975. **250**(13): p. 4972-9.
153. Marciniak, J., et al., *Optimization of an enzymatic method for the determination of lysosomal N-acetyl-beta-D-hexosaminidase and beta-glucuronidase in synovial fluid.* Clin Chem Lab Med, 2006. **44**(8): p. 933-7.

154. Walker, P.G., J.W. Woollen, and R. Heyworth, *Studies on glucosaminidase. 5. Kidney N-acetyl-beta-glucosaminidase and N-acetyl-beta-galactosaminidase.* Biochem J, 1961. **79**: p. 288-94.
155. Zwierz, K., et al., *The degradation of glycoconjugates in the human gastric mucous membrane.* Acta Med Acad Sci Hung, 1981. **38**(2): p. 145-52.
156. Casal, J.A., E. Cano, and J.C. Tutor, *Beta-hexosaminidase isoenzyme profiles in serum, plasma, platelets and mononuclear, polymorphonuclear and unfractionated total leukocytes.* Clin Biochem, 2005. **38**(10): p. 938-42.
157. Casal, J.A., et al., *Thermodynamic study of beta-N-acetylhexosaminidase enzyme heterogeneity in human seminal plasma.* Clin Chim Acta, 2005. **355**(1-2): p. 55-60.
158. Perez, L.F., et al., *Thermodynamic characterisation of the mutated isoenzyme A of beta-N-acetylhexosaminidase in GM2-gangliosidosis B1 variant.* Clin Chim Acta, 1999. **285**(1-2): p. 45-51.
159. Perez, L.F. and J.C. Tutor, *Assay of beta-N-acetylhexosaminidase isoenzymes in different biological specimens by means of determination of their activation energies.* Clin Chem, 1998. **44**(2): p. 226-31.
160. O'Brien, J.S., et al., *Ganglioside GM2 N-acetyl-beta-D-galactosaminidase and asialo GM2 (GA2) N-acetyl-beta-D-galactosaminidase; studies in human skin fibroblasts.* Clin Genet, 1977. **11**(3): p. 171-83.
161. Svennerholm, L., et al., *The assay of sphingolipid hydrolases in white blood cells with labelled natural substrates.* Clin Chim Acta, 1979. **92**(1): p. 53-64.
162. Clements, P.R., *Determination of sialylated and neutral oligosaccharides in urine by mass spectrometry.* Curr Protoc Hum Genet, 2012. **Chapter 17**: p. Unit17 10.
163. Klein, A., et al., *Sphingolipid activator protein D (sap-D) stimulates the lysosomal degradation of ceramide in vivo.* Biochem Biophys Res Commun, 1994. **200**(3): p. 1440-8.
164. Liu, Y., et al., *A genetic model of substrate deprivation therapy for a glycosphingolipid storage disorder.* J Clin Invest, 1999. **103**(4): p. 497-505.
165. Sango, K., et al., *Mice lacking both subunits of lysosomal beta-hexosaminidase display gangliosidosis and mucopolysaccharidosis.* Nat Genet, 1996. **14**(3): p. 348-52.
166. Yamashita, T., et al., *Enhanced insulin sensitivity in mice lacking ganglioside GM3.* Proc Natl Acad Sci U S A, 2003. **100**(6): p. 3445-9.
167. Tsui, Z.C., et al., *A method for profiling gangliosides in animal tissues using electrospray ionization-tandem mass spectrometry.* Anal Biochem, 2005. **341**(2): p. 251-8.
168. Tsuji, D., et al., *Specific induction of macrophage inflammatory protein 1-alpha in glial cells of Sandhoff disease model mice associated with accumulation of N-acetylhexosaminyl glycoconjugates.* J Neurochem, 2005. **92**(6): p. 1497-507.
169. Tsuji, D., et al., *Therapeutic evaluation of GM2 gangliosidoses by ELISA using anti-GM2 ganglioside antibodies.* Clin Chim Acta, 2007. **378**(1-2): p. 38-41.
170. Triggs-Raine, B.L., et al., *Sequence of DNA flanking the exons of the HEXA gene, and identification of mutations in Tay-Sachs disease.* Am J Hum Genet, 1991. **49**(5): p. 1041-54.

171. Zampieri, S., et al., *Molecular and functional analysis of the HEXB gene in Italian patients affected with Sandhoff disease: identification of six novel alleles.* Neurogenetics, 2009. **10**(1): p. 49-58.
172. Zhang, Z.X., et al., *Impact of premature stop codons on mRNA levels in infantile Sandhoff disease.* Hum Mol Genet, 1994. **3**(1): p. 139-45.
173. Schepers, U., et al., *Molecular analysis of a GM2-activator deficiency in two patients with GM2-gangliosidosis AB variant.* Am J Hum Genet, 1996. **59**(5): p. 1048-56.
174. Lowden, J.A., et al., *Carrier detection in Sandhoff disease.* Am J Hum Genet, 1978. **30**(1): p. 38-45.
175. Wilcken, B. and V. Wiley, *Newborn screening.* Pathology, 2008. **40**(2): p. 104-15.
176. ACMG, *Newborn screening: toward a uniform screening panel and system.* Genet Med, 2006. **8 Suppl 1**: p. 1S-252S.
177. Marsden, D. and H. Levy, *Newborn screening of lysosomal storage disorders.* Clin Chem, 2010. **56**(7): p. 1071-9.
178. Ohashi, T., *Enzyme replacement therapy for lysosomal storage diseases.* Pediatr Endocrinol Rev, 2012. **10 Suppl 1**: p. 26-34.
179. Kaback, M.M., T.J. Nathan, and S. Greenwald, *Tay-Sachs disease: heterozygote screening and prenatal diagnosis--U.S. experience and world perspective.* Prog Clin Biol Res, 1977. **18**: p. 13-36.
180. Kaback, M.M., *Population-based genetic screening for reproductive counseling: the Tay-Sachs disease model.* Eur J Pediatr, 2000. **159 Suppl 3**: p. S192-5.
181. Scott, S.A., et al., *Experience with carrier screening and prenatal diagnosis for 16 Ashkenazi Jewish genetic diseases.* Hum Mutat, 2010.
182. Andrews, L., *Assessing Genetic Risks:Implications for Health and Social Policy*, ed. L.B. Andrews, et al. 1994: The National Academies Press.

# Chapter 2: A Polymerase Chain Reaction-Based Genotyping Assay for Detecting a Novel Sandhoff Disease-Causing Mutation

The contents of this chapter have been peer-reviewed and published by Fitterer et al. in 2012 [1]. Written permission has been obtained from the publisher for inclusion of the publication in this thesis.

## 2.1 Abstract

Sandhoff disease is a rare genetic disorder, however, some northern Saskatchewan communities have a high incidence of the disease (for which the causative mutation has not been described). We discovered a novel mutation causing Sandhoff disease in this community and validated a molecular assay to detect the mutant allele.

DNA sequencing was used to search for mutations in the *HEXB* gene from the most recently affected patient. A polymerase chain reaction (PCR)-based genotyping assay was subsequently designed and validated to detect a novel single-nucleotide deletion using DNA isolated from newborn screening cards. The c.115delG mutation was found in exon 1 of the *HEXB* gene from 4 patients with clinical presentation of Sandhoff disease. Herein we describe a novel *HEXB* mutation that is shared among 4 patients with Sandhoff disease, as well as a validated PCR-based genotyping assay that can reliably detect the mutant allele. Because the 4 patients from this community share a common c.115delG mutation in the coding region of the *HEXB* gene, it may be possible to offer an effective preventative screening program for Sandhoff disease using this assay.

## 2.2 Introduction

Sandhoff disease (Online Mendelian Inheritance in Man no. 268800) is a rare autosomal recessive lysosomal storage disorder caused by mutations in the hexosaminidase-B (*HEXB*) gene (5q13). Mutations in *HEXB* result in a deficiency of the β-hexosaminidase A (HexA; E.C. 3.2.1.52) and β-hexosaminidase B (HexB; E.C. 3.2.1.52) β subunit [2]. HexA is a heterodimer of α-β subunits, and HexB is a β-β homodimer [3]. A third form β-hexosaminidase-S (HexS; E.C. 3.2.1.52) may also be present as a homodimer of α-α subunits [4]. At the time of this writing over thirty *HEXB* mutations resulting in deficient hexosaminidase activity have been described [5]; which lead to gangliosidoses, the accumulation of GM2 ganglioside in cells of the central nervous system, and progressive neural degradation [2]. Severe Sandhoff disease causing mutations result in substantial loss of enzymatic function, leading to symptoms that become apparent during infancy [2]. The infantile form of the disease typically causes death by 2-3 years of age. The incidence of Sandhoff disease among the general population has been estimated at 1 in 422,000 births, with a carrier frequency of 1 in 310 [6]; however, high incidences have been reported in certain groups residing in isolation or in areas with a high degree of consanguinity [7, 8]. Communities in northern Saskatchewan, Canada, are known to have a high incidence of Sandhoff disease [9], although the causative mutation has never been described.

The current fluorometric method used to diagnose Sandhoff disease in Saskatchewan [10] requires a patient's family to travel up to 600 km to the testing facility in order to provide fresh blood samples. A convenient alternative would be to diagnose Sandhoff disease using dried blood spots (DBSs). DBSs can be collected in remote

communities and mailed to the testing facility, thereby alleviating the need for patient travel. Furthermore, in many cases the DBSs collected for the newborn screening program can be analyzed so that further sample collection can be avoided. Enzymatic diagnosis of Sandhoff disease from DBS has been described [11] and is effective for detecting affected patients. However, it lacks the ability to reliably distinguish carriers. Thus, we sought to identify the mutation causing Sandhoff disease in northern Saskatchewan and develop a polymerase chain reaction (PCR)-based assay that could be used to confirm the diagnosis of affected individuals and detect carriers using DNA isolated from DBSs.

## 2.3 Materials and methods

### **2.31** *Primers and probes*

The reference sequence for chromosome 5 (NT_006713.15) was obtained from GenBank and used to design primers and probes for sequencing and for the allelic discrimination assay (Tables 2.1 and 2.2 respectively). The oligonucleotide designs were accomplished using Primer Express software, version 3.0 (Applied Biosystems, Carlsbad, CA). Primers were synthesized by Sigma Genosys (Sigma-Aldrich, St. Louis, MO). Probes were synthesized by Applied Biosystems.

Table 2.1. Primers used for amplification and sequencing of the HEXB gene

| Oligo | Sequence | Size | Location |
|---|---|---|---|
| HEXB-1FW[2] | GGCAGACCGGGCGGAAAGCAG | 425 | Promoter |
| HEXB-1RV[2] | TGCGCAGTGGGGTGGTGAGGG | | Intron 1 |
| HEXB_Ex2_F | TGGGCAGCATGGATTTGAGGAGT | 305 | Intron 1 |
| HEXB_Ex2_R | GTCCGTCAGCGAGCACCTGG | | Intron 2 |
| HEXB_Ex3_F | AGTGTGCGGAGGAGGGAGGG | 379 | Intron 2 |
| HEXB_Ex3_R | GGCGGCAAAGTTTTGCTGTGCT | | Intron 3 |
| HEXB_Ex4_F | TGCCTTACCTGGTTATGAGTCTGTTTG | 265 | Intron 3 |
| HEXB_Ex4_R | ACCGGCTAAGACAAATATCTGGGGA | | Intron 4 |
| HEXB_Ex5_F | TCCCCAGATATTTGTCTTAGCCGGT | 446 | Intron 4 |
| HEXB_Ex5_R | GGGGACCTCCACGTCACAGC | | Intron 5 |
| HEXB_Ex6_F | CCAGAGGCTTTTTGTATACTGATGTGTCA | 871 | Intron 5 |
| HEXB_Ex6_R | TCAGGGCTTCTACCTTGTTAAATTTCGATT | | Intron 6 |
| HEXB-7FW[1] | ACAATTTCCAGGATCAAATCTACG | 272 | Intron 6 |
| HEXB-7RV[2] | ACTCCAGCCTGGGTGACAGAACAA | | Intron 7 |
| HEXB-8FW[2] | GGCAAAGAGACAGGATTCAGGA | 1534 | Intron 7 |
| HEXB-9RV[2] | TGGAAAATACTTGGAGTCACCA | | Intron 9 |
| HEXB_Ex9_F | GTCTGCACAACTGATGTTAGGCATGT | 328 | Intron 8 |
| HEXB_Ex9_R | AGCAAGCAGTGGGTATTGCCTCC | | Intron 9 |
| HEXB_Ex10_F | TGACAGGCTAGATCGTGGTCCCT | 519 | Intron 9 |
| HEXB_Ex10_R | TGCAAATTCCCAGTTGCCCCA | | Intron 10 |
| HEXB_Ex11_F | TGTAGTTAGTGACCACTTTGGACCTCA | 556 | Intron 10 |
| HEXB_Ex11_R | ACCATCAGCTTCACAGTAACCCACA | | Intron 11 |
| HEXB_Ex12_13_F | AGGTCCTGCTAACCACGGGCA | 669 | Intron 11 |
| HEXB_Ex12_13_R | AGCAACTCAAGATGGAGTTCTAAGTTACAC | | Intron 13 |
| HEXB_Ex14_F | TCCATCTTGAGTTGCTTTAATTTTCTTCCCT | 550 | Intron 13 |
| HEXB_Ex14_R | GGCACTGGAGCCACCCGTTT | | 3’ untranslated region |

[1]Previously described by Zhang et al. [12]
[2]Previously described by Zampieri et al. [13]

Table 2.2. Primers and probes for allelic discrimination assay

| Oligo | Sequence and Labels | Final conc. |
|---|---|---|
| HEXB 1bf Forward | GCTGGCGGCGATGTTG | 0.9µM |
| HEXB 2bf Reverse | GCTGTGGCTGATGTAGAAGTTCTC | 0.9µM |
| HEXB Wt | [VIC]–TGGTGCAG**G**TGGCGGA-[Q,MGB][1] | 0.25µM |
| HEXB c.115delG | [NED]-TGGTGCAGTGGCGGA –[Q, MGB] | 0.25µM |

[1]Bold G is the site of the single nucleotide deletion.

**2.32** *DNA extraction and quantification*

Genomic DNA was isolated from DBSs on newborn screening cards. All cards were collected less than 1 year before testing and were stored at room temperature. Two 3-mm disks from each card were punched into 96-well plates and mixed with 250 µl of Tris-ethylenediaminetetraacetic acid (TE) buffer (pH, 8.0) for 30 minutes at room temperature before the DNA extraction. The extraction was then carried out by using the DNA Mini-prep kit (Promega, Madison, WI) with a 200 µl elution volume. The magnetic particle extraction procedure was automated on the BioSprint 96 (Qiagen, Venlo, the Netherlands). Genomic DNA was quantified by using the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, Carlsbad, CA) on the Applied Biosystems 7500 Fast instruments.

**2.33** *Real-time PCR*

A TaqMan genotyping assay was developed to differentiate between the wild type (H) and mutant (h) *HEXB* alleles (see Table 2.2 for oligo sequences) using the QuantiTect Probe PCR Master Mix (Qiagen). Total reaction volume for the assay was 25 µl. The PCR was carried out on Applied Biosystems 7500 Fast Real-Time PCR instruments using SDS software, version 1.4. Each run included homozygous mutant (h/h), heterozygous (H/h), homozygous wild type (H/H), and no template controls. Thermal cycling conditions consisted of 95°C for 15 min, followed by 45 cycles of 94°C for 1 min and 65°C for 30 sec. Pre- and post-reads of each PCR plate were taken in addition to capturing real-time PCR data. The SDS version 1.4 allelic discrimination template was used to analyze the pre- and post-read fluorescence and to assign genotype calls on the basis of the ratio of change in normalized reporter signal (∆Rn) for each

probe. Outliers that were not automatically called as one of the three possible genotypes were marked as undetermined and repeated. Extracts that produced an undetermined result were quantified, and the amount of template added to the genotyping PCR was increased to within the optimal range of input DNA for the assay. The *HEXB* gene from persistent outliers was sequenced.

### **2.34** *Sequencing*

Exons 1 through 14 were amplified and sequenced using the oligonucleotides listed in Table 2.1. Each exon was first amplified by PCR, followed by purification with ExoSAP-IT (USB, Cleveland OH) according to the manufacture's protocol. DNA sequencing was performed in both forward and reverse directions using the ABI BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) according to the manufacturer's instructions. Purification of cycle sequencing reactions used the BigDye Xterminator Purification Kit (Applied Biosystems). Purified cycle sequencing reactions were analyzed on an Applied Biosystems 3130xl Genetic Analyzer. Sequence data was viewed using BioNumerics 6.5 (Applied Maths).

### **2.35** *Validation*

The validation procedure was based largely on the protocol described in the Clinical Laboratory Validation of Multiplex Nucleic Acid Assays document "MM17-A: Verification of Multiplex Nucleic Acid Assays" [14]. Performance specifications, including the optimal range of input DNA, accuracy, precision, and analytical specificity, were evaluated for the genotyping assay.

The optimal range of input DNA for the assay was determined by running 5 replicates of a serial dilution series for each genotype and performing regression analysis on the curves generated for threshold cycle vs. nanograms of input DNA in Excel® (Microsoft). A panel of 37 samples consisting of all three possible genotypes was used to determine accuracy by comparing the genotyping assay results with sequencing data for the same panel.

Repeatability was assessed by analyzing the mean and standard deviation of ΔRn for each probe on 5 replicates of DNA extracted from one, two, or three DBS disks repeated three times in one day. Only the H/H genotype was used to assess the repeatability of the entire assay due to a lack of h/h and H/h sample material. Post-extraction repeatability was examined for all three genotypes and included 5 replicates of high-, medium-, and low- concentration genomic DNA relative to the optimal range. The post-extraction repeatability experiment was also repeated three times in one day. Reproducibility was examined by assaying each genotype and a negative control on fifteen independent runs using different kit lot numbers and three different 7500 fast real-time PCR instruments. Mean and standard deviation for ΔRn were also used to quantify reproducibility.

Spiking experiments were used to assess the possibility of PCR inhibition. A sufficient number of samples were spiked with no matrix, low concentration, and high concentration of matrix to examine the possibility of inhibition with 95 % confidence. Real-time fluorescence data were used to investigate the specificity of the probes for their respective target alleles in the multiplex reaction. Amplicon from the TaqMan assay was sequenced to ensure amplification of only the target region of the *HEXB* gene. The

possibility of carry-over contamination during the automated DNA extraction was assessed by running the real-time PCR assay with blanks interspersed throughout a full plate of DBS extracts. Finally, 246 newborn screening cards were assayed to determine the expected range of ΔRn values from clinical samples.

### **2.36** *Ethics approval*

Ethics approval for the use of residual DBSs to be used in the development and validation of this assay was obtained from the Research Ethics Board at the University of Regina and the Bio-medical Research Ethics Board at the University of Saskatchewan.

## 2.4 Results

### **2.41** *Sequencing*

The *HEXB* genes from the most recent patients born with Sandhoff disease in Saskatchewan and the child's mother were sequenced. A single nucleotide deletion of the guanine residue at position 115 in the coding region of exon one was detected in both alleles of the *HEXB* gene from the child (Figure 2.1) and only a single allele from the mother. The c.115delG mutation was then confirmed in three additional cases of Sandhoff disease from northern Saskatchewan. The sequence variation was named according to the Human Genome Variation Society guidelines [15].

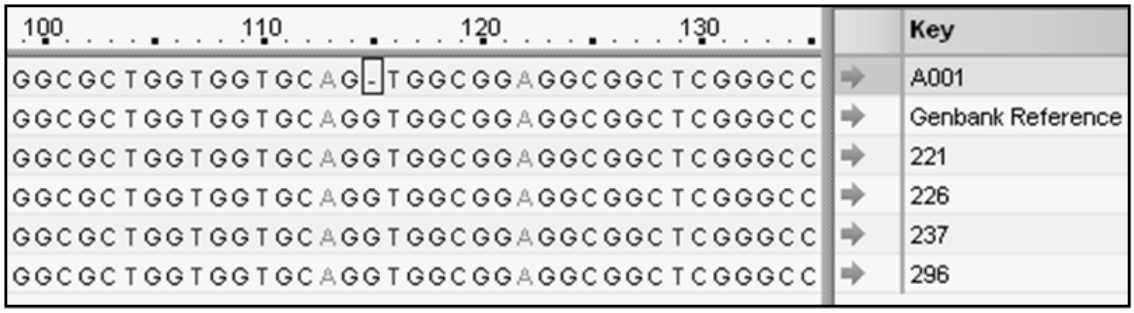



**Figure 2.1.** The site of the single nucleotide deletion detected in exon one of the coding region in the *HEXB* gene. (A001) A patient possessing two c.115delG alleles. (GeneBank Reference) The GeneBank reference sequence NT_006713.15 used for comparison. (221, 226, 237, 296) Healthy patient controls.

**2.42** *Validation*

The optimal range of input DNA for the *HEXB* genotyping assay was 0.25-120 ng of genomic DNA per 25 µl reaction. $R^2$ values of the curves of crossing threshold (Ct) vs. DNA input for each genotype were greater than 0.99 within the optimal range (Figure 2.2). Increasing the DNA input to 250 ng per reaction produced $R^2$ values of less than 0.99. Accuracy of the assay was shown to be 100 % (95 % confidence interval, 90.5 % - 100 %).

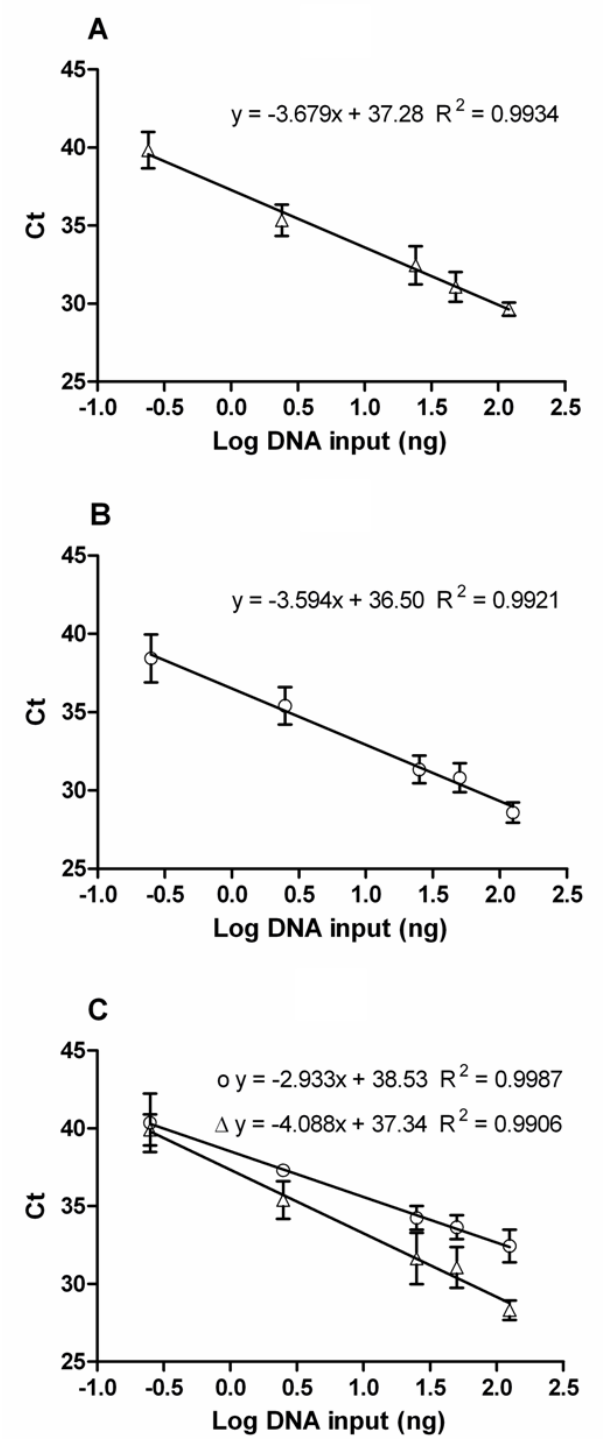



**Figure 2.2.** Linear regression for threshold cycle (Ct) vs. log DNA input. Points are mean +/- standard deviation for 5 replicates of serially diluted DNA ranging from 0.25 to 125 ng. (A) H-allele reporter with H/H template. (B) h-allele reporter with h/h template. (C) H-allele reporter with H/h template (△) and h-allele reporter with H/h template (o).

Assay repeatability for the H/H genotype produced mean +/- standard deviation ΔRn values for the H allele reporter ranging from 0.146 +/- 0.023 to 0.244 +/- 0.015 (Figure 2.3A). Additional repeatability and reproducibility experiments resulted in low variability, as can be seen in the genotype clustering of Figure 2.3B and 2.3C, respectively. The genotype for all 246 clinical samples was H/H (Figure 2.3D).

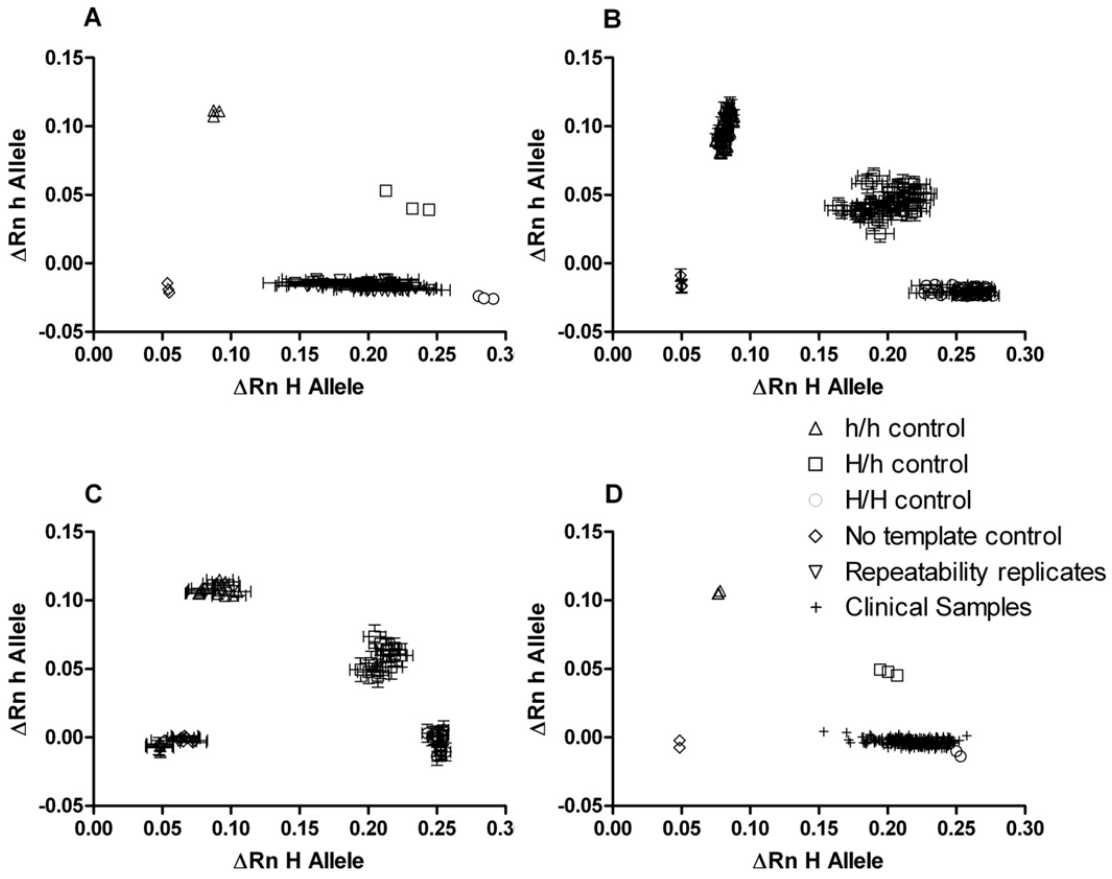



**Figure 2.3.** Precision study for the *HEXB* allelic discrimination assay. (A) Assay repeatability. Only the H/H genotype was included due to a lack of H/h and h/h sample material. Five replicates of one, two, and three 3 mm DBS disks were repeated three times in one day. (B) Post-extraction repeatability. Five replicates of each genotype assayed at high, medium, and low DNA input were repeated three times in one day. (C) Reproducibility of each genotype assayed on fifteen independent runs. (D) The distribution of 246 clinical samples. ΔRn, change in normalized reporter signal.

Spiking experiments showed no statistically significant inhibitory effects on PCR due to the extraction procedure. Cross-reactivity was observed between the H-specific probe and the h/h template (Figure 2.4). However, the amount of nonspecific binding of the H-allele probe to the mutant allele was insufficient to interfere with accurate genotype calling. No cross-reactivity was observed between the mutant probe and wild-type template. No carry-over contamination was observed in any of the blank extractions. The amount of genomic DNA assayed from the clinical samples ranged from 8.96 to 61.88 ng as a result of the variable amount of DNA present on two 3-mm DBS disks. The $\triangle$Rn values for the H allele reporter during clinical sampling ranged from 0.153 to 0.258 (mean +/- standard deviation of 0.219 +/- 0.0171).

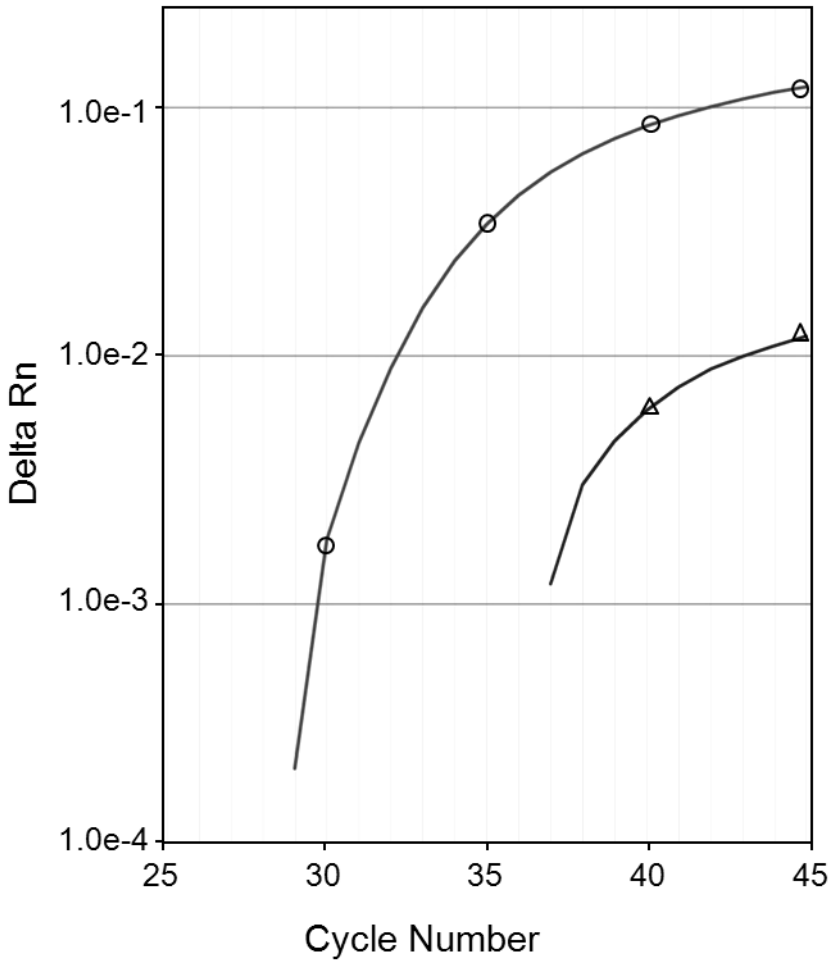



**Figure 2.4.** Non-specific binding of the Wt probe to h/h template. NED® reporter signal from the mutant probe (∆) relative to the VIC® reporter signal from the wild-type probe (o) generated as a result of nonspecific binding of the wild-type probe to homozygous mutant template. Rn, normalized reporter signal.

## 2.5 Discussion

The purpose of this investigation was to identify the mutation causing Sandhoff disease in northern Saskatchewan and develop a PCR-based assay that could be used to confirm a diagnosis of the disease using DNA isolated from DBSs. We have shown that Sandhoff disease in northern Saskatchewan has been caused by a single nucleotide deletion in the coding region of the *HEXB* gene and we have also validated a TaqMan genotyping assay for detecting the mutant allele. Our genetic sequencing results revealed the c.115delG mutation in the *HEXB* gene from 4 patients clinically and biochemically diagnosed with Sandhoff disease from the northern community. Validation of the genotyping assay has proven it to be accurate and precise. Accuracy was shown by comparing results from the genotyping assay to genetic sequencing data. Evidence of precise genotype calling was demonstrated by the distinct genotype clustering in our repeatability and reproducibility studies. The variable amount of DNA able to be isolated from the DBS disks was the largest factor affecting repeatability. However, this variable can be controlled by quantifying the amount of DNA in each extract and standardizing the amount of template presented to the assay. Despite the variable amount of DNA isolated from DBSs, the fluorescence data from 2 µl of extract were consistently sufficient to allow for automated allele calling when analyzing clinical samples.

Furthermore, the reading frame shift caused by the c.115delG mutation in exon one prevents proper transcription and translation of more than 93% of the β-subunit required to form the HexA and HexB enzymes. The effect that this mutation has on the expected gene products is consistent with the clinical outcome of patients with Sandhoff disease from the northern region. Similar to the case studies presented by Lowden and

colleagues in the late 1970s [9], the patients with Sandhoff disease from the last decade have all presented with infantile onset and die of the disease by 2-3 years of age.

Currently there is no treatment for Sandhoff disease aside from supportive care. This has led health authorities to seek a means of prevention in order to alleviate the burden that this disease inflicts upon families and health care systems [8]. Preventative programs often focus on premarital or prenatal screening. Ribeiro and colleagues have found that the allelic discrimination methodologies such as the one described here are rapid, cost-effective, and highly reproducible [16]. In this northern community it appears that the c.115delG mutation is predominantly responsible for Sandhoff disease. As such, the assay we present could serve as an effective screening tool for a program aimed at preventing Sandhoff disease in Saskatchewan. However, further understanding of the mutant allele frequency and the extent to which the allele has disseminated throughout northern Saskatchewan would be required. In summary, we present a novel mutation in the *HEXB* gene causing Sandhoff disease in Saskatchewan as well as a reliable genotyping assay able to detect the mutant allele.

## 2.6 References:


1. Fitterer, B.B., et al., *A polymerase chain reaction-based genotyping assay for detecting a novel Sandhoff disease-causing mutation.* Genet Test Mol Biomarkers, 2012. **16**(5): p. 401-5.
2. Mahuran, D.J., *Biochemical consequences of mutations causing the GM2 gangliosidoses.* Biochim Biophys Acta, 1999. **1455**(2-3): p. 105-38.
3. Beutler, E., W. Kuhl, and D. Comings, *Hexosaminidase isozyme in type O Gm2 gangliosidosis (Sandhoff-Jatzkewitz disease).* Am J Hum Genet, 1975. **27**(5): p. 628-38.
4. Ikonne, J.U., M.C. Rattazzi, and R.J. Desnick, *Characterization of Hex S, the major residual beta hexosaminidase activity in type O Gm2 gangliosidosis (Sandhoff-Jatzkewitz disease).* Am J Hum Genet, 1975. **27**(5): p. 639-50.
5. *Human Gene Mutation Database. www.hgmd.org/, accessed December 3, 2010.*
6. Meikle, P.J., et al., *Prevalence of lysosomal storage disorders.* JAMA, 1999. **281**(3): p. 249-54.
7. Drousiotou, A., et al., *Sandhoff disease in Cyprus: population screening by biochemical and DNA analysis indicates a high frequency of carriers in the Maronite community.* Hum Genet, 2000. **107**(1): p. 12-7.
8. Kaya, N., et al., *GM2 gangliosidosis in Saudi Arabia: Multiple mutations and considerations for future carrier screening.* Am J Med Genet A, 2011. **155**(6): p. 1281-4.
9. Lowden, J.A., et al., *Carrier detection in Sandhoff disease.* Am J Hum Genet, 1978. **30**(1): p. 38-45.
10. O'Brien, J.S., et al., *Tay-Sachs Disease.* New England Journal of Medicine, 1970. **283**(1): p. 15-20.
11. Chamoles, N.A., et al., *Tay-Sachs and Sandhoff diseases: enzymatic diagnosis in dried blood spots on filter paper: retrospective diagnoses in newborn-screening cards.* Clin Chim Acta, 2002. **318**(1-2): p. 133-7.
12. Zhang, Z.X., et al., *Impact of premature stop codons on mRNA levels in infantile Sandhoff disease.* Hum Mol Genet, 1994. **3**(1): p. 139-45.
13. Zampieri, S., et al., *Molecular and functional analysis of the HEXB gene in Italian patients affected with Sandhoff disease: identification of six novel alleles.* Neurogenetics, 2009. **10**(1): p. 49-58.
14. Wilson JA, Z.M., Jacobson JW, Kalman L, Krunic N, Matthijs G, Pratt VM, Schoonmaker MM, Tezak Z, *Verification and Validation of Multiplex Nucleic Acid Assays; Approved Guideline*. CLSI document MM17-A. Vol. 28. 2008, Clinical Laboratory Standards Institute: Wayne, PA.
15. *Human Genome Variation Society. www.hgvs.org/mutnomenrecs-DNA.html#del, accessed June 15, 2011.*
16. Ribeiro, D., A.J. Duarte, and O. Amaral, *Rapid and cost-effective method for the detection of the c.533G>A mutation in the HEXA gene.* Genet Test Mol Biomarkers, 2011. **15**(3): p. 123-6.

# Chapter 3: Tandem Mass Spectrometer Assay for the Detection of Sandhoff Disease Using Dried Blood Spots

## 3.1 Abstract

Sandhoff disease is a rare genetic disorder characterised by deficient β-hexosaminidase activity. The traditional methods for measuring β-hexosaminidase activity are labour intensive for laboratory staff and in some cases require patients to travel hundreds of kilometres to the testing facility in order to provide fresh blood samples. Described herein is a novel MS/MS-based assay for detecting Sandhoff disease using residual dried blood spots on newborn screening cards. Testing dried blood spots allows for remote sample collection and transportation through regular mail to the testing facility. Assay design was consistent with other lysosomal storage disease assays in order to facilitate multiplex screening for at least 6 other lysosomal storage diseases. β-hexosaminidase activity was measured with novel synthetic substrates. The MS/MS-based assay was validated and normal ranges were established for Total Hex activity (63.2 – 147.3 umol/h/L of blood), HexB activity (7.5 – 33.3 umol/h/L of blood) and %HexA (72.3 – 91.7%) using freshly collected samples. Precision ranged from 9.3 – 18.0%CV. A cut-off of <10% of the low end of the normal range for Total Hex activity and greater than 92% HexA resulted in a positive predictive value for Sandhoff disease of 100% (n = 4). Using this method Sandhoff disease can be detected retrospectively from residual dried blood spots stored for up to 5 years.

## 3.2 Introduction

Sandhoff disease (Online Mendelian Inheritance in Man no. 268800) is a rare autosomal recessive lysosomal storage disorder caused by diminished activity of the β-hexosaminidase A (HexA; E.C. 3.2.1.52) and β-hexosaminidase B (HexB; E.C. 3.2.1.52) enzymes [1]. HexA is a heterodimer of α-β subunits produced by the *HEXA* and *HEXB* genes respectively, and HexB is a β-β homodimer [2]. A third isoenzyme HexS also exists as an α-α homodimer and may be responsible for minor residual enzyme activity in Sandhoff disease patients [3]. Compromised β-hexosaminidase activity leads to the distinct clinical phenotype of gangliosidoses, the accumulation of GM2 ganglioside in cells of the central nervous system, and progressive neural degradation [1]. The hydrolysis of GM2 by HexA requires the GM2-activator protein, a substrate specific-cofactor. Accumulation of GM2 can be the result of mutations in the *HEXB* gene (Sandhoff disease) affecting the β-subunit, mutations in the *HEXA* gene (Tay-Sachs disease) affecting the α-subunit, or mutations in the GM2A gene (AB variant) affecting the GM2-activator [2, 4]. Tay-Sachs disease patients will have diminished HexA but normal HexB activity. Sandhoff disease patients will lack both HexA and HexB activity since they both possess the common β-subunit. Patients with the AB variant will show normal HexA and HexB activity when assayed with synthetic substrates but lack the ability to hydrolyse GM2. The presentation of symptoms for any of the three disorders of β-hexosaminidase deficiency can manifest at different stages of life producing infantile, juvenile, and adult onset variants of the disorders with varying degrees of residual enzyme activity.

The isoforms of β-hexosaminidase can hydrolyse terminal N-acetyl-D-hexosamine residues from N-acetyl-D-hexosaminides but only HexA can hydrolyse the terminal N-acetyl-galactose (GalNAc) residue from GM2 [5, 6]. The crystal structure of β-hexosaminidase and the mechanism of N-acetyl-D-hexosamine hydrolysis have been described [7, 8]. The mechanism of β-hexosaminidase hydrolysis can be seen in Figure 3.1. The standard fluorometric method for assaying β-hexosaminidase activity takes advantage of the β-hexosaminidase-specific hydrolytic activity using the 4-methylumbelliferyl-2-acetamido-2-deoxy-β-D-glucopyranoside synthetic substrate (also known as 4-MUG or MUG) [5]. Two methods for differentiating HexA and HexB activity in clinical samples have been established. Specificity for only HexA has been achieved with the use of a sulphate derivative of MUG, 4-methylumbelliferyl-N-acetyl-β-D-glucosaminide-6-sulfate (MUGS) [9]. Alternately, the thermostability of HexB at 52°C has been exploited to denature the HexA isoenzyme allowing the measurement of HexB activity alone [10].

With recent advances in lysosomal storage disease therapies, neonatal screening for these disorders has begun [11]. However, the current methodologies for detecting Sandhoff disease lack throughput, are labour intensive, and require fresh blood samples. The assay described herein is intended to address the shortcomings of the current methodologies by providing a semi-high throughput technique to measure β-hexosaminidase activity using tandem mass spectrometry (MS/MS). This assay has the potential to be multiplexed with other lysosomal storage disease assays in order to facilitate cost-effective screening. Using a novel synthetic substrate, a new method for the detection of Sandhoff disease from dried blood spots is presented.

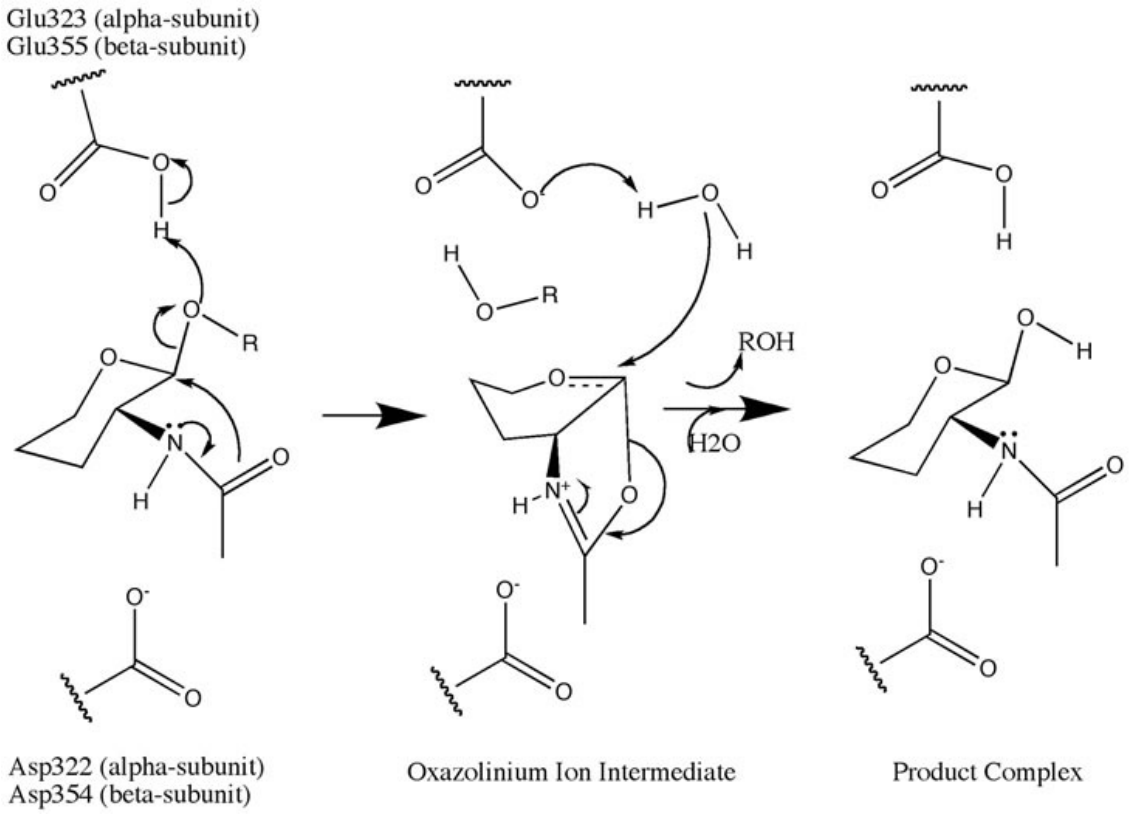



**Figure 3.1**. Mechanism for β-hexosaminidase hydrolysis of terminal N-acetyl-D-hexosamine residues. An enzyme-substrate complex is formed involving a glutamate residue, an aspartate residue, and a terminal N-acetyl-D-hexosamine residue. Glutamate acts as an acid donating a proton to the glycosidic oxygen of the substrate. The aspartate residue orients the C2-acetamido moiety in such a way that it can undergo nucleophilic attack from the oxygen atom on carbon 1. Aspartate stabilizes the intermediate ion while the glutamate deprotonates a water molecule allowing it to perform a second nucleophilic attack on carbon 1 of the ring regenerating the enzyme and releasing the product. Figure taken with permission from [8].

## 3.3 Materials and methods

### **3.31** *Substrates and internal standards*

The β-hexosaminidase specific substrates and internal standards were synthesized in the laboratory of Dr. Michael Gelb at the University of Washington. The structures for the compounds can be seen in Figure 3.2. To prepare the substrates and internal standards for use, all materials were individually dissolved in methanol. Reagent cocktails for measuring Total Hex and HexB activities were made up at 0.23 µM of their respective substrate and internal standard such that each reaction would receive $3.4 \times 10^{-3}$ µmol of substrate and internal standard per 15 µl of cocktail added to each 25 µl reaction. The Total Hex reagent cocktail contained the non-deuterated substrate with internal standard whereas the HexB cocktail contained deuterated substrate with internal standard. Reagent cocktails were dried under nitrogen and stored at -20°C prior to use.

When needed the dried reagents were allowed to equilibrate to room temperature prior to being reconstituted in 8.7 mL of 0.04 M citrate-phosphate buffer (pH 4.4) and 300 µl of sodium taurocholate (120 g / L). Aliquots of convenient volumes were prepared and frozen at -20°C until needed.

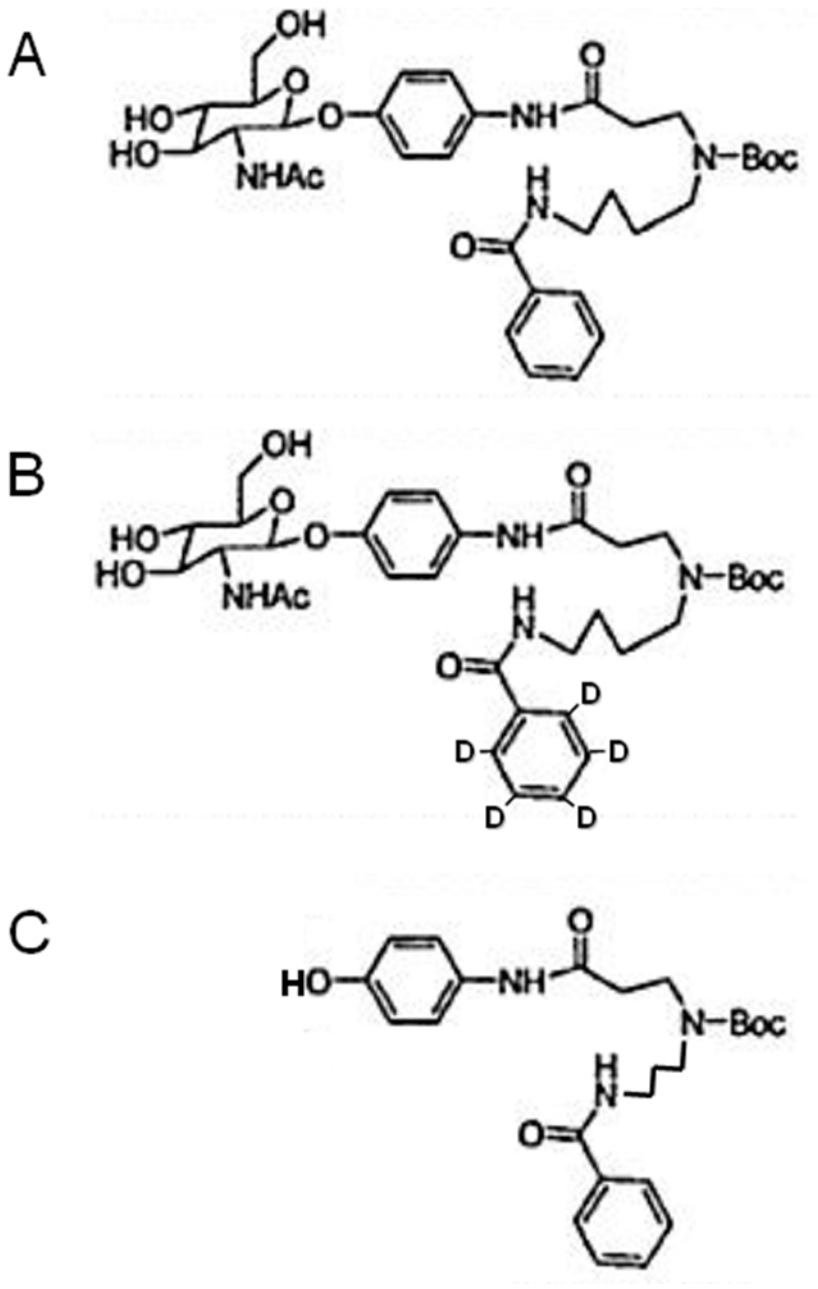



**Figure 3.2.** Synthetic substrates and internal standard used to assay β-hexosaminidase activity.

Compounds were synthesized in the laboratory of Michael Gelb at the University of Washington. (A) Substrate added to the Total Hex reaction. (B) Substrate added to the HexB reaction. (C) Internal standard added to each reaction.

**3.32** *β-Hexosaminidase assay*

The sample preparation procedure was developed to be consistent with other lysosomal storage disease assays in order to facilitate multiplexing in the future. The method for assaying other lysosomal storage disorders has been described by Gelb et al. [11] and is outlined below with relevant modifications for the β-Hexosaminidase assay.

From each newborn screening card a 3 mm punch was taken and deposited into an individual well of a 96-well polypropylene micro plate. Polypropylene materials were required to withstand the organic solvents used downstream. Blood was eluted from the punches with 70 µl of extraction buffer (20 mM sodium-phosphate monobasic pH 7.1). The extraction plate was then incubated at 37°C for 1 hour while shaking at 850 RPM.

While incubating, a fresh plate was prepared by adding 15 µl of Total Hex reagent cocktail to each well (Total Hex plate). The Total Hex reaction was then started by adding 10 µl of each extract to the appropriate wells of the Total Hex plate. The Total Hex plate was then sealed and incubated at 37°C with shaking at 200 RPM overnight (approximately 22 hours). The dried blood spot extract plate was then re-sealed and incubated at 52°C for 1 hour with no shaking in order to heat inactivate β-Hexosaminidase A. Heat inactivation of HexA is common to other β-hexosaminidase assays [12, 13] but was optimized here for these assay conditions.

During the 52°C incubation, 15 µl of the HexB reagent cocktail was added to each well of a new 96-well micro plate (HexB plate). Following the 52°C incubation, 10 µl of heat-inactivated extract was added to the corresponding wells of the HexB plate. The HexB plate was then sealed and incubated overnight at 37°C while shaking at 200 RPM

for approximately 22 hours. Exact incubation start and stop times for each plate were recorded for calculating product generation per hour.

After approximately 22 hours the reactions in each plate were stopped by adding 100 µl of 50:50 (v/v %) methanol / ethylacetate to each well. The 125 µl total volumes from corresponding wells of the Total Hex and HexB plates were then combined in a polypropylene deep-well plate for product purification. The 5 deuteriums on the HexB product allowed for determination of each enzyme's activity while multiplexing the two reactions. If multiplexing with other lysosomal storage disease assays was to take place it would be at this stage that the reactions would be combined.

**3.33** *Product purification*

To the combined quenched reactions, 400 µl of ethylacetate and 400 µl of distilled water were added and mixed to facilitate a liquid-liquid extraction. After mixing the plate was sealed and centrifuged at 14.7 G for 2 minutes. Following centrifugation 300 µl of the organic phase from each well was transferred to a new 96-well polypropylene deep well plate. The material in the organic phase was then dried under nitrogen then reconstituted in 200 µl of 95/5 (v/v %) ethylacetate / methanol for solid phase extraction.

A solid phase extraction apparatus for 96-well plates was prepared by adding approximately 100 mg of silica gel to each well of the filter plate. An empty 96-well polypropylene plate was loaded into the apparatus and the filter plate was washed with 1 ml of 95/5 (v/v %) ethylacetate / methanol under vacuum. The wash plate was discarded and a fresh 96-well polypropylene plate was loaded to collect the product. The 200 µl

volume of reconstituted product was then transferred to the corresponding wells of the filter plate followed by the addition of 4 x 400 µl of 95/5 (v/v %) ethylacetate / methanol under vacuum. The filtered product was then dried down under nitrogen. Prior to detection the material was reconstituted in 100 µl of 80/20 (v/v %) methanol/water with 0.2% formic acid and transferred to a 96-well micro plate. If product detection was not taking place same-day then the plate containing the dried product was sealed and frozen at -20°C until needed.

**3.34** *MS/MS analysis*

Product detection was carried out using an API-2000 mass spectrometer (AB Sciex Concord, Ontario, CA) with Ionics Upgrade (Ionics Bolton, Ontario, CA), coupled to an Agilent HPLC. Analysis was done using flow injection with positive electrospray, and multiple reaction monitoring to identify the target compounds. The mobile phase consisted of 80/20 (v/v %) methanol / water with 0.2% formic acid and a flow rate of 150 µl per minute. Optimized instrument-specific run parameters for β-hexosaminidase product detection are summarized in Table 3.1 and Table 3.2.

**Table 3.1**. Compound specific optimized MS/MS run parameters

| Parameter | Total Hex Product | HexB Product | Internal Standard |
|---|---|---|---|
| Transition | 456.3 > 356.3 | 461.3 > 361.3 | 442 > 342 |
| DP | 41 | 41 | 45 |
| EP | 6 | 6 | 6 |
| CE | 18 | 18 | 15 |
| CXP | 23 | 23 | 24 |
| FP | 275 | 275 | 300 |

*DP: declustering potential; EP: entrance potential; CE: collision energy; CXP: collision cell exit potential; FP: focusing potential.*

**Table 3.2**. Shared MS/MS instrument settings for all experiments

| Parameter | Setting |
|---|---|
| CAD | 6 |
| CUR | 25 |
| GS1 | 25 |
| GS2 | 50 |
| IS | 5500 |
| TEM | 300 |

*CAD: collision gas; CUR curtain gas; GS1 &GS2 ion source gasses; IS ion spray voltage; TEM: temperature.*

Peak areas were calculated using Analyst software (Applied Biosystems). Peak areas for the products and internal standards were exported to Microsoft® Excel in order to calculate β-hexosaminidase enzyme activities. The ratio of peak area for the products to the peak area of the internal standard for each reaction was used to determine the enzyme activity relative to the amount of internal standard added to each reaction, the incubation time, and the volume of eluted blood added to each reaction. HexA enzyme activity and subsequently % HexA were calculated by subtracting the HexB activity from the Total Hex activity.

### **3.35** *β-Hexosaminidase assay validation*

In order to validate the β-hexosaminidase assay for detecting Sandhoff disease the following parameters were examined; 1) matrix blanks were analyzed to detect any interfering substances, 2) the possibility of spontaneous substrate degradation was examined by measuring product formation in blanks, 3) substrate depletion for each reaction (Total Hex and HexB) was evaluated by stopping the assay at time intervals from 2 – 47 hours to ensure that the reaction was not going to completion, 4) precision was assessed for both inter-run and intra-run variation.

In order to distinguish individuals affected with Sandhoff disease from those unaffected it was necessary to establish a normal range for Total Hex, HexB and % HexA activities. Normal ranges were calculated using EP Evaluator® (Data Innovations, South Burlington, Vermont, USA). To assess accuracy, dried blood spots from four diagnosed Sandhoff disease patients were analyzed in order to determine the positive predictive and negative predictive values.

Results from this assay for the four Sandhoff disease patients were compared to results obtained from the Metabolic Disease Laboratory in Saskatoon using the standard fluorometric method for assaying β-hexosaminidase activity. Furthermore, the assay was used to analyze a blind panel of β-hexosaminidase deficient samples sent from an external laboratory.

### **3.36** *β-Hexosaminidase stability on Saskatchewan newborn screening cards*

In Saskatchewan newborn screening cards are stored at 4°C for up to one year after collection and then at room temperature thereafter for a total of 21 years. In order to ascertain the ability of this method to detect β-hexosaminidase activity from Saskatchewan newborn screening cards retrospectively, the assay was used to analyze cards collected from 2005 – 2009.

### **3.37** *Ethics approval*

Ethics approval for the use of residual dried blood spots for the development and validation of this assay was obtained from the Research Ethics Board at the University of Regina and the Biomedical Research Ethics Board at the University of Saskatchewan. To ensure the anonymity of patients and their families all samples were de-identified prior to testing.

## 3.4 Results

### **3.41** *Assay optimization for Sandhoff disease*

The optimal reaction buffer to assay β-hexosaminidase activity in dried blood spots and white blood cells has been previously described [14] and was used here. Incubation time for the heat inactivation of HexA was assessed at 15 minute intervals from 45 - 90 minutes. A 60 min incubation at 52°C was sufficient for complete inactivation of HexA. Individual sample run time was set to 1 min with peak elution occurring around 0.2 min (Figure 3.3). Peaks were well shaped with no tailing and heights several orders of magnitude above any background signal.

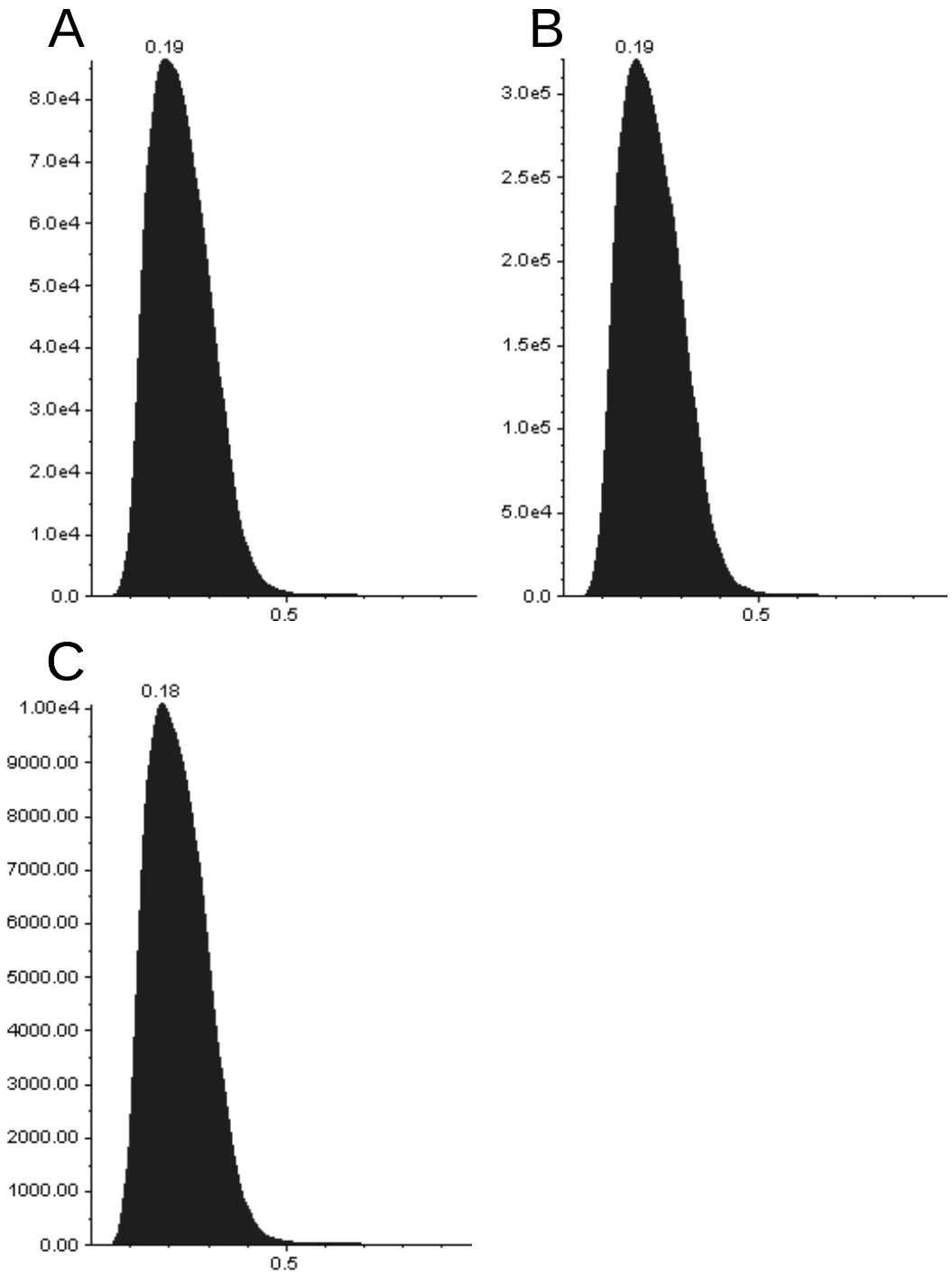



**Figure 3.3.** Optimized MS/MS product and internal standard detection.

The Y-axis is the analyte signal intensity. The X-axis is time (minutes). Panel A is the 456 > 356 transition for the Total Hex Product. Panel B is the 442 > 342 transition for the internal standard. Panel C is the 461 > 361 transition for the HexB product.

**3.42** *β-Hexosaminidase assay validation*

Matrix blanks (dried blood spot punches which received no substrates or internal standard) produced no product or internal standard signal precluding the presence of any interfering substances from the matrix. Blanks that included newborn screening card punches free of blood and blanks which had extraction buffer but no punch and received the substrate/internal standard cocktails were found to produce signal albeit several orders of magnitude smaller than the signal received from normal patient samples. This signal was assumed to be from spontaneous substrate degradation. As such blanks were included on every run and the minor spontaneous product generation was subtracted from test samples.

The amount of product generated by Total Hex or HexB over a 47 hour period is presented in Figure 3.4. Hydrolysis of the synthetic substrate by β-hexosaminidase A and β-hexosaminidase B continues after the 22 hour incubation time used by the assay as shown at the 47 hour mark. Continued product formation after the 22 hour incubation time indicates that neither the Total Hex nor HexB reactions are going to completion. Therefore, substrate depletion is not a factor for the assay run with the outlined conditions.

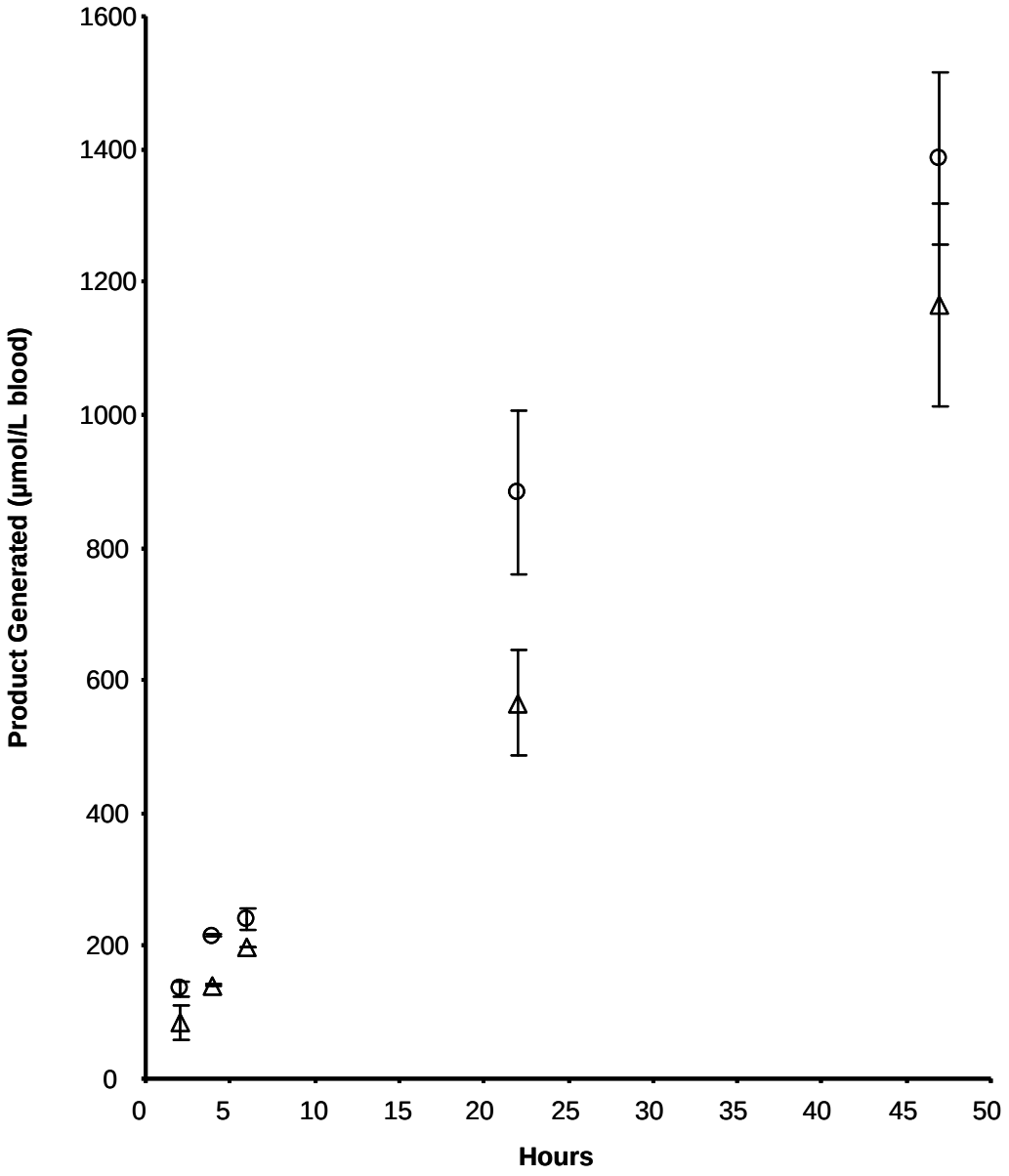



**Figure 3.4**. Product generated by Total Hex or HexB over 47 hours.

The enzymatic reaction continues past the assay incubation time of 22 hours precluding the possibility of substrate depletion. Circles represent Total Hex activity. Triangles represent HexB activity.

Precision of the assay was determined by analyzing replicate dried blood spots. Intra-day precision as a result of five runs resulted in %CV values of 13.7% for Total Hex and 9.3% for HexB (n = 6). Inter-day precision was evaluated over two days and resulted in %CV of 10.0% for Total Hex and 18.0% for HexB (n = 6).

To establish a normal range for β-hexosaminidase levels in newborns using this assay, 400 dried blood spot samples from Saskatchewan newborns were analyzed. The normal ranges for the measured analytes are summarized in Table 3.3. The range for % HexA was broad though typically 3 – 10 times greater than % HexB.

β-hexosaminidase activities from the four Sandhoff disease affected patients were significantly lower than activities from the unaffected population and in some cases undetectable (Table 3.3). For the purpose of screening, a cut-off was established such that a screen positive result for Sandhoff disease would have a Total Hex enzyme activity of less than 10% of the low end of the normal range (< 6.3 µmol/h/L blood), in addition to a higher than normal % HexA (> 92%). With these cut-offs the positive predictive value and negative predictive values are both 100% for detecting Sandhoff disease.

**Table 3.3**. β-hexosaminidase normal ranges and Sandhoff disease patient values

| Parameter | Normal Range | Affected Patient Average* |
|---|---|---|
| Total Hex | 63.2 - 147.3 | 1.23 |
| HexB | 7.5 - 33.3 | 0.01 |
| % HexA | 72.3 - 91.7 | 99.75 |

*β-Hexosaminidase activities are presented as µmol/h/L of blood.*
**Average values from 4 diagnosed Sandhoff disease patients*

In Saskatchewan, suspected cases of Sandhoff disease are diagnosed using the standard fluorometric enzyme assay run at the Metabolic Disease Laboratory in Saskatoon [13]. The normal ranges and average values using the fluorometric assay to determine plasma β-hexosaminidase activity are shown in Table 3.4. The normal ranges from the fluorometric method are established in an age-group dependent manner while normal range for the MS/MS assay described here were determined using dried blood spots from newborns. Comparison of the results for the Sandhoff patients from the MS/MS assay and the fluorometric method used in Saskatoon correlated for all four patients. When a panel of blinded samples from an outside laboratory containing dried blood spot punches from Sandhoff disease and Tay-Sachs disease patients was analyzed, the assay successfully identified deficient β-hexosaminidase activity in the Sandhoff patients but was unable to distinguish any abnormal activity in the Tay-Sachs patients. Furthermore, differentiation of Sandhoff disease or Tay-Sachs disease carriers from normal patients was not possible.

**Table 3.4**. β-hexosaminidase activities using the fluorometric method on plasma

| Parameter | Normal Range (0-1 yr) | Normal Range (1-3 yr) | Affected Patient Average |
|---|---|---|---|
| Total Hex | 390 - 2622 | 532 - 1748 | 14.5 |
| HexB | 50 - 1458 | 267 - 995 | 0 |
| % HexA | 13 - 56 | 24.1 - 52.2 | 100 |

*β-Hexosaminidase activities are presented as µmol/h/L of blood.*

**3.43** *β-Hexosaminidase stability on Saskatchewan newborn screening cards*

Normal ranges for Total Hex and HexB activity were determined by year (Table 3.5). As was expected Total Hex and HexB activities decreased in a time dependent manner but were still detectable in samples that were 5 years old at the time of analysis. When samples older than 5 years were analyzed the lower end of the normal range decreased significantly while the high normal values remained approximately the same. In addition, the % HexA ranges remained fairly constant suggesting that HexA and HexB degrade at similar rates.

**Table 3.5**. β-hexosaminidase activity from stored newborn screening cards

| Year | Total Hex Normal Range | HexB Normal Range | % A Normal Range |
|---|---|---|---|
| 2005 (n = 62) | 30.5 - 108.1 | 1.2 - 10.7 | 87.2 - 96.6 |
| 2006 (n = 124) | 23.1 - 110.9 | 1.9 - 12.2 | 84.5 - 93.7 |
| 2007 (n = 130) | 32.8 - 115.4 | 2.6 - 23.4 | 78.3 - 94.7 |
| 2008 (n = 153) | 38.6 - 112.2 | 2.9 - 18.1 | 77.4 - 93.5 |
| 2009 (n = 62) | 52.8 - 111.5 | 3.2 - 15.0 | 83.9 - 94.3 |

*β-Hexosaminidase activities are presented as µmol/h/L of blood.*

### 3.5 Discussion

The assay presented here has been validated for the purpose of detecting Sandhoff disease from freshly collected dried blood spots or those that have been stored for up to 5 years at room temperature. Similar to the standard assay for measuring β-hexosaminidase activity in plasma at the Metabolic Disease Laboratory in Saskatoon, a one hour heat inactivation step was ideal for abolishing HexA activity.

No interfering substances were detected in the matrix although some spontaneous substrate degradation was detected in our blanks (data not shown). The amount of product detected in the blanks was in the range of 2 to 3 orders of magnitude smaller than that obtained for patient samples. As such the product generated from spontaneous substrate degradation would have little impact on our results though the blanks were still subtracted from patient samples on each run.

Because only a small amount of substrate is added to each reaction it would be detrimental to the assay if either reaction was going to completion during the incubation period. To show that this was not the case, product generation was measured at several time points over a 47 hour period. The results show that product formation continues after 22 hours indicating that substrate depletion does not occur under the assay conditions.

The %CV values obtained for assay precision were around the generally acceptable cut-off of 15%. Dried blood spots are known to be more variable than other sample types such as fresh blood or plasma due to the uneven soaking of blood across the newborn screening card. The variability in the amount of blood on each punch used in our precision studies may account for the %CV values obtained.

Furthermore, there was clear distinction between the normal ranges for Total Hex and %HexA and the values obtained for the four Sandhoff disease patients tested (Table 3.3). β-hexosaminidase activity was undetectable in three cases and only a very small level of residual activity remained in the fourth. The four patients whose material was used for this stage of the assay validation were all affected with the infantile variant of the disease and as such presented with the manifestation of symptoms during the first two years of life. Patients with the infantile variant typically have no residual β-hexosaminidase activity however; juvenile or adult onset Sandhoff patients can have residual activity. It would be worthwhile to obtain dried blood spot material from juvenile and adult onset patients to see if the cut-off values are able to distinguish those patients from the unaffected population. It is possible that the cut-offs would need to be adjusted to identify juvenile and adult onset patients however, other β-hexosaminidase assays are able to make the distinction [15]. Fortunately, Sandhoff disease is extremely rare, thought due to its rarity sample material from juvenile and adult onset patients has not yet been obtained.

The normal range data shows a much higher % HexA than % HexB. Previous studies have indicated % HexA in plasma to be between 44% and 70% [16]. During the validation of this assay % HexA was observed at levels greater than 90%. It would also be important to conduct substrate specificity studies between these synthetic substrates and the HexA and HexB enzymes separately. By conducting a substrate specificity study we may be able to determine if this discrepancy is due to substrate specificity for each enzyme or an assay parameter favouring HexA.

To further validate the MS/MS β-hexosaminidase assay it was compared to the standard fluorometric assay run at the Metabolic Disease Laboratory in Saskatoon [12, 13]. Results from all four Sandhoff disease patients correlated (Table 3.3 and Table 3.4). Additionally, a panel of blinded samples from an external laboratory containing Sandhoff disease patients, Tay-Sachs disease patients, as well as carriers for both diseases was tested. The MS/MS assay properly identified all Sandhoff disease patients in the panel however, carriers for both diseases were indistinguishable from the normal range similar to other β-hexosaminidase assays [17]. The inability to identify carriers may be explained by up regulation of the healthy HexA or HexB allele by the cell to compensate for the pathogenic allele. What is noteworthy is the inability of the assay to detect patients affected by Tay-Sachs disease. Theoretically these patients should have no HexA activity but retain HexB activity. As such a % HexA of zero would be expected along with Total Hex activity equal to that of HexB for Tay-Sachs patients. Further troubleshooting and optimization of the assay will be required if the assay is to be used for the identification of individuals with Tay-Sachs disease.

Moving forward, optimization of the extraction buffer to pH 4.4 rather than pH 7.1 may help to stabilize the HexA and HexB enzymes during elution of the blood and heat inactivation. HexA and HexB are both lysosomal enzymes and exist at a native pH of 4.4. Although HexB has been shown to be stable at 52°C in a pH 4.4 buffer it has not been shown to be stable at 52°C in a pH 7.1 buffer. The compounded effect on the enzyme from both pH and temperature deviations from optimum will likely be detrimental to HexB activity. This point is discussed further in chapter 5. Adjusting any assay parameters such as the extraction buffer pH will affect the normal ranges

highlighting the need for additional validation if Tay-Sachs identification is to be included, though as was shown during validation, the assay works for detecting Sandhoff disease.

The effect of long term storage on the β-hexosaminidase activity on newborn screening cards was investigated. The maximum length of time that β-hexosaminidase activity could be reliably detected on cards stored for one year at 4°C then at room temperature thereafter was 5 years. Consequently, retrospective screening of populations with a high incidence of Sandhoff disease should be limited to 5 years using this MS/MS assay.

The traditional method of detecting Sandhoff disease requires patients to provide a fresh blood sample at a testing facility. In Saskatchewan this can be problematic since communities where the incidence of Sandhoff disease is high [18] can be up to 600 km away from the testing facility. The ability to measure β-hexosaminidase on dried blood spots that can be collected in remote centers and transported through regular mail to the testing facility is valuable to the investigation of suspected cases of Sandhoff disease in these remote communities. Furthermore, the ability to detect Sandhoff disease from dried blood spots allows for presymptomatic newborn screening for this disease, an important consideration given the progressive and irreversible nature of the disorder. With the potential to enrol patients in clinical trials for Sandhoff disease treatments, and as new treatment options become available, the ability to screen newborns for Sandhoff disease will become increasingly valuable.

The MS/MS assay described here provides additional advantages over the classic fluorometric method for measuring β-hexosaminidase levels. The MS/MS-based assay

has been validated in a 96-well high-throughput format in order to facilitate large-scale retrospective or prospective newborn screening. In addition, the assay has been designed to allow for multiplexing with assays that detect six other lysosomal storage diseases [11]. The ability to test for multiple lysosomal storage diseases and only perform a single liquid/liquid extraction, solid-phase extraction, and MS/MS injection for each sample is of great value to newborn screening laboratories operating with finite personnel and financial resources.

In conclusion this MS/MS assay for measuring β-hexosaminidase activity has been validated for identifying Sandhoff disease using residual dried blood on newborn screening cards and has been designed with the potential to be easily multiplexed with other lysosomal storage disease assays in order to facilitate cost effective high-throughput screening.

### 3.6 References


1. Mahuran, D.J., *Biochemical consequences of mutations causing the GM2 gangliosidoses.* Biochim Biophys Acta, 1999. **1455**(2-3): p. 105-38.
2. Beutler, E., W. Kuhl, and D. Comings, *Hexosaminidase isozyme in type O Gm2 gangliosidosis (Sandhoff-Jatzkewitz disease).* Am J Hum Genet, 1975. **27**(5): p. 628-38.
3. Ikonne, J.U., M.C. Rattazzi, and R.J. Desnick, *Characterization of Hex S, the major residual beta hexosaminidase activity in type O Gm2 gangliosidosis (Sandhoff-Jatzkewitz disease).* Am J Hum Genet, 1975. **27**(5): p. 639-50.
4. Heng, H.H., et al., *Refined mapping of the GM2 activator protein (GM2A) locus to 5q31.3-q33.1, distal to the spinal muscular atrophy locus.* Genomics, 1993. **18**(2): p. 429-31.
5. Hou, Y., R. Tse, and D.J. Mahuran, *Direct determination of the substrate specificity of the alpha-active site in heterodimeric beta-hexosaminidase A.* Biochemistry, 1996. **35**(13): p. 3963-9.
6. Kytzia, H.J. and K. Sandhoff, *Evidence for two different active sites on human beta-hexosaminidase A. Interaction of GM2 activator protein with beta-hexosaminidase A.* J Biol Chem, 1985. **260**(12): p. 7568-72.
7. Mark, B.L., et al., *Crystal structure of human beta-hexosaminidase B: understanding the molecular basis of Sandhoff and Tay-Sachs disease.* J Mol Biol, 2003. **327**(5): p. 1093-109.
8. Lemieux, M.J., et al., *Crystallographic structure of human beta-hexosaminidase A: interpretation of Tay-Sachs mutations and loss of GM2 ganglioside hydrolysis.* J Mol Biol, 2006. **359**(4): p. 913-29.
9. Bayleran, J., P. Hechtman, and W. Saray, *Synthesis of 4-methylumbelliferyl-beta-D-N-acetylglucosamine-6-sulfate and its use in classification of GM2 gangliosidosis genotypes.* Clin Chim Acta, 1984. **143**(2): p. 73-89.
10. Kaback, M.M., T.J. Nathan, and S. Greenwald, *Tay-Sachs disease: heterozygote screening and prenatal diagnosis--U.S. experience and world perspective.* Prog Clin Biol Res, 1977. **18**: p. 13-36.
11. Gelb, M.H., et al., *Direct multiplex assay of enzymes in dried blood spots by tandem mass spectrometry for the newborn screening of lysosomal storage disorders.* J Inherit Metab Dis, 2006. **29**(2-3): p. 397-404.
12. O'Brien, J.S., et al., *Tay-Sachs Disease.* New England Journal of Medicine, 1970. **283**(1): p. 15-20.
13. Okada, S. and J.S. O'Brien, *Tay-Sachs disease: generalized absence of a beta-D-N-acetylhexosaminidase component.* Science, 1969. **165**(3894): p. 698-700.
14. Chamoles, N.A., et al., *Hurler-like phenotype: enzymatic diagnosis in dried blood spots on filter paper.* Clin Chem, 2001. **47**(12): p. 2098-102.
15. Mitsumoto, H., et al., *Motor neuron disease and adult hexosaminidase A deficiency in two families: evidence for multisystem degeneration.* Ann Neurol, 1985. **17**(4): p. 378-85.
16. Casal, J.A., E. Cano, and J.C. Tutor, *Beta-hexosaminidase isoenzyme profiles in serum, plasma, platelets and mononuclear, polymorphonuclear and unfractionated total leukocytes.* Clin Biochem, 2005. **38**(10): p. 938-42.

17. Chamoles, N.A., et al., *Tay-Sachs and Sandhoff diseases: enzymatic diagnosis in dried blood spots on filter paper: retrospective diagnoses in newborn-screening cards.* Clin Chim Acta, 2002. **318**(1-2): p. 133-7.
18. Lowden, J.A., et al., *Carrier detection in Sandhoff disease.* Am J Hum Genet, 1978. **30**(1): p. 38-45.

# Chapter 4: Investigating the Incidence and Carrier Frequency of Sandhoff Disease in Northern Saskatchewan

## 4.1 Abstract

Sandhoff disease is a rare progressive neurodegenerative genetic disorder with a high incidence among certain isolated communities and ethnic groups around the world. Previous reports have shown a high occurance of Sandhoff disease in northern Saskatchewan. Newborn screening cards from northern Saskatchewan were retrospectively screened in order to investigate the incidence and determine the carrier frequency of Sandhoff disease in these communities. PCR-based screening was conducted for the c.115delG variant that was previously found in 4 Sandhoff disease patients from this area. The carrier frequency for this allele was estimated to be ~1:27. MS/MS-based screening along with genetic sequencing allowed for the identification of additional variants based on low Total Hex activity and high % HexA relative to c.115delG carriers. In total 4 pathogenic variants were discovered in the population (c.115delG, c.619A>G, c.1601G>T, and c.1652G>A) of which two are previously unreported (c.1601G>T and c.1652G>A). The combined carrier frequency of these alleles in the study area was estimated at ~ 1:15. Based on the number of cases of Sandhoff disease from this area we estimate the incidence to be ~1:390 corresponding to a child being born with the disease every 1-2 years on average.

## 4.2 Introduction

Sandhoff disease (Online Mendelian Inheritance in Man no. 268800) is a rare autosomal recessive lysosomal storage disorder caused by mutations in the hexosaminidase-B (*HEXB*) gene (5q13). Mutations in *HEXB* result in a deficiency of the β-hexosaminidase A (HexA; E.C. 3.2.1.52) and β-hexosaminidase B (HexB; E.C. 3.2.1.52) β subunit [1]. HexA is a heterodimer of α-β subunits, and HexB is a β-β homodimer [2]. A third form β-hexosaminidase-S (HexS; E.C. 3.2.1.52) may also be present as a homodimer of α-α subunits [3]. Normally HexA is responsible for the degradation of GM2 ganglioside however; when Sandhoff disease mutations are present the diminished HexA activity leads to the progressive accumulation of GM2 in neuronal cells and irreversible neuronal degradation [1]. The symptoms of Sandhoff disease can manifest at different stages of life corresponding to the amount of residual enzyme activity and the mutations present in the *HEXB* gene [1]. Symptoms manifest before one year and lead to death typically by four years of age for the infantile onset form of the disorder. Juvenile and adult onset forms of the disease are also possible.

Estimates of Sandhoff disease causing mutations in the general population vary slightly at 1:310 based on the prevalence of Sandhoff disease in Australia between 1980-1996 [4] and 1:276 (n = 32,342) in non-Jewish Americans based on serum β-hexosaminidase levels [5]. There is also suggestion that non-Jewish German descendants have a lower frequency of Sandhoff disease causing mutations while those of Mexican or Central American descent have a higher frequency for such mutations [5]. Other reports have been made for isolated communities or those with a high degree of consanguinity having high Sandhoff disease carrier frequencies. The IVS-2+1 G>A splice site mutation

has been implicated as the predominant allele responsible for Sandhoff disease in Argentina [6]. In Saudi Arabia the high degree of consanguinity has led to a very high incidence for many autosomal recessive conditions including Sandhoff disease where several different mutations have been described though patients are typically homozygous for a family-specific allele [7]. Cyprus has the highest reported Sandhoff disease carrier frequency among its Christian Maronite community at 1:7 (n = 244) [8]. Sandhoff disease has also been reported among French Canadians and those of French descent due primarily to a 16 kb deletion [9]. The high prevalence of the 16 kb deletion among French Canadian populations is suspected to be the result of a founder effect and subsequent genetic drift.

Some northern Saskatchewan communities also have a high incidence of Sandhoff disease [10] where all cases have been the infantile variant. Previously we found the c.115delG mutation in the *HEXB* gene from several Sandhoff disease patients born in this area [11]. At this time we have been unable to determine a direct relationship between Sandhoff disease in northern Saskatchewan where the community is largely Métis (individuals of mixed French/aboriginal descent) and the reports of Sandhoff disease among other French Canadian populations.

In this study, variants in the *HEXB* gene and aberrant β-hexosaminidase levels from newborn screening cards collected from individuals born in northern Saskatchewan were retrospectively investigated. Our objectives were to determine the frequency of the c.115delG variant previously found in several affected patients, investigate the possibility of other *HEXB* variants in the population, characterise those variants via in silico analysis, estimate the frequency of all Sandhoff disease causing variants in the

population, estimate the incidence of Sandhoff disease, and compare the frequency of Sandhoff disease causing mutations in our study population to the frequency in the general population. Several estimates of the frequency of Sandhoff disease causing alleles in the general population have previously been described [4, 5] however, the studies have used restricted population sampling. As such a global sample of *HEXB* genes was considered by analyzing the data provided by the 1000 Genomes project in order to estimate the frequency of Sandhoff disease causing alleles in the global population.

## 4.3 Materials and methods

### **4.31** *Study area and selection of newborns to screen*

A high incidence of Sandhoff disease among northern Saskatchewan communities has been known since the late 1970's [10]. In order to establish the incidence and carrier frequency of Sandhoff disease in these northern Saskatchewan communities, we selected our study area based on the first three digits of the postal code listed for each newborn's mother. Consequently, children born to mothers whose first three digits of the postal code matched the first three digits from mothers who previously had children affected by Sandhoff disease were chosen to be included in our study. The long-term retention of newborn screening cards in Saskatchewan began in 2000. Thus, our retrospective analysis was limited to newborns born between 2000 and 2012. A total of 1561 individuals were included in the study.

**4.32** *c.115delG genotyping assay*

The c.115delG genotyping assay described in chapter two was used to screen for the common mutation found among Sandhoff disease patients from northern Saskatchewan. The assay for detecting the c.115delG mutation has been previously described (Fitterer et al. [11]) and was used to determine each individuals genotype for this particular variant. To summarize, this assay uses real-time PCR to detect the presence of the c.115delG allele in blood eluted from two 3 mm dried blood spot punches. The PCR assay was used to screen 1561 individuals from the study area born between 2000 and 2012. All specimens found to contain the c.115delG allele were confirmed by repeating the assay.

**4.33** *β-hexosaminidase enzyme assay*

β-hexosaminidase enzyme activity in residual dried blood spots on newborn screening cards was measured according to the protocol described in chapter 3. Briefly, blood from a single 3 mm punch from each newborn screening card was eluted and separate aliquots were incubated with two different synthetic substrates. Denaturation of HexA at 52°C in one aliquot allowed for determination of HexB alone while the other aliquot measured total β-Hexosaminidase (Total Hex) activity. Product detection was carried out by tandem mass spectrometry (MS/MS). The ability of the β-hexosaminidase assay to detect affected Sandhoff disease individuals and carriers was previously established in chapter 3. Validation of the assay used freshly collected dried blood samples similar to conditions that would be expected for the routine testing of newborns in a newborn screening facility.

However, due to the long-term storage that had occurred for many of the samples included in our retrospective analysis, special consideration was taken for identifying aberrant β-hexosaminidase activity from older specimens. Total Hex activity remaining on the residual dried blood spots was found to diminish in a time dependent manner. As such, normal ranges for Total Hex activity and % HexA for the purpose of identifying affected individuals were established for each year of our retrospective study (Chapter 3 Table 3.5). The MS/MS assay was used to measure β-hexosaminidase activity from 760 newborn screening cards collected between 2005 and 2009.

The work done in chapter 3 indicated that the MS/MS-based assay was not capable of distinguishing Sandhoff disease carriers from the normal population. It is possible for individuals who carry Sandhoff disease causing mutations to exhibit normal levels of β-hexosaminidase activity. However, they may also have β-hexosaminidase activity in the lower end of the normal range [5, 12]. One of our objectives was to identify Sandhoff disease causing mutations in the population. Ideally the *HEXB* gene from all individuals born in the study area would have been sequenced but due to the time consuming and expensive restrictions of genetic sequencing we chose to narrow our search for potential Sandhoff disease causing mutations in the population using the MS/MS assay.

In order to identify Sandhoff disease causing mutations in the population we analyzed the MS/MS data and established a potential carrier cut-off for individuals who fell within limits of Total Hex activity and % HexA for their respective birth year. Samples that met these criteria were repeated by the MS/MS assay. The *HEXB* gene was

then sequenced from samples that repeatedly fell within the cut-off range for potential carriers.

**4.34** *Sequencing the HEXB gene*

Genetic Sequencing of the *HEXB* gene was carried out as previously described [11]. Exons one through fourteen of the *HEXB* gene and several nucleotides from the flanking intronic regions were amplified by PCR. The genetic sequence was determined using dye-terminator sequencing. Data analysis was conducted using BioNumerics v6.5 (Applied Maths). Variants were determined relative to the *HEXB* reference sequence NT_006713.15 from Genbank. An alignment was created for each exon and sequence differences were highlighted. Genetic sequence variations and protein level variations were named according to the Human Genome Variation Society guidelines [13].

**4.35** *Analysis of genetic variants*

The deleterious or benign nature of non-synonymous single nucleotide polymorphisms (nsSNPs) was investigated using the online prediction tools PolyPhen-2 [14] and PROVEAN [15] in addition to further in silico analysis. To briefly summarize the PolyPhen-2 prediction tool, a BLAST search and multiple sequence alignment is conducted using the reference sequence input by the user. The nsSNP is entered then evaluated for its effect on protein function by considering the impact on sequenced based features such as annotated locations in the sequences, position-specific independent count scores for two amino acid variants appearing at a single loci in the BLAST search, structural features such as disruption of hydrophobic cores or protein-protein interactions,

secondary structure, changes in protein surface area, as well as additional structural and calculated parameters. In addition PolyPhen-2 provides the variant mapped onto a 3D model of the protein if the structure is known. The positive predictive value for Polyphen-2 is between 73% and 92% with a false positive rate of about 20% [14].

PROVEAN is the next generation of the SIFT tool (used by the 1000 Genomes Project to evaluate nsSNPs). PROVEAN collects a set of homologous and distantly related sequences from the NCBI protein database and arranges them into clusters. Each sequence is assigned a delta score using the BLOSUM62 substitution matrix and an average delta score is calculated for each cluster. The average delta score for all clusters is then calculated from all average cluster scores to yield the PROVEAN score. A variant is predicted to be deleterious if the PROVEAN score is equal to or less than the threshold which is set at -2.5 using default settings. The accuracy for predicting the deleterious outcome of SNPs using PROVEAN is estimated at about 79% [15]. The protein reference sequence used for both PolyPhen-2 and PROVEAN predictions was P07686.

Appendix A provides a quick reference to amino acid codes and sctuctures to assist with understanding the nomenclature heavy protein level variation discussion.

### **4.36** *Comparison to the 1000 Genomes dataset*

In addition to screening the population in our study area, the publicly available genome data provided by the 1000 Genomes project was also analyzed [16]. At the time of this analysis the 1000 Genomes pilot data included the human genomes from 1092 individuals. These 1092 individuals come from the populations listed in Table 4.1. The 1000 Genomes dataset was compared to our study data to see if any of the variants found

among the Saskatchewan cohort were also present in that dataset. Clinically relevant *HEXB* variants present among the 1092 individuals were analyzed and discussed. PolyPhen-2 and SIFT scores for the variants present in the 1000 Genomes dataset were considered along with in silico analysis and discussion of reports in the literature for specific variants in order to determine the pathogenicity of each variant. The frequency of Sandhoff disease causing alleles in the global sample was then estimated for comparison to the northern Saskatchewan population.

The 1000 Genomes database not only includes the variants found in the 1092 individuals whose genomes were sequenced but additional variants from multiple other databases including dbSNP, the Human Gene Mutation Database, and COSMIC among others. In order to filter only those variants from the 1092 individuals the data was accessed in the following manner. In the 1000 Genomes browser the search term "HEXB" was used. The "variations in gene" page was then selected. From the variations in gene page the appropriate transcript was selected. In the case of the *HEXB* gene this was the HEXB-001 CCDS (consensus coding sequence) transcript. The variation table for this transcript was selected from the left hand menu. Under the "Configure this page" option several criteria were changed. All variation sources were selected. All variation classes were selected. The consequence types selected were essential splice site, splice donor, splice acceptor, stop gained, frameshift, stop lost, initiator codon, inframe insertion, inframe deletion, missense, splice region, incomplete terminal codon, stop retained, synonymous, and coding sequence. The variant table was then downloaded for all of the variants included in the above parameters. The downloaded table was opened in Microsoft Excel® and any variants that did not possess

a minor allele frequency (MAF) greater than zero were excluded. The MAF field was only populated (MAF > 0) for variants found in the 1092 individuals included in the 1000 Genomes project.

| **Table 4.1** *Populations and superpopulations included in the 1000 Genomes project* | | |
|---|---|---|
| Super Population Code | Population Code | Description |
| ASN | CHB | Han Chinese in Beijing, China |
| ASN | JPT | Japanese in Tokyo, Japan |
| ASN | CHS | Southern Han Chinese |
| ASN | CDX | Chinese Dai in Xishuanagbanna, China |
| ASN | KHV | Kinh in Ho Chi Minh City, Vietnam |
| EUR | CEU | Utah Residents (CEPH) with Northern and European ancestry |
| EUR | TSI | Toscani in Italia |
| EUR | FIN | Finnish in Finland |
| EUR | GBR | British in England and Scotland |
| EUR | IBS | Iberian population in Spain |
| AFR | YRI | Yoruba in Ibadan, Nigeria |
| AFR | LWK | Luhya in Webuye, Kenya |
| AFR | GWD | Gambian in Western Divisions in the Gambia |
| AFR | MSL | Mende in Sierra Leone |
| AFR | ESN | Esan in Nigera |
| AFR | ASW | Americans of African Ancestry in SW USA |
| AFR | ACB | African Carribeans in Barbados |
| AMR | MXL | Mexican Ancestry from Los Angeles USA |
| AMR | PUR | Puerto Ricans from Puerto Rico |
| AMR | CLM | Colombians from Medellin, Colombia |
| AMR | PEL | Peruvians from Lima, Peru |
| SAN | GIH | Gujarati Indian from Houston, Texas |
| SAN | PJL | Punjabi from Lahore, Pakistan |
| SAN | BEB | Bengali from Bangladesh |
| SAN | STU | Sri Lankan Tamil from the UK |
| SAN | ITU | Indian Telugu from the UK |
| | | |
| AFR | | African |
| AMR | | As Mixed American |
| ASN | | East Asian |
| EUR | | European |
| SAN | | South Asian |

4**.37** *Ethics approval*

Special consideration was taken in order to protect the anonymity of the individuals, families, and communities whose residual dried blood spots were analyzed in this study. All patient information and personal identifiers were removed prior to analysis. Ethics approval for the use of residual dried blood spots was obtained from the University of Regina research ethics board and the University of Saskatchewan Biomedical Research Ethics Board.

## 4.4 Results

**4.41** *c.115delG genotyping assay*

The results from the c.115delG mutation screening can be seen in Table 4.2. A total of 1561 newborns born between 2000 and 2012 were screened. We found 57 individuals who had a single copy of the c.115delG allele and 3 individuals who were homozygous for the allele and determined to be Sandhoff disease affected patients. As such, the carrier frequency of the c.115delG allele in the study area is estimated to be approximately 1:27.

**Table 4.2** *HEXB* variants found in the northern Saskatchewan population

| Variant | Exon | # alleles found/# alleles analyzed | Consequence | Previously Described | Polyphen-2 Prediction* | PROVEAN Prediction** |
|---|---|---|---|---|---|---|
| *c.115delG* | 1 | 63/3122 | Reading Frame Shift | Yes [11] | N/A | N/A |
| *c.362A>G* | 2 | 3/48 | *p.Lys121Arg* | No | 0.00 Benign | -0.353 Neutral |
| *c.619A>G* | 5 | 16/48 | *p.Ile207Val* | Yes [17] | 0.281 Benign | -0.964 Neutral |
| *c.1601G>T* | 13 | 1/48 | *p.Cys534Phe* Loss of C534-C551 bond | No *c.1601G>A* described [18] | 1.00 Probably Damaging | -10.029 Deleterious |
| *c.1652G>A* | 14 | 1/48 | *p.Cys551Tyr* Loss of C534-C551 bond | No | 1.00 Probably Damaging | -9.773 Deleterious |

**PolyPhen-2 prediction score can range from 0 (benign) to 1 (probably damaging) with a default cut-off of 0.432.*

***PROVEAN scores are deleterious if less than -2.5 and neutral if greater than -2.5 using default settings.*

**4.42** *β-hexosaminidase enzyme assay*

Of the 760 newborn screening cards screened from the study area, four of the samples fell in the cut-off range for affected individuals established in chapter three. Those four individuals were confirmed to be the Sandhoff disease affected patients by genetic sequencing of the *HEXB* gene. An additional 35 samples were below the thresholds that were established for potential carriers for each year.

The results from the β-hexosaminidase activity screening and the c.115delG mutation screening were compared. It was found that only 21 of 35 individuals in the potential carrier range possessed the c.115delG mutation. A further 17 of the remaining 725 individuals had a single copy of the c.115delG mutation and were missed by the original cut-offs for potential carriers. As such the cut-off for each year was adjusted so that all of the samples known to contain the c.115delG allele would be included and the data was reanalyzed. The yearly cut-off values for Total Hex and % HexA that were used are listed in Table 4.3. Subsequently 76 samples fell below the adjusted cut-offs of which 4 were affected individuals and 34 carried the c.115delG mutation. The remaining 38 samples lacked the c.115delG allele. These 38 were repeated and the *HEXB* gene was sequenced from those samples that repeatedly fell below the adjusted cut-offs.

**Table 4.3** Cut-offs for Sandhoff disease carriers by year based on the Total Hex and % HexA of individuals found to carry the c.115delG mutation.

| Year | Total Hex | % HexA |
|---|---|---|
| 2005 | 37.41 | 92.6 |
| 2006 | 32.25 | 93.1 |
| 2007 | 40.84 | 91.2 |
| 2008 | 49.18 | 89.4 |
| 2009 | 57.89 | 92.2 |

**The units for Total Hex are µmol/hour/L of blood.*

***Cut-offs are Total Hex activity less than the values stated and % HexA greater than the values stated.*

**4.43** *HEXB genetic sequencing*

Sequencing of the *HEXB* gene from individuals with low β-hexosaminidase activity revealed three mutations in addition to the c.115delG variant previously found in the northern Saskatchewan population. These mutations include c.362A>G, c.619A>G, and c.1601G>T. Furthermore, only three of the four individuals with no hexosaminidase activity (Sandhoff disease affected individuals) were homozygous for the c.115delG allele. The fourth patient was a compound heterozygote having one copy of the gene containing the c.115delG mutation and a c.1652G>A polymorphism in the other copy. A complete list of the genetic variants and the frequency of each allele that was found is shown in Table 4.2.

**4.44** *Analysis of genetic variants*

PolyPhen-2 and PROVEAN predictions were in agreement for all 4 nsSNPs found among the population from the study area. The c.362A>G nsSNP causes a change of Lys at position 121 of the β-hexosaminidase β subunit to Arg. Both of these amino acids share long positively charged side chains and thus have similar potential for electrostatic interactions. A PolyPhen-2 score of 0.00 (benign) and a PROVEAN score of -0.353 (neutral) correlate in predicting the c.362A>G variant to be tolerated.

The second nsSNP c.619A>G results in a conversion of Ile at position 207 to Val. The difference due to this conversion is essentially the loss of a methyl group on the side chain of the amino acid while remaining small and non-polar. PolyPhen-2 and PROVEAN scores for the c.619A>G SNP were 0.281 (benign) and -0.964 (neutral) respectively.

Both c.1601G>T and c.1652G>A result in the loss of key Cys residues responsible for the formation of disulfide bonds. The c.1601G>T variant changes Cys at position 534 to Phe while c.1652G>A results in the change of Cys at position 551 to Tyr. PolyPhen-2 and PROVEAN scores for the c.1601G>T variant were 1.00 and -10.029 respectively and for the c.1652G>A variant were 1.00 and -9.773 indicating a probably damaging or deleterious outcome for both of these nsSNPs.

Although the c.115delG allele was not analyzed by either program, its pathogenic nature is clear due to the induced reading frame shift in the coding sequence of exon one. A summary of the variants found, the number of times each allele was detected and the PolyPhen-2 and PROVEAN predictions can be seen in Table 4.2.

**4.45** *Comparison to the 1000 Genomes data set*

A total of 19 variants were present in the coding region of the *HEXB* gene or the adjacent intronic splice regions from the 1092 individuals included in the 1000 Genomes data set. The potential pathogenicity and the frequency of those variants was analyzed. The 19 variants include 5 synonymous DNA-level variations (Table 4.4), one splice region variant (Table 4.4), and 13 missense mutations (Table 4.5). None of the Sandhoff disease causing alleles found in the Saskatchewan study population were present in the 1000 Genomes data set although two other variants matched mutations found in the Saskatchewan population.

The synonymous variants present in the 1000 Genomes dataset include c.276C>T, c.978G>C, c.1035A>C, c.1051T>C, and c.1251G>A. All of the synonymous variants were rare compared to the c.362A>G and c.619A>G mutations found in both

Saskatchewan and the 1000 Genomes data set having minor allele frequencies of less than 0.0074 and only a single occurrence found in the 2184 alleles sampled for c.978G>C, c.1035A>C, and c.1251G>A. The c.722-4A>G splice region variant in the *HEXB* gene from the 1000 Genomes dataset was located -4 nucleotides from the 5-prime end of exon 7. The synonymous variants are benign however the outcome of the splice region variant requires explanation. In order to consider the clinical relevance of genetic variations in splice regions it is important to recognize the areas of the gene that must be conserved in order to facilitate proper splicing. In order to facilitate proper splicing the following sequence regions must be conserved; GURAGU at the 5' end of the intron, NNAG at the 3' end of the intron, a sequence of 7 pyrimidines upstream from the 3' end of the intron, a YNYURAY branch site upstream from the pyrimidine box, a G at the 5' end of the downstream exon and an AG at the 3' end of the upstream exon [19]. The c.722-4A>G variant is located at the 3' end of the intron. Given that the consensus sequence for a 3' intron splice site is NCAG or NNAG in higher eukaryotes such as humans [19, 20] it is likely that even though A is fairly conserved at the location 4 nucleotides upstream from this intron/exon boundary of the *HEXB* gene, the G variant will be tolerated since any of the four nucleotides are observed at this location of the splice region. Therefore all of the variants listed in Table 4.4 are benign and not included when estimating the frequency of Sandhoff disease causing alleles in the global population.

Of the 13 missense variants present in the 1000 Genomes dataset, Polyphen-2 and SIFT predictions agree that 8 are benign polymorphisms (Table 4.5). The 8 missense variants predicted to be benign are c.185C>T, c.251A>G, c.362A>G, c.449C>A,

c.619A>G, c.922C>G, c.1258A>G, and c.1437A>C. The c.185C>T [21], c.362A>G [22], and c.619A>G [17, 23, 24] variants have been described in the literature. Both c.185C>T and c.362A>G have proven to be benign polymorphisms when examined in clinical samples however, c.619A>G has been implicated in adult onset GM1 gangliosidosis [17, 23, 24]. No reports were found for the other 5 missense variants predicted to be benign and based on the positive predictive value and accuracy for the Polyphen-2 and SIFT tools they are assumed to be as such. Therefore of these 8 missense variants only c.619A>G was considered when calculating the frequency of pathogenic variants.

The two prediction tools also agree that 4 of the missense mutations (c.214C>T, c.923C>T, c.1066G>A, and c.1250C>T) are probably damaging or deleterious and disagree on 1 variant, c.1367A>C, for which the Polyphen-2 prediction is possibly damaging whereas the SIFT prediction is benign. Of the 4 missense mutations for which the prediction tools outcome of potentially pathogenic was in agreement, 3 are of unknown significance and 1 has been described in clinical samples previously.

The c.214C>T variant results in the change of Leu at position 72 to Phe. Leu 72 is located at a region of the peptide chain between an alpha helix and beta sheet on the periphery of the peptide with about 42% conservation of Leu at this position. Leu 72 is not a critical residue for either the active site or protein secondary structure however, both prediction tools suggest that a change to Phe at this position will be detrimental to protein function based on the lack of Phe at this position in related sequences. The c.214C>T variant was found a total of 13 times in the American, Asian, and European superpopulations of the 1000 Genomes dataset but was not found in the African group.

Two variants were present at position 308 of the peptide sequence, c.922C>G (p.Pro308Ala) and c.923C>T (p.Pro308Leu). Both Polyphen-2 and SIFT agree that that the former is a benign variant due to the presence of alanine at position 308 in other closely related sequences however, both prediction tools suggest that the change of Pro 308 to Leu will be damaging to the beta subunit. At position 308 of the alignment of related *HEXB* sequences proline is the most common having moderate conservation however, Ala and Ile are also present among other amino acid variants. Since the structural difference between Ile and Leu is essentially the placement of a methyl group on the side chain the pathogenicity of the c.923C>T variant is questionable yet predicted to be damaging due to there being no leucine at position 308 in any of the aligned sequences. The c.923C>T allele was rare showing up only a single time in an individual of Mexican ancestry living in Los Angeles.

The c.1066G>A variant is also of unknown significance. The peptide locus is between an alpha helix and beta strand structures. Both prediction tools suggest that the subsequent p.Val356Met amino acid change will be detrimental to the enzymes function based on the high conservation of Val at position 356 (~80%). The c.1066G>A variant was only found in a single Southern Han Chinese individual.

The c.1250C>T variant has been described by Wakamatsu et. al and Gomez-Lira et. al in compound heterozygous patients suffering from juvenile and adult onset Sandhoff disease respectively [22, 25]. This variant results in the change of Pro at position 417 to Leu. The Pro at position 417 is responsible for a turn on the periphery of the secondary structure of the beta subunit that properly orients the adjacent alpha helix and beta strands. Without Pro at position 417 it would be expected that the polypeptide

would not fold properly. However, using transfection experiments Wakamatsu et. al showed that this variant actually causes the activation of a cryptic splice site which results in the loss of exon 11 from the processed transcript accounting for the loss of β-hexosaminidase activity [22]. Interestingly, this allele was observed once in each of the Finnish, Italian, and Japanese populations included in the 1000 Genomes sample which correlates well with the previous literature reports from Wakamatsu et. al (Japanese patient) and Gomez-Lira et. al (Italian patient) [22, 25].

For the single variant which the two prediction tools disagreed upon the Polyphen-2 score of 0.673 indicates that amino acid variation caused by c.1367A>C will likely be damaging to the protein's function as a result of the Tyr to Ser change at position 456. Tyr 456 is located in a region of the peptide responsible for protein-protein interaction between the subunits that combine to form β-hexosaminidase A and B [26, 27]. In fact Tyr 456, and its neighbours, Asp 452 and Tyr 450 are responsible for hydrogen bonding between the subunits that stabilize the active site of the enzyme [26, 27]. Banerjee et. Al have previously described and characterized the c.1367A>C variant as being non-functional by transfecting the c.1367A>C *HEXB* gene into COS-7 cells where they observed no functional β-hexosaminidase B formed [17, 23]. The authors go on to explain how further in silico analysis of the variant predicts a dramatic change in the folding of the β-subunit. The c.1367A>C variant was observed only a single time in the 1000 Genomes dataset in an Italian individual.

Therefore, as a result of the Polyphen-2 and SIFT predictions along with the mutation analysis presented here and reports from the literature the following alleles were used to determine the frequency of potentially pathogenic alleles in the global population

sample: c.619A>C, c.214C>T, c.923C>T, c.1066G>A, c.1250C>T, and c.1367A>C. Subsequently, the combined frequency for these potentially pathogenic alleles in the 1000 Genomes sample was 0.15934. If we exclude c.619A>C and only include the potentially Sandhoff disease causing alleles the frequency is 0.0087 or a Sandhoff disease carrier rate of roughly 1:57.

**Table 4.4** Synonymous and splice region *HEXB* variants found in the 1000 Genomes dataset

| Variant | Global MAF[1] | Highest frequency population | Consequence |
|---|---|---|---|
| c.276C>T | 0.00137 | YRI | Synonymous |
| c.978G>C | 0.00046 | GBR | Synonymous |
| c.1035A>C | 0.00046 | JPT | Synonymous |
| c.1051T>C | 0.00733 | IBS | Synonymous |
| c.1251G>A | 0.00046 | YRI | Synonymous |
| c.772-4A>G | 0.04991 | IBS | Splice Region Variant |

[1] *Global MAF: Global minor allele frequency*

[2] *SIFT scores are deleterious if less than or equal to 0.05 and tolerated if greater than 0.05.*

[3] *PolyPhen-2 prediction score can range from 0 (benign) to 1 (probably damaging) with a default cut-off of 0.432.*

**Table 4.5** Missense *HEXB* variants found in the 1000 Genomes dataset

| Variant | Global MAF[1] | Highest frequency population | Consequence | SIFT[2] | Polyphen-2[3] |
|---|---|---|---|---|---|
| c.185C>T | 0.02060 | TSI | *p.Ser62Leu* | 1 Tolerated | 0.001 Benign |
| c.251A>G | 0.00046 | CHB | *p.Asn84Ser* | 1 Tolerated | 0 Benign |
| c.362A>G | 0.20742 | LWK | *p.Lys121Arg* | 0.65 Tolerated | 0 Benign |
| c.449C>A | 0.00641 | ASW | *p.Tyr150Asn* | 0.18 Tolerated | 0.021 Benign |
| c.619A>G | 0.15064 | JPT | *p.Ile207Val* | 0.31 Tolerated | 0.157 Benign |
| c.922C>G | 0.00046 | ASW | *p.Pro308Ala* | 0.08 Tolerated | 0.012 Benign |
| c.1258A>G | 0.01236 | LWK | *p.Ile420Val* | 1 Tolerated | 0.02 Benign |
| c.1437A>C | 0.00046 | ASW | *p.Gln479His* | 0.15 Tolerated | 0.001 Benign |
| | | | | | |
| c.214C>T | 0.00595 | IBS | *p.Leu72Phe* | 0.001 Deleterious | 0.49 Possibly damaging |
| c.923C>T | 0.00046 | MXL | *p.Pro308Leu* | 0.001 Deleterious | 0.642 Possibly Damaging |
| c.1066G>A | 0.00046 | CHS | *p.Val356Met* | 0 Deleterious | 0.999 Probably Damaging |
| c.1250C>T | 0.00137 | JPT | *p.Pro417Leu* | 0.01 Deleterious | 0.503 Possibly Damaging |
| c.1367A>C | 0.00046 | TSI | *p.Tyr456Ser* | 0.19 Tolerated | 0.673 Possibly Damaging |

[1] *Global MAF: Global minor allele frequency*

[2] *SIFT scores are deleterious if less than or equal to 0.05 and tolerated if greater than 0.05.*

[3] *PolyPhen-2 prediction score can range from 0 (benign) to 1 (probably damaging) with a default cut-off of 0.432.*

## 4.5 Discussion

The purpose of this study was to better understand the incidence of Sandhoff disease and the frequency of Sandhoff disease causing mutations in northern Saskatchewan. When 1561 individuals were retrospectively assayed using the c.115delG allelic discrimination assay 57 heterozygotes were discovered. Therefore, we estimate the carrier frequency for this allele to be 1:27 in the study population. As a result of finding 4 affected individuals in our retrospective study we estimate the incidence of Sandhoff disease in this population to be 1:390.

The β-hexosaminidase enzyme assay reliably detected 4 Sandhoff disease affected individuals and proved to be semi-reliable for detecting carriers. In order to achieve a 100% positive predictive value for identifying c.115delG carriers, 32% (20 / 62) of the samples that repeatedly fell below the cut-offs that were established in conjunction with the c.115delG assay, did not carry the c.115delG mutation. If genetic sequencing of these 20 samples was not conducted they may have been misclassified as false positives. However, upon further investigation using genetic sequencing, it was revealed that 17 of the 20 samples that did not have the c.115delG allele instead possessed one or more of the nsSNPs listed in Table 4.2. This left < 5% (3 / 62) of samples with β-hexosaminidase activity in the carrier range but no identified mutation in the *HEXB* gene. It is possible that there are *HEXA* variants present in these 3 outliers however, *HEXA* mutation analysis has not been completed at this time.

A total of 4 nsSNPs were found among our study group in addition to the c.115delG mutation. Of these 4 nsSNPs, 2 cause the loss of disulphide bonds in the resulting polypeptide. In silico analysis and online prediction tools indicate the loss of

the disulphide bonds to be detrimental to β-hexosaminidase activity. No report of the c.1601G>T or c.1652G>A polymorphisms could be found in the dbSNP database or 1000 Genomes project data however, there are reports of a c.1601G>A (similar loss of disulphide bond) substitution being responsible for the infantile form of Sandhoff disease in a Japanese patient [18]. As such the available information is indicative that the c.1601G>T and c.1652G>A alleles are disease causing.

The other 2 nsSNPs, c.362A>G and c.619A>G, have more ambiguous outcomes. The c.362A>G variant has been described in the online dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/) and reported to be benign based on transfection experiments [22]. In silico analysis and online prediction tools suggest that this allele is tolerated by the enzyme however, we saw reduced Total Hex activity and elevated % HexA in these individuals. It is possible that this allele causes reduced β-hexosaminidase activity within the range of tolerance so as not to cause the disease phenotype when present in a homozygous or compound heterozygous state.

Originally the c.619A>G variant was implicated as the cause of an adult form of Sandhoff disease in a compound heterozygous patient [17]. Subsequent reports of this allele have shown it to be capable of α/β-subunit dimerisation, incapable of β/β-subunit dimerisation [23], and tolerated in homozygous individuals producing β-hexosaminidase activity consistent with carriers [21, 24] and thus, not responsible for GM2 build up due to β-hexosaminidase A deficiency. Our data correlates with the observations presented in the literature. Three c.619A>G homozygotes were detected along with 9 heterozygotes all showing β-hexosaminidase activity consistent with the range established for Sandhoff disease carriers. However, one individual possessed both c.619A>G and c.1601G>T

variants. Similar to the compound heterozygote patient described by Banerjee et al. [17] the individual we found possessing the c.619A>G and c.1601G>T variants may be at risk for developing an adult form of β-hexosaminidase B deficiency with motor neuron disease.

Interestingly, since our initial report of the c.115delG allele in northern Saskatchewan it has also been detected in France [28]. The c.115delG allele was paternally inherited by a French patient whose father was of Vietnamese descent [28]. This report indicates that the c.115delG allele is more widely disseminated throughout the global population than just northern Saskatchewan and as such our PCR-based assay may be useful for the detection of Sandhoff disease in other populations.

The major advantage of using an assay to measure enzyme activity compared to single variant mutation analysis is that it can detect carriers irrespective of the mutation present. The disadvantage is that the carrier range typically overlaps with the normal range leading to the possibility of missing some carriers and a high false positive rate. We observed <5% of individuals that fell in the range for Sandhoff disease carriers that had no *HEXB* mutations and a further 5% which possessed *HEXB* polymorphisms predicted to be neutral. The possibility of *HEXA* variants present in the samples thought to be false positive carriers may bring down the false positive rate of the MS/MS assay even further. The major advantage of the PCR-based assay is that it does not produce false positives though it can only detect a single variant per primer and probe set and those variants must be known ahead of time. The use of both techniques to screen the northern Saskatchewan population has provided the opportunity to take advantage of the benefits of both methodologies.

In total the number of individuals found to carry a Sandhoff disease causing allele or c.619A>G as detected by our biochemical screening and genetic sequencing analysis was 51 (n = 760). The c.619A>G variant though not the cause of Sandhoff disease was included in this calculation since it has been implicated in adult onset β-hexosaminidase B deficiency. As such the combined carrier frequency for these mutations in northern Saskatchewan is estimated at 1:15. Given the birthrate in this area and the estimated incidence of 1:390 births, a child born with Sandhoff disease every 1 - 2 years on average is to be expected in these communities.

Application of the Hardy-Weinberg Principle to the observed incidence of Sandhoff disease in northern Saskatchewan suggests a carrier frequency for Sandhoff disease causing mutations of 1:39 (calculations not shown). However, when the Chi-squared test is applied to our observed allele frequencies and expected number of normal, affected, and carrier individuals according to the Hardy-Weinberg Principle we fail to reject the null hypothesis that the population is in Hardy-Weinberg equilibrium.

One of our objectives was to compare the frequency of Sandhoff disease causing mutations in the northern Saskatchewan population to frequency estimates in the general population from the literature as well as data we compiled from a global sample of individuals whose genomes were sequenced for the 1000 Genomes project. With this information we can get a sense for how elevated the frequency of Sandhoff disease causing alleles are in northern Saskatchewan compared to other populations. This year alone several groups have used the 1000 Genomes data as a control sample while conducting mutation analysis or determining pathogenic allele frequencies in populations

for various diseases [29-34]. Similarly, we have used the 1000 Genomes data to compare to our analysis of the *HEXB* gene while investigating Sandhoff disease in Saskatchewan.

Of the 19 variants found in the 1000 Genomes dataset 5 have the potential to cause Sandhoff disease. The pathogenicity of several of these variants is corroborated by reports in the literature whereas for others the exact clinical outcome has not been determined. Based on the results from prediction tools the carrier rate of Sandhoff disease causing variants may be as high as 0.0174 in the global sampling. However, due to the error associated with SIFT and Polyphen-2 predictions [14, 15] and without confirming the pathogenicity of all of the potentially pathogenic alleles identified by the 1000 Genomes project the rate may be as low as 0.00549 in the global sample. The carrier rate for Sandhoff disease in the Saskatchewan sample was 0.0460. As such the carrier frequency for Sandhoff disease causing mutations in Saskatchewan is about 2.5-8 times higher than that among individuals sampled by the 1000 Genomes project and ~12-14 times higher than previous estimates made for the general population [4, 5]. Given the high birthrate of Sandhoff disease affected children in northern Saskatchewan an elevated carrier rate as was determined here was expected.

Interestingly, two coding region variants found in the *HEXB* gene from Saskatchewan individuals were also present in the 1000 Genomes dataset. These variants were c.362A>G and c.619A>G. The frequencies of these variants in the Saskatchewan individuals was (number of alleles observed/number of alleles analyzed) 3/48 (0.063) and 16/48 (0.333) for the c.362A>G and c.619A>G variants respectively. In the 1000 Genomes dataset the frequencies were 0.207 and 0.151 for the c.362A>G and c.619A>G variants respectively. The frequency for the c.362A>G variant in northern Saskatchewan

was found to be lower than any of the subpopulations sampled by the 1000 Genomes project, the lowest being 0.125 in the Southern Han Chinese population whereas the frequency for the c.619A>G allele was higher in the Saskatchewan dataset than any of the 1000 Genomes populations. In addition, these two variants had the highest frequencies for any of the variants detected in the *HEXB* gene from the 1000 Genomes sample. The highest frequencies for these variants were Luhya individuals living in Webuye, Kenya for c.362A>G and in the Japanese individuals living in Tokyo for the c.619A>G variant. Both of these variants were ubiquitous throughout all of the populations included in the 1000 Genomes project to various degrees. The presence of these two alleles in all of the populations included in the 1000 Genomes sample suggests that these are ancestral variants that have disseminated globally during human migration. The discrepancy between the allelic frequencies for the two variants that are present in both the Saskatchewan population and the 1000 Genomes dataset is likely the result of random genetic drift over the last ~12,000 years. This is the amount of time since the last migration of people from Eurasia to North America across the land bridge which is theorized to have once connected the two continents [35]. The introduction of European alleles into the gene pool over the last several hundred years is also a likely factor. Furthermore, both of these variants seem to be neutral in terms of selection for or against the allele supporting genetic drift as the cause of the discrepant frequency.

A total of 19 individuals in the 1000 Genomes data set were carriers for an allele predicted to cause Sandhoff disease by either SIFT of Polyphen-2. Thus if all of the variants predicted to be damaging do cause Sandhoff disease then the carrier rate among the 1000 Genomes sample is about 1 in 57 which is significantly higher than Sandhoff

disease carrier estimates in the general population (1 in 276 to 1 in 310) [4, 5]. A carrier rate of 1 in 57 in the general population along with an autosomal recessive pattern of inheritance would suggest an incidence of Sandhoff disease of roughly 1 in 13000 births. Sandhoff disease in the general population is extremely rare with an estimated incidence of 1:422,000 (n = 4.2 million) [4]. This discrepancy may be due to the error associated with SIFT and Polyphen-2 predictions, underestimates of the carrier frequency by the previously published works, or the result of randomly sampling a relatively small subset of the global population. Given the rarity of Sandhoff disease based on its incidence in the general population the carrier rate in the general population should be much lower than the Polyphen-2 and SIFT predictions suggest.

Of the 19 *HEXB* variants found in the 1000 Genomes dataset 5 were missense mutations with deleterious predictions. Two of these have been previously described in the literature as being pathogenic whereas the other three are of questionable consequence. The results from the prediction tools would lead us to believe that these three missense variants will also cause Sandhoff disease. In order to accurately determine the consequence of these three variants without them being present in clinical cases, transfection experiments could be performed to analyze the *HEXB* transcripts and polypeptides produced from these mutant *HEXB* genes for aberrant biochemical behaviour. Determining the effect of these mutations on β-hexosaminidase activity would confer a higher degree of confidence in estimating the frequency of pathogenic variants in the 1000 Genomes sample. Assuming that some of the variants predicted to be deleterious or damaging end up being tolerated then this analysis could explain why

the carrier rate for Sandhoff disease causing mutations in the 1000 Genomes sample was so high.

It is also interesting to note that the combined frequency for all of the *HEXB* variants in the 1000 Genomes data set was 0.468 whereas in the very specific set of Saskatchewan individuals that had the *HEXB* gene sequenced, the frequency of variants was almost twice as high at 0.875. As such the 1000 Genomes data seems to indicate that nearly half of the individuals of the global population will have some sort of benign or pathogenic coding region or splice region variant in their *HEXB* gene. The frequency of variants in the *HEXB* gene from the Saskatchewan dataset is likely higher due to the nature in which this study sample was selected. The Saskatchewan individuals were selected for having aberrant β-hexosaminidase activity therefore introducing a bias in the Saskatchewan sample selection towards individuals with *HEXB* variants.

Reports in the literature for the carrier frequency of Sandhoff disease range from 1:310 in the general population to 1:7 in isolated communities with a high degree of consanguinity [4-9, 36]. Similarly, the Saskatchewan cohort is remote and isolated, with a high degree of consanguinity [10] and has an elevated carrier frequency for Sandhoff disease causing mutations relative to the general population. In other communities where there is a high carrier frequency for Sandhoff disease, programs have been established or proposed to ameliorate the impact of the disease on families and the community [7, 8, 36]. As such, a program aimed at preventing Sandhoff disease in northern Saskatchewan is being considered. If a carrier screening program for Sandhoff disease in Saskatchewan is to be established a multiplex PCR could be developed to detect all of the pathogenic variants discovered in this study in order to facilitate better carrier screening with a single

assay. Although the c.115delG allele is the major variant responsible for the Sandhoff disease in Saskatchewan, other pathogenic variants are present in the population highlighting the need for special consideration when selecting an assay to detect this disease.

## 4.6 References


1. Mahuran, D.J., *Biochemical consequences of mutations causing the GM2 gangliosidoses.* Biochim Biophys Acta, 1999. **1455**(2-3): p. 105-38.
2. Beutler, E., W. Kuhl, and D. Comings, *Hexosaminidase isozyme in type O Gm2 gangliosidosis (Sandhoff-Jatzkewitz disease).* Am J Hum Genet, 1975. **27**(5): p. 628-38.
3. Ikonne, J.U., M.C. Rattazzi, and R.J. Desnick, *Characterization of Hex S, the major residual beta hexosaminidase activity in type O Gm2 gangliosidosis (Sandhoff-Jatzkewitz disease).* Am J Hum Genet, 1975. **27**(5): p. 639-50.
4. Meikle, P.J., et al., *Prevalence of lysosomal storage disorders.* JAMA, 1999. **281**(3): p. 249-54.
5. Cantor, R.M., et al., *Sandhoff disease heterozygote detection: a component of population screening for Tay-Sachs disease carriers. II. Sandhoff disease gene frequencies in American Jewish and non-Jewish populations.* Am J Hum Genet, 1987. **41**(1): p. 16-26.
6. Kleiman, F.E., et al., *Sandhoff disease in Argentina: high frequency of a splice site mutation in the HEXB gene and correlation between enzyme and DNA-based tests for heterozygote detection.* Hum Genet, 1994. **94**(3): p. 279-82.
7. Kaya, N., et al., *GM2 gangliosidosis in Saudi Arabia: Multiple mutations and considerations for future carrier screening.* Am J Med Genet A, 2011. **155**(6): p. 1281-4.
8. Drousiotou, A., et al., *Sandhoff disease in Cyprus: population screening by biochemical and DNA analysis indicates a high frequency of carriers in the Maronite community.* Hum Genet, 2000. **107**(1): p. 12-7.
9. Andermann, E., et al., *Genetic variants of Tay-Sachs disease: Tay-Sachs disease and Sandhoff's disease in French Canadians, juvenile Tay-Sachs disease in Lebanese Canadians, and a Tay-Sachs screening program in the French-Canadian population.* Prog Clin Biol Res, 1977. **18**: p. 161-88.
10. Lowden, J.A., et al., *Carrier detection in Sandhoff disease.* Am J Hum Genet, 1978. **30**(1): p. 38-45.
11. Fitterer, B.B., et al., *A polymerase chain reaction-based genotyping assay for detecting a novel Sandhoff disease-causing mutation.* Genet Test Mol Biomarkers, 2012. **16**(5): p. 401-5.
12. Cantor, R.M., et al., *Sandhoff disease heterozygote detection: a component of population screening for Tay-Sachs disease carriers. I. Statistical methods.* Am J Hum Genet, 1985. **37**(5): p. 912-21.
13. *Human Genome Variation Society. www.hgvs.org/mutnomenrecs-DNA.html#del, accessed June 15, 2011.*
14. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.
15. Choi, Y., et al., *Predicting the functional effect of amino Acid substitutions and indels.* PLoS One, 2012. **7**(10): p. e46688.
16. *A map of human genome variation from population-scale sequencing.* Nature, 2010. **467**(7319): p. 1061-73.

17. Banerjee, P., et al., *Molecular basis of an adult form of beta-hexosaminidase B deficiency with motor neuron disease.* Biochem Biophys Res Commun, 1991. **181**(1): p. 108-15.
18. Kuroki, Y., et al., *A novel missense mutation (C522Y) is present in the beta-hexosaminidase beta-subunit gene of a Japanese patient with infantile Sandhoff disease.* Biochem Biophys Res Commun, 1995. **212**(2): p. 564-71.
19. Faustino, N.A. and T.A. Cooper, *Pre-mRNA splicing and human disease.* Genes Dev, 2003. **17**(4): p. 419-37.
20. Iida, Y., *Categorical discriminant analysis of 3'-splice site signals of mRNA precursors in higher eukaryote genes.* J Theor Biol, 1988. **135**(1): p. 109-18.
21. Zhang, Z.X., et al., *A second, large deletion in the HEXB gene in a patient with infantile Sandhoff disease.* Hum Mol Genet, 1995. **4**(4): p. 777-80.
22. Wakamatsu, N., et al., *A novel exon mutation in the human beta-hexosaminidase beta subunit gene affects 3' splice site selection.* J Biol Chem, 1992. **267**(4): p. 2406-13.
23. Banerjee, P., et al., *Preferential beta-hexosaminidase (Hex) A (alpha beta) formation in the absence of beta-Hex B (beta beta) due to heterozygous point mutations present in beta-Hex beta-chain alleles of a motor neuron disease patient.* J Biol Chem, 1994. **269**(7): p. 4819-26.
24. Redonnet-Vernhet, I., et al., *Significance of two point mutations present in each HEXB allele of patients with adult GM2 gangliosidosis (Sandhoff disease) homozygosity for the Ile207-->Val substitution is not associated with a clinical or biochemical phenotype.* Biochim Biophys Acta, 1996. **1317**(2): p. 127-33.
25. Gomez-Lira, M., et al., *A common beta hexosaminidase gene mutation in adult Sandhoff disease patients.* Hum Genet, 1995. **96**(4): p. 417-22.
26. Maier, T., et al., *The X-ray crystal structure of human beta-hexosaminidase B provides new insights into Sandhoff disease.* J Mol Biol, 2003. **328**(3): p. 669-81.
27. Mark, B.L., et al., *Crystal structure of human beta-hexosaminidase B: understanding the molecular basis of Sandhoff and Tay-Sachs disease.* J Mol Biol, 2003. **327**(5): p. 1093-109.
28. Gaignard, P., et al., *Characterization of seven novel mutations on the HEXB gene in French Sandhoff patients.* Gene, 2013. **512**(2): p. 521-6.
29. Hussin, J., et al., *Rare allelic forms of PRDM9 associated with childhood leukemogenesis.* Genome Res, 2013. **23**(3): p. 419-30.
30. Jouan, L., et al., *Investigation of rare variants in LRP1, KPNA1, ALS2CL and ZNF480 genes in schizophrenia patients reflects genetic heterogeneity of the disease.* Behav Brain Funct, 2013. **9**: p. 9.
31. Pickett, L.A., et al., *Functional consequences of a novel variant of PCSK1.* PLoS One, 2013. **8**(1): p. e55065.
32. Song, C., et al., *A genome-wide scan for breast cancer risk haplotypes among African American women.* PLoS One, 2013. **8**(2): p. e57298.
33. Thompson, E.R., et al., *Analysis of RAD51D in ovarian cancer patients and families with a history of ovarian or breast cancer.* PLoS One, 2013. **8**(1): p. e54772.
34. Wang, E., et al., *Identification of Functional Mutations in GATA4 in Patients with Congenital Heart Disease.* PLoS One, 2013. **8**(4): p. e62138.

35. Greenberg, J.H., C.G. Turner, and S.L. Zegura, *The Settlement of the Americas: A Comparison of the Linguistic, Dental, and Genetic Evidence.* Current Anthropology, 1986. **27**(5): p. 477-497.
36. Warner, T.G., et al., *Prenatal diagnosis of infantile GM 2 gangliosidosis type II (Sandhoff disease) by detection of N-acetylglucosaminyl-oligosaccharides in amniotic fluid with high-performance liquid chromatography.* Prenat Diagn, 1986. **6**(6): p. 393-400.

# Chapter 5: Discussion

In this study the mutations causing Sandhoff disease in northern Saskatchewan and the frequency of those mutations were analyzed using newly designed and validated assays. The PCR-based assay was validated for detecting the c.115delG allele that was found in four recent Sandhoff disease patients in the province. The MS/MS-based assay was validated for detecting Sandhoff disease patients harbouring a variety of *HEXB* mutations. Based on the retrospective screening of newborns from northern Saskatchewan and the subsequent genetic sequencing data, four different pathogenic mutations were found in the population with a combined estimated carrier frequency of ~1:15. Three of these pathogenic variants were described here for the first time. Comparison to the 1000 Genomes data set showed that variants in the *HEXB* gene are common with roughly half of the individuals included in that sample possessing either a coding region or splice region genetic variant.

## **5.1** *Carrier Screening for Sandhoff Disease*

One of the primary goals at the outset of this project was to evaluate the presence of Sandhoff disease in Saskatchewan and collect data so that it could be determined whether or not a prevention program would be warranted and possible. It has been shown that detection of carriers can be accomplished using the methods described here as well as methods described elsewhere [1-8]. However, our methods come with the added benefits outlined in chapter 2 and 3, namely high-throughput, the potential to be multiplexed with other assays for cost and labor savings [9], no need for additional sample collection in most cases, and the ability to collect the sample as a dried blood spot

in a remote location and have it mailed to the testing facility preventing the need for patient travel.

Furthermore, the carrier frequency of Sandhoff disease in Saskatchewan (~1:15) and the severity of the disease are similar to the carrier frequencies and severity for other genetic diseases that are already screened for either as a means of prevention or by newborn screening for early treatment in other jurisdictions. Some examples include the Tay-Sachs disease prevention program among the Ashkenazi Jewish population with an estimated carrier frequency of about 1:27 [10], screening for cystic fibrosis among the general population with an estimated carrier frequency of 1:38 though this frequency is highly variable among different ethnicities [11], screening for β-thalassemia and α-thalassemia in Sardinia with reported carrier frequencies of 1:10 and 1:4 respectively [12], and screening for aspartylglucosaminuria in Finnish couples with an estimated carrier frequency of 1:36 [13], among others. Although the population studied in Saskatchewan is smaller than these other examples the carrier frequency and severity of disease are sufficiently high to warrant consideration for some type of intervention.

Intervention in the form of a carrier screening program aimed at preventing Sandhoff disease is technically feasible as shown by the work presented here. However, prior to the implementation of such a program a cost-benefit analysis should be conducted and additional details would need to be addressed such as when to offer testing and how to educate participants as to the meaning of their test results. Although premarital screening has worked for reducing the incidence of Tay-Sachs disease among Ashkenazi Jews, offering testing to individuals upon reaching reproductive age may be more beneficial for these Saskatchewan communities. Programs targeted to high-school

aged individuals have proven effective for reducing the level of Tay-Sachs disease among French-Canadians [14] and have had high uptake and retention of knowledge among Sardinians [12]. As such offering Sandhoff disease carrier screening to high-school aged individuals living in these northern communities may be appropriate.

Another factor to consider is migration. The communities that were part of this study are relatively small and isolated. The population size in these centers is relatively stable based on Statistics Canada census data. Along with the high birthrate this data suggests that people from these centers move elsewhere perhaps due to a high level of unemployment. A nearby larger center may receive a large percentage of these individuals and therefore have a lesser and yet still significant carrier frequency for Sandhoff disease. A carrier frequency study of this larger center may show that the pathogenic alleles have disseminated to a larger portion of the province than what was originally studied and suggest that a potential Sandhoff disease carrier screening program be expanded to a larger portion of northern Saskatchewan than just the small communities that were included in this study.

The laboratory methods presented here have laid the foundation for a preventative carrier screening program for Sandhoff disease in northern Saskatchewan. Now that other pathogenic variants besides the c.115delG mutation have been found in the northern population the PCR-based assay could be modified to detect all four of the mutations in a single multiplex reaction. Such a multiplex should detect the vast majority of Sandhoff disease carriers in the population. However, the presence of other disease causing variants is still a possibility. Furthermore, neither of the two assays were ideal for carrier screening alone. Even a multiplex PCR assay will still miss other pathogenic variants

and the carrier range for the MS/MS based assay has significant overlap with the normal range. Therefore tandem molecular and biochemical based screening is still appropriate. Biochemical outliers that do not possess any of the known variants should continue to be worked up by genetic sequencing.

One major shortcoming of the biochemical assay for measuring β-hexosaminidase activity was its inability to detect Tay-Sachs disease. The use of the labelled and unlabelled substrates along with heat inactivation of β-hexosaminidase A in one of two reactions should allow for the determination of each enzyme's individual activity and therefore distinction of both Tay-Sachs and Sandhoff diseases. The inability to detect Tay-Sachs disease using dried blood spots from known Tay-Sachs patients limits the utility of the assay to only detecting Sandhoff disease (GM2 gangliosidosis type II) and thereby restricts its use by laboratories who want to detect the other GM2 gangliosidosis disorders (Tay-Sachs disease and the AB variant or GM2 activator protein deficiency).

The MS/MS-based assay's inability to detect Tay-Sachs disease was also investigated. Preliminary data suggests that optimization of the dried blood spot extraction buffer to a pH 4 rather than a pH 7 may overcome this issue. From inception other β-hexosaminidase assays have used a buffer of pH 4 [15]. Due to an oversight when developing the β-hexosaminidase assay described in chapter 3, a common dried blood spot extraction buffer with a pH of 7 for measuring non-lysosomal enzymes and metabolites was used to elute the blood from the filter paper. The heat-inactivation fraction of this elution was then heated to abolish β-hexosaminidase A at a pH of 7. The incorrect buffer pH combined with heating of the β-hexosaminidase enzymes during the assay is likely detrimental to β-hexosaminidase B while also denaturing β-

hexosaminidase A as intended due to the combined stress on the enzymes caused by deviation from thermal and pH optima. If the combined stress of heating at a non-optimal pH is having an effect on the measured β-hexosaminidase B activity then our values for both β-hexosaminidase A and β-hexosaminidase B may be inaccurate. As such the values for β-hexosaminidase A may be higher than actual while β-hexosaminidase B values lower.

Another possibility for the assay's inability to detect Tay-Sachs disease is the specificity of each major form of the enzyme for the synthetic substrate. It is possible that based on the chemical structure of the substrate β-hexosaminidase B may be better able to hydrolyze the substrate than β-hexosaminidase A. To test this theory each form of the enzyme could be isolated and quantified then their activities for hydrolyzing the synthetic substrate measured and compared. As such, studies of the specificity of β-hexosaminidase A and B for this new synthetic substrate may be valuable.

Once the cause of the assay's inability to detect Tay-Sachs disease is found any protocol changes to the biochemical assay to allow for detection of Tay-Sachs will require re-validation including new normal range values for β-hexosaminidase A and β-hexosaminidase B. Regardless, the assay as described in chapter 3 was validated for detecting Sandhoff disease proving it useful for our study.

Furthermore, we set out to design assays that could collect the data for our retrospective study but also translate into optimized diagnostic assays for a laboratory to use for preventative carrier screening. Given the research into therapy for gangliosidosis, the potential exists for an approved Sandhoff disease treatment in the coming years. At such a time as a treatment becomes available it would also be possible to use these assays

for newborn screening in order to identify affected newborns prior to the manifestation of symptoms and offer treatment. Additionally, these methods could be used prior to the approval of therapy for Sandhoff disease as a newborn screening test in order to identify affected newborns so that they may have the option of participating in clinical trials.

**5.2** *Elevated Sandhoff disease-causing allele frequency*

The frequency of an allele in a population can become elevated by several mechanisms such as a founder event, random genetic drift, or selection for the allele. The prevailing theory thus far has been that the high incidence of Sandhoff disease in northern Saskatchewan is due to a founder effect caused by introduction of the disease allele by Europeans during early settlement of Western Canada with genetic drift towards a high frequency of the disease causing allele enhanced by the small isolated nature of the communities where it is present along with the high degree of consanguineous matings.

However, there has also been suggestion that some lysosomal storage disease causing alleles, and specifically those causing GM2 gangliosidosis, may confer a heterozygote advantage in the form of resistance to tuberculosis [16, 17]. Historically, Saskatchewan has had tuberculosis epidemics most notably around the time of the founding of the Anti-Tuberculosis League in 1911 when roughly 1000 people a year were being infected with tuberculosis and the mortality rate due to infection was around 1-2 per day [18]. Communities in northern Saskatchewan are mostly made up of First Nations people or people of mixed First Nation and European descent. Roughly 40 years earlier Tuberculosis was rare amongst First Nations peoples but by the 1880's Tuberculosis had become the number one cause of death among First Nations people

living on the Canadian prairies [19]. As such, tuberculosis infection would have applied significant selective pressure for any genetic alleles contributing to resistance against the infection in heterozygotes. The theory that the c.115delG allele confers a heterozygote advantage in the form of Tuberculosis resistance would seem to fit with the historical selective pressure of Tuberculosis epidemics in these populations and the high Sandhoff disease carrier rate in these communities. It is possible that the high incidence of the c.115delG alleles in northern Saskatchewan is not only due to a founder effect but also likely due to selective pressure and heterozygote advantage as has been suggested for the high frequency of lysosomal storage disease causing alleles in other populations [20].

**5.3** *Power of the 1000 Genomes and value for the future of medical genetics*

Thus far data provided by the 1000 Genomes project has been used in medical genetics studies, genome-wide association studies, as an exclusionary filter for monogenic diseases, tumour studies, the advancement of genotyping arrays, as well as studies of allele frequency for rare disease causing variants [21]. Here we used the 1000 Genomes project data to compare variants in the *HEXB* gene from our Saskatchewan sample to the global population. Overall the rate of variation in the *HEXB* gene was higher in the Saskatchewan sample than globally which was not unexpected since Sandhoff disease is rare in the general population and the Saskatchewan sample was selected with bias towards *HEXB* variants.

The 1000 Genomes project was designed to characterise over 95% of genetic variants that are present in the human population and have allele frequencies of >1% as well as coding region variants with allele frequencies as low as 0.1% [21]. However,

variants of low frequency and rare variants as defined by the 1000 Genomes project (<5% and <0.5% respectively) greatly out number the amount of higher frequency variants [21]. As such the power of the 1000 Genomes analysis to detect small variants (SNPs and small indels) such as those present in *HEXB* gene from northern Saskatchewan individuals that contribute to rare disorders such as Sandhoff disease is of interest. The power of the 1000 Genomes analysis to detect small variants that were present in the pilot sample a single time has been estimated at ~25% up to as much as 90% if the variant was present 5 times in the sample [21]. This power is largely based on the likelihood of a variant being present in the sample and the quality and coverage of sequencing which is dictated by the technology used [21]. As such the ability of the 1000 Genomes project to detect extraordinarily rare variants may be limited. It will be interesting to see if any of the disease causing alleles found in northern Saskatchewan are picked up by the 1000 Genomes project once all of the 2500 genomes have been completely sequenced. If the Sandhoff disease causing variants found in Saskatchewan do not show up in the 2500 genomes it will suggest that those allele frequencies are <0.1% globally. The rarity and diversity of disease causing alleles again highlights the limitations of PCR-based screening where only a small and specific subset of mutations can be detected compared to biochemical assays and whole genome sequencing which do not discriminate based on the variants present.

Furthermore, tests are currently being developed to detect recessive alleles in consanguineous couples using high-throughput next generation sequencing technologies such as those used to generate the 1000 Genomes data [22]. Consanguineous couples are at increased risk for having offspring with autosomal recessive disorders such as

Sandhoff disease due to the sharing of alleles through descent. By sequencing the genomes of consanguineous couples, shared heterozygosity for pathogenic alleles in many genes could allow for the prevention of a broad spectrum of disorders including Sandhoff disease. Prevention of such disorders would be accomplished by identifying a couple's risk for specific autosomal recessive disorders and allowing for pre-conception planning. This approach could be further expanded to all couples not only those with a high degree of consanguinity. As such future preventative screening for Sandhoff disease in northern Saskatchewan where the rate of consanguinity is high may involve genome sequencing using high-throughput next generation technologies similar to those used to generate the 1000 Genomes data.

### **5.4** *Future Work*

1. Implement a carrier screening program in northern Saskatchewan
2. Modify the MS/MS assay to also detect Tay-Sachs disease
3. Multiplex the PCR assay to detect the other pathogenic variants found in northern Saskatchewan
4. Screen a sample of nearby northern Saskatchewan communities to check for Sandhoff disease carriers elsewhere
5. Network with groups studying gangliosidosis therapy so that future affected individuals may have the opportunity to be included in clinical trials

## **5.5** *Conclusion*

We report on the high incidence of Sandhoff disease in Saskatchewan relative to reports of the incidence among other populations and those found in the 1000 Genomes data. Evidence is presented which supports the case for a preventative carrier screening program to reduce the number of Sandhoff disease affected births in the province. The genetic variants responsible for the disease and their frequencies are described. Genetic analysis of the *HEXB* gene from affected patients revealed a common c.115delG variant shared among 4 affected patients however, a total of three mutations causing Sandhoff disease were found in the population. The carrier rate for these pathogenic alleles is estimated to be 12-14 times higher in northern Saskatchewan than in the general population which is comparable to other genetic diseases that are screened for as a means of prevention. Based on the methodologies presented and the evidence gathered for the incidence and carrier frequency, a preventative carrier screening program for Sandhoff disease in northern Saskatchewan is recommended.

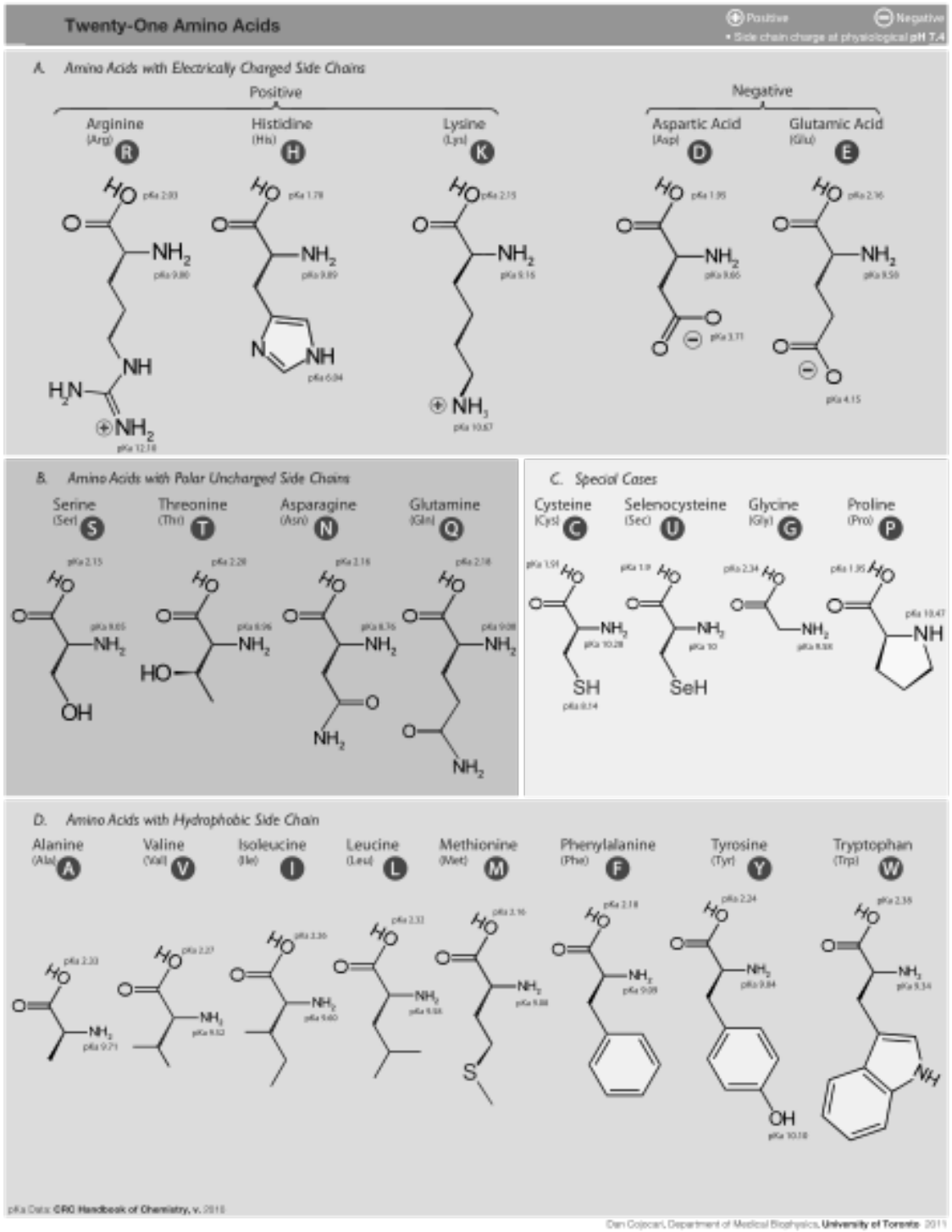
**5.6** References:


1. Carmody, P.J., M.C. Rattazzi, and R.G. Davidson, *Tay-Sachs disease--the use of tears for the detection of heterozygotes.* N Engl J Med, 1973. **289**(20): p. 1072-4.
2. Nakagawa, S., et al., *Human hexosaminidase isozymes. Assay of platelet activity for heterozygote identification during pregnancy.* Clin Chim Acta, 1978. **88**(2): p. 249-56.
3. Nakagawa, S., et al., *Human hexosaminidase isozymes. III. Distribution and activity of isozymes in peripheral blood leukocytes and platelets.* J Lab Clin Med, 1978. **91**(6): p. 922-8.
4. Navon, R. and B. Padeh, *Urinary test for identification of Tay-Sachs genotypes.* J Pediatr, 1972. **80**(6): p. 1026-30.
5. O'Brien, J.S., et al., *Tay-Sachs Disease.* New England Journal of Medicine, 1970. **283**(1): p. 15-20.
6. Saifer, A., J. Amoroso, and G. Perle, *Identification of Tay-Sachs by hexosaminidase analysis of urine and tear samples.* Adv Exp Med Biol, 1976. **68**: p. 339-66.
7. Singer, J.D., E. Cotlier, and R. Krimmer, *Hexosaminidase A in tears and saliva for rapid identification of Tay-Sachs disease and its carriers.* Lancet, 1973. **2**(7838): p. 116-9.
8. Suzuki, Y., P.H. Berman, and K. Suzuki, *Detection of Tay-Sachs disease heterozygotes by assay of hexosaminidase A in serum and leukocytes.* J Pediatr, 1971. **78**(4): p. 643-7.
9. Spacil, Z., et al., *High-throughput assay of 9 lysosomal enzymes for newborn screening.* Clin Chem, 2013. **59**(3): p. 502-11.
10. Kaback, M.M., *Population-based genetic screening for reproductive counseling: the Tay-Sachs disease model.* Eur J Pediatr, 2000. **159 Suppl 3**: p. S192-5.
11. Rohlfs, E.M., et al., *Cystic fibrosis carrier testing in an ethnically diverse US population.* Clin Chem, 2011. **57**(6): p. 841-8.
12. Cao, A., et al., *Thalassaemia and glucose-6-phosphate dehydrogenase screening in 13- to 14-year-old students of the Sardinian population: preliminary findings.* Community Genet, 2008. **11**(3): p. 121-8.
13. Syvanen, A.C., et al., *Convenient and quantitative determination of the frequency of a mutant allele using solid-phase minisequencing: application to aspartylglucosaminuria in Finland.* Genomics, 1992. **12**(3): p. 590-5.
14. Mitchell, J.J., et al., *Twenty-year outcome analysis of genetic screening programs for Tay-Sachs and beta-thalassemia disease carriers in high schools.* Am J Hum Genet, 1996. **59**(4): p. 793-8.
15. Leaback, D.H. and P.G. Walker, *Studies on glucosaminidase. 4. The fluorimetric assay of N-acetyl-beta-glucosaminidase.* Biochem J, 1961. **78**: p. 151-6.
16. Myrianthopoulos, N.C. and S.M. Aronson, *Population dynamics of Tay-Sachs disease. I. Reproductive fitness and selection.* Am J Hum Genet, 1966. **18**(4): p. 313-27.
17. Myrianthopoulos, N.C., A.F. Naylor, and S.M. Aronson, *Founder effect in Tay-Sachs disease unlikely.* Am J Hum Genet, 1972. **24**(3): p. 341-2.

18. *The Encyclopedia of Saskatchewan.* *http://esask.uregina.ca/entry/tuberculosis_control.html. Accessed: April 23, 2013.*
19. Daschuk, J.W., P. Hackett, and S. MacNeil, *Treaties and tuberculosis: First Nations people in late 19th-century western Canada, a political and economic transformation.* Can Bull Med Hist, 2006. **23**(2): p. 307-30.
20. Chakravarti, A. and R. Chakraborty, *Elevated frequency of Tay-Sachs disease among Ashkenazic Jews unlikely by genetic drift alone.* Am J Hum Genet, 1978. **30**(3): p. 256-61.
21. *A map of human genome variation from population-scale sequencing.* Nature, 2010. **467**(7319): p. 1061-73.
22. Makrythanasis, P. and S.E. Antonarakis, *High-throughput sequencing and rare genetic diseases.* Mol Syndromol, 2012. **3**(5): p. 197-203.

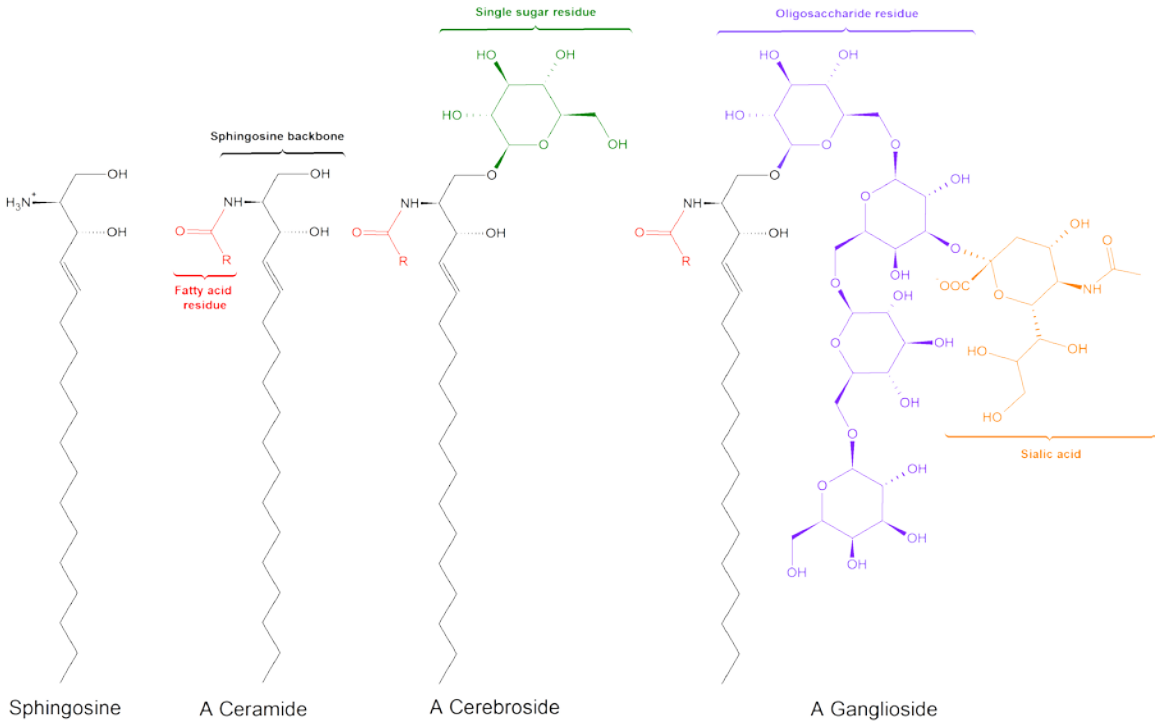
## Appendix

A. IUPAC amino acid codes (one letter, 3 letter, names)

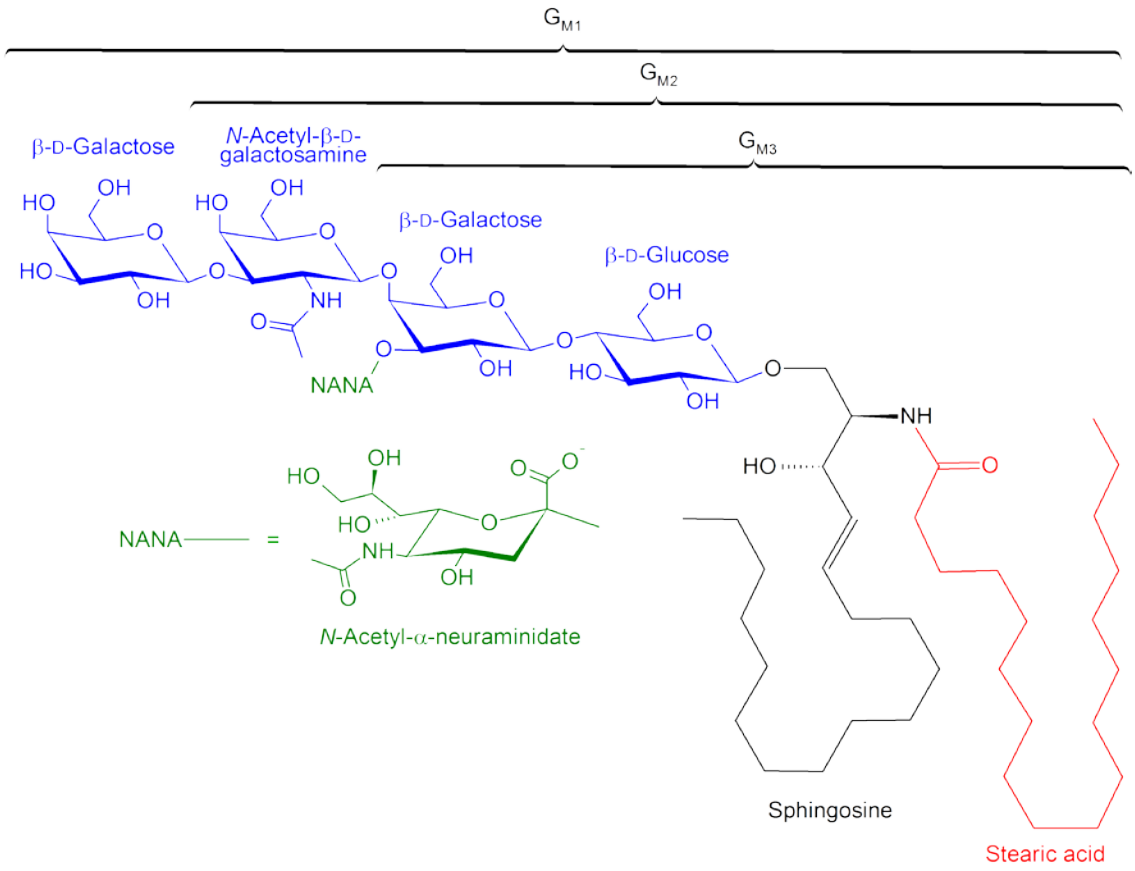



Taken From: http://en.wikipedia.org/wiki/File:Amino_Acids.svg

B. Ganglioside structures




Source: http://en.wikipedia.org/wiki/File:Sphingolipids_general_structures.png




Source: http://en.wikipedia.org/wiki/File:Structure_of_GM1,_GM2,_GM3.png

*Note GA2 is essentially GM2 lacking the NANA group.

C. University of Regina Research Ethics Board Approval

UNIVERSITY OF REGINA

OFFICE OF RESEARCH SERVICES

MEMORANDUM

DATE: November 24, 2009

TO: Mr. Braden Fitterer
Molecular Diagnostics
Saskatchewan Disease Control Laboratory
3211 Albert Street, Regina SK S4S5W6

FROM: Dr. Bruce Plouffe
Chair, Research Ethics Board

Re: **Genetic Mutation in the HEXB Gene Causing Sandhoff Disease in the Saskatchewan Population: Method Development for Genetic and Biochemical Screening and Allelic Frequency within Specific Communities (File # 38S0910)**

---

Please be advised that the University of Regina Research Ethics Board has reviewed your proposal and found it to be:

☑ 1. APPROVED AS SUBMITTED. Only applicants with this designation have ethical approval to proceed with their research as described in their applications. For research lasting more than one year (Section 1F). **ETHICAL APPROVAL MUST BE RENEWED BY SUBMITTING A BRIEF STATUS REPORT EVERY TWELVE MONTHS.** Approval will be revoked unless a satisfactory status report is received. Any substantive changes in methodology or instrumentation must also be approved prior to their implementation.

☐ 2. ACCEPTABLE SUBJECT TO MINOR CHANGES AND PRECAUTIONS (SEE ATTACHED). Changes must be submitted to the REB and approved prior to beginning research. Please submit a supplementary memo addressing the concerns to the Chair of the REB.** Do not submit a new application. Once changes are deemed acceptable, ethical approval will be granted.

☐ 3. ACCEPTABLE SUBJECT TO CHANGES AND PRECAUTIONS (SEE ATTACHED). Changes must be submitted to the REB and approved prior to beginning research. Please submit a supplementary memo addressing the concerns to the Chair of the REB.** Do not submit a new application. Once changes are deemed acceptable, ethical approval will be granted.

☐ 4. UNACCEPTABLE AS SUBMITTED. The proposal requires substantial additions or redesign. Please contact the Chair of the REB for advice on how the project proposal might be revised.

O. Bruce Plouffe

Dr. Bruce Plouffe
cc: Dr. Denis C. Lehotay, Dr. Christopher Yost, Dr. Nick Antonyshin
** supplementary memo should be forwarded to the Chair of the Research Ethics Board at the Office of Research Services (Research and Innovation Centre, Room 109) or by e-mail to research.ethics@uregina.ca

Phone: (306) 585-4775
Fax: (306) 585-4893