Biochemistry and Genetics of Tay-Sachs Disease

Roy A. Gravel, Barbara L. Triggs-Raine and Don J. Mahuran

ABSTRACT: Tay-Sachs disease is one of the few neurodegenerative diseases of known cause. It results from mutations of the HEXA gene encoding the α subunit of β-hexosaminidase, producing a destructive ganglioside accumulation in lysosomes, principally in neurons. With the determination of the protein sequence of the α and β subunits, deduced from cDNA sequences, the complex pathway of subcellular and lysosomal processing of the enzyme has been determined. More recently, detailed knowledge of the gene structure has allowed the determination of specific mutations causing Tay-Sachs disease. The high incidence of the disease in Ashkenazi Jews is attributed predominantly to three mutations present in high frequency, while in non-Jews some two dozen mutations have been identified thus far. The cataloguing of mutations has important implications for carrier screening and prenatal diagnosis for Tay-Sachs disease.

Résumé: Étude biochimique et génétique de la maladie de Tay-Sachs. La maladie de Tay-Sachs est une des seules maladies neurodégénératives dont on connaît la cause. Elle résulte de mutations du gène HEXA codant la sous-unité α de l’hexosaminidase-β, produisant une accumulation destructrice de ganglioside dans les lysosomes, principalement dans les neurones. À la suite de la détermination de la séquence protéique des sous-unités α et β, déduite à partir des séquences d’ADN complémentaire, la voie complexe de maturation subcellulaire et lysosomiale de l’enzyme a été déterminée. Récemment, la connaissance détaillée de la structure du gène a permis l’identification de mutations spécifiques causant la maladie de Tay-Sachs. La forte incidence de la maladie chez les juifs Ashkenazi est attribuée à la fréquence élevée de trois mutations prédominantes, alors que chez les non-juifs pas moins de deux douzaines de mutations ont été identifiées à date. Il est important de répertorier ces mutations pour le dépistage des porteurs et le diagnostic prénatal de la maladie de Tay-Sachs.


Tay-Sachs disease (GM2 gangliosidosis, B variant or type 1) is an autosomal recessive lysosomal storage disorder that results from mutation of the HEXA gene encoding the α subunit of β-hexosaminidase A (Hex A, structure αβ). In the absence of the enzyme activity, the lysosomal swelling and neuronal dysfunction resulting from accumulation of the GM2 ganglioside substrate lead to a progressive neurologic degeneration (reviewed in 1-3). There are many clinical variants of the disease, from infantile, lethal forms to variants compatible with survival into adulthood. The severity of the disease generally correlates with the level of residual Hex A activity,4 a finding implicating the existence of considerable genetic heterogeneity. This review will summarize our knowledge of the biochemistry and genetics of Tay-Sachs disease, which are now leading to an explosion in mutation identification with major implications for carrier screening and clinical understanding.

Structure of β-hexosaminidase

Lysosomal hexosaminidase occurs in two principal forms, Hex A and Hex B. Hex A is made up of one α and one β subunit, while Hex B is made up of two β subunits. The α-subunit is encoded by the HEXA gene on chromosome 15 and the β-subunit by the HEXB gene on chromosome 5. Comparison of the cDNA sequence and predicted amino acid sequence of the α and β subunits shows nearly 60% identity.5 The structural organization of the HEXA and HEXB genes is also similar. They are split into 14 exons spanning about 35 kb and 40 kb, respectively, and all but the first splice junction are located at identical positions in the aligned sequence.6 Thus, the genes appear to share a common ancestral origin.

Like other lysosomal hydrolases, hexosaminidase is synthesized through a subcellular pathway that includes the endoplasmic reticulum (ER) and Golgi and terminates in the lysosome or by secretion of the enzyme from the cell.3 Typical of glycoproteins, the first step is commitment to synthesis in the rough ER. This is mediated by a hydrophobic signal peptide at the amino terminus of the prepropeptide which allows translocation of the nascent polypeptide into the lumen of the ER. The α signal peptide is 22 amino acid residues long,7 while that of the β subunit is 42 amino acids.8 The signal peptides are cleaved on entry into the ER, permitting continued synthesis of the remaining, freely soluble α and β polypeptides.
The propolypeptides undergo co-translational glycosylation on selected asparagine (Asn) residues in the ER. This involves direct transfer of a mannone-rich moiety, Glc3Man9GlcNAc2, via a dolichol intermediate to the growing polypeptide. Removal of the three terminal glucose residues from the high-mannose structures and the initial trimming of one mannone residue also occurs in the ER. In addition, formation of the intrapolypeptide disulfide bonds and the association of the pro-α and pro-β chains occur. These events result in the formation of a catalytically active pro-Hex.9,10

The specific targeting of lysosomal enzymes, including hexosaminidase, to the lysosome requires generation of phosphomannosyl recognition markers. They are formed by the sequential action of a phosphotransferase that transfers UDP-N-acetylglucosamine to selected mannone residues11 and a phosphodiesterase oc-N-acetylglucosaminidase, that removes N-acetylglucosamine12 exposing the phosphomannosyl structures. These reactions occur in the late ER and cis-Golgi.13 Of the four oligosaccharide side chains in the P-subunit,14 the first and fourth are preferentially phosphorylated compared to the second and third.15

Kornfield et al.16 showed that the phosphotransferase has a specific affinity for lysosomal enzymes and does not phosphorylate other glycoproteins, thereby providing a mechanism for segregating proteins destined for the lysosome from those that are exported elsewhere. Biochemical and kinetic data suggest that lysosomal proteins have a common protein domain that enables them to bind to phosphotransferase.16 The protein domain does not seem to be associated with a specific amino acid sequence, although at least one lysine residue appears to play an important part in phosphotransferase recognition.17

In the trans-Golgi network, hexosaminidase containing phosphomannosyl recognition markers combines with the mannone-6-phosphate receptor to form a complex that will be shuttled to the lysosome.18 The lysosomal protein/mannose-6-phosphate receptor complex is delivered by clathrin coated vesicles to either a permanent "packaging" organelle that then transfers the free protein to a lysosome19 or a transient preslysosomal/late endosomal compartment that fuses with or becomes, through acidification, a lysosome.20 Dissociation of the receptor-ligand complex occurs due to the increased acidity of the organelle. The released receptors are recycled back to the trans-Golgi network to ferry additional ligands. Some of the proenzyme does not bind to mannone-6-phosphate receptors and is secreted, appearing in serum or other fluids and in cell culture media.

In the lysosome, hexosaminidase is subjected to extensive proteolytic and glycosidic modification. Depending on the extent of exposure to the lysosomal milieu, the oligosaccharides are variously cleaved, some to a limited extent, others much more so. The fourth oligosaccharide side chain of the P-subunit, for example, is reduced to a single asparagine-linked N-acetylglucosamine residue.14

The protein modifications occurring in the lysosome are equally extensive. The 67 kDa percursor a propolypeptide is processed to a 7 kDa (αn) N-terminal segment and a 54 kDa "mature" polypeptide.21 This is accompanied by the removal of 16 or 17 amino acid residues at the interval between the two sequences. Similarly, the 63 kDa β subunit is cleaved into βp, βr and βb polypeptides, also with removal of internal sequences.21,22 The biological reason for the proteolytic processing is not clear. These events are not required for enzymatic activity since the percursor forms are catalytically active.23 It is possible that the partially degraded enzyme resulting from exposure to the harsh environment of the lysosome is the resistant, but still functional, product of the protein’s adaptation to its lysosomal role.

Function of Hexosaminidase

Hexosaminidase cleaves the glycosidic bond, at the nonreducing end, of terminal β-N-acetylgalactosamine or β-N-acetylgalactosamine moieties of glycoconjugates, including glycolipids, glycoproteins, and glycosaminoglycans.2 While Hex A and Hex B are able to hydrolyse many of the same substrates, only Hex A has the capacity to utilize negatively charged substrates, including the primary substrate GM2 ganglioside. GM2 ganglioside also requires the presence of a water-soluble, lipid-binding protein cofactor known as the "GM2 activator".24 It forms a 1:1 complex with GM2 ganglioside that renders the whole complex water soluble. It acts both as a transport protein to deliver the ganglioside substrate to the lysosome and, as well, interacts with the Hex A to allow cleavage of the glycosidic bond by the α subunit.25 A C DNA encoding the GM2 activator has recently been isolated.26

Furthermore, only Hex A hydrolyses other naturally occurring, negatively charged substrates, such as terminal β-linked N-acetylgalactosamine-6-sulfate contained in keratan sulfate and chondroitin or dermanat sulfates.2 This difference in specificity is probably due to a unique binding site on the α-subunit capable of accommodating the negatively charged group of the substrate. This has been supported by studies using a GlcNAc-6-sulfate containing artificial substrate which is hydrolysed by Hex A and not by Hex B.25

Tay-Sachs Disease

Clinically, Tay-Sachs disease is associated with a wide spectrum of age at onset and expression.2 The classical infantile disease is characterized by onset at 3-5 months of age with developmental arrest, hyperacussis, macular cherry red spots and blindness, intractable seizures, and progressive neurological deterioration culminating in death at 3-5 years of age. Later onset forms are extremely variable. The clinical course may be dominated by signs of dementia and seizures, cerebellar dysfunction, atypical spinocerebellar degeneration, atypical motor neuron disease, dystonia, or acute psychosis.2,27

The immediate importance of identifying mutations in Tay-Sachs disease is to add DNA-based diagnostic testing to current enzymatic methods and to link specific mutations with defined clinical phenotypes. Both objectives make it desirable to identify all mutations causing deficiency of enzyme activity. This is proving to be a daunting task with a large number of mutations already defined (Table 1).

The effort to identify alleles in Tay-Sachs disease began with populations showing a high incidence of the disease. The first to yield its genetic basis was a 7.6 kb deletion of the 5' end of the HEXA gene prevalent in French Canadians of eastern Quebec.46 The mutation, possibly derived by misaligned recombination of nearby Alu sequences, removes the putative promoter region, exon 1 and part of intron 1. It is incompatible with the synthesis of the mRNA and protein product. Patients homozygous for this
A concerted effort to identify the mutation responsible for the disease in Ashkenazi Jews led to a simultaneous report by three groups that the mutation was at the donor splice site in intron 12.\textsuperscript{40-42} This mutation was shown to result in abnormal splicing and consequent instability of the mRNA.\textsuperscript{42} The effective result was absence of normal mRNA and, therefore, the absence of the α-subunit product. More significant than the actual identification of the mutation was the remarkable discovery that it was not the only one responsible for the infantile disease in Ashkenazi Jews. It had generally been anticipated that a single mutation, derived through a founder and thought to have become elevated in frequency by genetic drift\textsuperscript{48} or selective advantage among carriers\textsuperscript{49,50} would be identified. Such simplistic expectations are rapidly disappearing as many diseases, showing high carrier frequencies in particular ethnic groups or geographic isolates, are proving to have an abundance of distinct alleles.

To date, three common mutations have been found to account for Ashkenazi Tay-Sachs disease (Table 1): a 4 bp insertion in exon 11,\textsuperscript{39} the splice mutation in intron 12, and an amino acid substitution in exon 7.\textsuperscript{34-36} The first two produce the "classical" infantile disease. Neither produces a detectable enzyme by standard Northern blotting, although nuclear transcription has been shown to be normal for the insertion mutation\textsuperscript{51} and is likely so for the splice mutation. The exon 7 mutation, so far only seen in Jews coupled with one of the other two alleles, produces the adult disease with onset in the second or third decade. Clinically, there is lower motor neuron, pyramidal tract and cerebellar involvement and, in some cases, psychosis.\textsuperscript{27} Homozygous patients have now been identified among non-Jews, and they show clinical expression at the mild end of the adult disease phenotype.\textsuperscript{36}

Table 1. Mutations in the \textit{HEXA} Gene.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Location</th>
<th>Result</th>
<th>Class</th>
<th>Origin</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>G→T</td>
<td>-1 IVS-4</td>
<td>abnormal splicing</td>
<td>Infantile</td>
<td>Black</td>
<td>28</td>
</tr>
<tr>
<td>G509→A</td>
<td>Exon 5</td>
<td>Arg170→Gln</td>
<td>Infantile</td>
<td>Japanese</td>
<td>28</td>
</tr>
<tr>
<td>C532→T</td>
<td>Exon 5</td>
<td>Arg178→Cys</td>
<td>Infantile</td>
<td>Czechoslovakian</td>
<td>30</td>
</tr>
<tr>
<td>G533→A</td>
<td>Exon 5</td>
<td>Arg178→His</td>
<td>Infantile</td>
<td>Diverse</td>
<td>31, 30</td>
</tr>
<tr>
<td>G570→A</td>
<td>Exon 5</td>
<td>abnormal splicing</td>
<td>Juvenile</td>
<td>Tunisian</td>
<td>32</td>
</tr>
<tr>
<td>G749→A</td>
<td>Exon 7</td>
<td>Gly250→Asp</td>
<td>Juvenile</td>
<td>Lebanese</td>
<td>33</td>
</tr>
<tr>
<td>G805→A</td>
<td>Exon 7</td>
<td>Gly269→Ser</td>
<td>Adult</td>
<td>Diverse</td>
<td>34-36</td>
</tr>
<tr>
<td>ΔTTC910-912</td>
<td>Exon 8</td>
<td>ΔPhe304 or Phe305</td>
<td>Infantile</td>
<td>Moroccan</td>
<td>37</td>
</tr>
<tr>
<td>G1260→C</td>
<td>Exon 11</td>
<td>Trp420→Cys</td>
<td>Infantile</td>
<td>Jewish</td>
<td>28</td>
</tr>
<tr>
<td>+TATC1278</td>
<td>Exon 11</td>
<td>Frameshift</td>
<td>Infantile</td>
<td>Irish/German</td>
<td>38</td>
</tr>
<tr>
<td>G→C</td>
<td>Exon 11</td>
<td>+1 IVS-12</td>
<td>abnormal splicing</td>
<td>Infantile</td>
<td>Ashkenazi Jewish</td>
</tr>
<tr>
<td>G1444→A</td>
<td>Exon 13</td>
<td>Glu482→Lys</td>
<td>Infantile</td>
<td>Ashkenazi Jewish</td>
<td>40-42</td>
</tr>
<tr>
<td>G1496→A</td>
<td>Exon 13</td>
<td>Arg499→His</td>
<td>Juvenile</td>
<td>Italian</td>
<td>43</td>
</tr>
<tr>
<td>ΔC1510</td>
<td>Exon 13</td>
<td>Fram e-shift</td>
<td>Infantile</td>
<td>Scottish/Irish</td>
<td>44</td>
</tr>
<tr>
<td>G1511→A</td>
<td>Exon 13</td>
<td>Arg504→His</td>
<td>Juvenile</td>
<td>Italian</td>
<td>45</td>
</tr>
<tr>
<td>Δ7.6 Kb</td>
<td>5' end</td>
<td>no mRNA</td>
<td>Infantile</td>
<td>Assyrian</td>
<td>44</td>
</tr>
</tbody>
</table>

Carrier testing has been available to Ashkenazi Jews for about 20 years. Recently, several reports have appeared that have investigated the distribution of mutant alleles among Jews and have compared the fidelity of enzyme and DNA testing.\textsuperscript{52-54} Taken together, over 400 enzymatically defined carriers were examined for the three common alleles. While the data cannot formally be combined because the criteria for testing and defining carriers were slightly different among the three reports, all approximate the combined distribution of insertion mutation, 81\%, splice mutation, 16\%, and adult onset mutation, 3\%. In addition, from 5-18\% of carriers did not have one of the known mutations. These latter individuals likely include those with yet to be identified mutations, those with carrier-level enzyme results due to unknown biological factors and those that were low-normal because the cut-off used to designated carriers is biased to include a higher proportion of normal individuals.

In addition to the adult mutation, the 4 bp insertion has also been seen in non-Jewish populations. Indeed, it appears to be the most prevalent mutation in non-Jews. It was identified in 8/33 enzymatically determined carriers in one of the studies\textsuperscript{53} and 4/20 obligate carriers in another.\textsuperscript{52} To date, sixteen mutations have been identified in all populations (Table 1). Many of these have been seen in single families only, while others may prove to be associated with specific ethnic groups even though the overall carrier frequency in such groups might be low.

With such large numbers of alleles being identified in Tay-Sachs and other diseases, it will be important to identify cost-effective strategies for screening carriers for mutations. For example, will it be appropriate to develop screening tests that will detect all known mutations, or will it be acceptable to screen for only those mutations considered to be prevalent in the population in question? In Tay-Sachs carrier testing, the availability of a good enzymatic test is allowing us to temporarily set this issue aside. However, it is being brought into sharp focus...
with the discovery in cystic fibrosis of a very common single allele and an extremely large number of very rare alleles. Further, the ability to more accurately predict the clinical phenotype in Tay-Sachs disease (infantile, juvenile, adult, clinically normal) using mutation identification will ultimately demand the use of DNA testing.

The combination of classical biochemistry and genetic analysis with the new technologies of molecular biology has allowed us to define in exquisite detail the biosynthesis, processing, and structure of hexosaminidase. The molecular basis of mutation in Tay-Sachs disease is now yielding to the DNA technologies. The important challenge still before us is to understand how mutation of this enzyme initiates the cascade of events that leads to profound neurodegenerative disease.

REFERENCES


