The arylsulphatase A gene and molecular genetics of metachromatic leucodystrophy

Maria Luiza Barth, Anthony Fensom, Ann Harris

The lysosomal storage disorders share one common characteristic: the accumulation of a particular substrate inside the lysosome. Each disorder in this group is caused either by the deficiency of a lysosomal enzyme responsible for one step of the degradation pathway of a substrate, lack of a transporter involved in the movement of a small molecule across the lysosomal membrane, or absence of a low molecular weight activator protein required for in vivo hydrolysis of a sphingolipid. There are now more than 30 lysosomal storage diseases where the biochemical defect is well known and fully characterised. In this group, metachromatic leucodystrophy (MLD) is an autosomal recessive disorder that has been extensively studied over the past few decades. The overall incidence of MLD is estimated to be 1:40 000 and the enzymatic defect associated with this disease was discovered in the 1960s. More recently, the molecular biology of MLD has been elucidated, adding to the well established biochemical knowledge. Systematic mutation analyses have been carried out in an attempt to establish genotype–phenotype relationships. A clear understanding of the molecular basis of this disease will be helpful in the genetic counselling of affected families.

Historical aspects

In 1910, Alzheimer described metachromatic staining in the nervous system of an adult patient with diffuse sclerosis. The storage of metachromatic material not only in brain but in other tissues in a similar patient was described by Witte in 1921. A detailed study was reported in 1925 by Scholz which described clinicopathological findings in a family with progressive leucodystrophy. However, metachromatic glycolipid inclusions were not observed owing to the staining procedure used at the time. Thirty years later, using a different stain, Peiffer was able to find metachromatic granules in frozen sections of the original patient of Scholz. The storage of sulphatides in tissues from patients with MLD was discovered by two independent groups in 1958 and the build up of this substrate was proven to be responsible for the metachromasia observed in affected patients. In 1963, Austin et al proved that MLD was associated with the deficiency of a lysosomal enzyme, arylsulphatase A (ASA). This enzyme catalyses the first step in the degradation pathway of cerebrosides such as sphingosyl ceramide sulphate. In 1989, Stein et al reported the cloning of the ASA cDNA and, one year later, the genomic structure of the ASA gene was described by Kreising et al.

Clinical forms of MLD

There are three main clinical forms of MLD based on the age of onset of the disease. The late infantile form is the most common type of MLD. Patients with this form of the disease develop clinical features between the ages of 15 months and 2 years and the onset is characterised by loss of acquired motor skills. Hypotonias and ataxias are present in most patients and mental deterioration accompanies disease progression until the stage when patients are no longer able to establish any contact with their surroundings. This form of MLD has a short course and death generally occurs one to seven years after onset.

In the juvenile form of MLD the onset of symptoms usually occurs between the ages of 4 and 12 years and is characterised by motor dysfunction in younger patients or by mental illness with features of psychosis, dementia, and emotional disorders in older children. Walking and speaking difficulties are observed a few years later. Progressive neurological problems occur although the course of the disease is less severe than in the late infantile forms. The majority of juvenile MLD cases are fatal, but in a few cases the course of the disease can last for more than 20 years.

The onset of symptoms of patients with the adult form of MLD can vary within a wide age range, from 19 to 46 years. In general, it can be observed at any age beyond puberty. The progress of the disease is slow and five to 10 year survival is common. The initial symptoms are mostly changes in behaviour, which makes early diagnosis difficult because it is often attributed to psychiatric problems. A number of patients show symptoms of schizophrenia, anxiety, and general emotional instability. During the final stages of the disease, patients become severely affected, lose speech, and become immobile and incontinent.

Allelic and non-allelic MLD variants

In addition to the three clinical forms of MLD described above, which result from deficiencies
in ASA activity caused by mutations in the ASA gene, three associated conditions are known, one without and two with clinical consequences.

The first is allelic and is characterised by low ASA activity in healthy persons. This condition is termed ASA pseudodeficiency. It was first described in healthy relatives of MLD patients who were found to have low ASA activities similar to affected patients. Persons who present this phenomenon do not store metachromatic material and have normal excretion of urinary sulphatide. A pseudodeficient allele at the ASA locus was delineated by Schaap et al. in 1981 and was defined by Gieselmann et al. in 1989 after the cloning of the ASA cDNA. The allele is characterised by two mutations in the ASA gene: A→G transitions at base 1049 and base 1620. The first causes an asparagine to serine substitution (N350S), which leads to the loss of an N-glycosylation site, and the second changes the first polyadenylation signal downstream from the stop codon. The N350S mutation may occur alone and is not responsible for reduced ASA activity. The pseudodeficiency allele has a high frequency in general populations, which is estimated to be between 7 and 15%, and it may coexist in families carrying an MLD associated mutation. Therefore, the differential diagnosis has great importance for genetic counselling. One non-allelic variant of MLD is multiple sulphatase deficiency (MSD). In this autosomal recessive disease at least seven different sulphatases are deficient. These enzymes include ASA, arylsulphatase B, steroid sulphatase, and other sulphatases involved in the degradation of mucopolysaccharides. Patients show less advanced development in the presymptomatic period than that of children with late infantile MLD, and they lose the ability to sit, stand, and speak during the second year of life. Mucopolysaccharidosis-like features can be observed and death usually occurs before the end of the first decade of life, although there is also a form with onset in the neonatal period and early death. It has been suggested that the mutation responsible for MSD is in an enzyme in the endoplasmic reticulum which performs post-translational modification of sulphatase precursors. This hypothesis is based on the fact that ASA can be induced in cells of patients with MSD when certain culture conditions are used; therefore, it is likely that the structural gene for this sulphatase is intact. The gene responsible for MSD has not yet been defined.

The other non-allelic condition is sphingolipid activator protein B (SAP-B) deficient MLD. In this case, the ASA gene is normal and the mutation is located in the gene that codes for SAP-B, which makes the cerebroside sulphate accessible to the enzyme. SAP-B is one of four activator proteins called SAP-A, -B, -C, and -D, which are derived from a common precursor, prosaposin, each one acting as an activator of hydrolysis of different sphingolipids. The correct function of ASA depends on the integrity of SAP-B. Patients who are affected by this form of the disease show signs of juvenile MLD including sulphatiduria, have a normal ASA enzyme, but abnormal ASA activity owing to the lack of SAP-B that is required for enzyme activity in vivo. The prosaposin cDNA as well as the prosaposin gene have been cloned and sequenced.

**Enzyme defect and diagnostic tests**

ASA has been purified from a variety of sources including human liver, placenta, and urine. The enzyme has a high content of proline, aspartic acid, and glutamic acid. ASA forms a dimer at pH 4.5, but above pH 6.5 it exists as a monomer with a molecular weight of approximately 100 kDa. After purification, the enzyme consists of two non-identical subunits of different sizes depending on the source. Its active site contains two or more arginine residues.

The main laboratory test to diagnose MLD is the measurement of ASA activity. However, other biochemical or molecular tests have to be performed owing to the fact that low ASA activity is not necessarily indicative of MLD. The enzyme assay uses p-nitrocatechol sulphate as an artificial substrate instead of the natural cerebrosidesulphate, under carefully controlled conditions designed to minimise interference from others arylsulphatases, most notably arylsulphatase B. The most convenient way to perform the assay was devised by Lee-Vaupel and Conzelmann in 1987 and is carried out with long incubations at 0°C. Under these conditions, arylsulphatase B shows very little activity, which allows a good discrimination of MLD patients. This assay is usually used with leucocytes or cultured fibroblasts for postnatal studies or amniotic cells or chorionic villi for prenatal diagnosis. Urinary sulphatide excretion and measurement of sulphatide degradation by cultured fibroblasts are additional useful postnatal tests.

**Structure of the arylsulphatase A gene**

The ASA gene is located on chromosome 22, distal to q13, and the full length human ASA cDNA was cloned and sequenced in 1987. The predicted amino acid sequence comprises 507 residues, which includes a putative signal peptide of 18 residues that is removed after translocation in the endoplasmic reticulum. The amino acid sequence contains three potential N-glycosylation sites, two in exon 3 and one in exon 6.

The ASA cDNA hybridises to three mRNA species of 4.8, 3.7, and 2.1 kb on Northern blot analysis of total RNA from human fibroblasts. These species arise owing to the use of different polyadenylation signals. The first two species each account for about 50% of the total mRNA while the 2.1 kb species accounts for about 40%. After purification of poly A + RNA, the majority of the 3.7 and 4.8 kb mRNA species are lost. The 2.1 kb species is responsible for 90% of poly A + mRNA and the two larger species for about 5% each.

The structure of the ASA gene was determined by Kreysing et al. and includes about...
3.2 kb of genomic DNA divided into eight exons. The promoter region is characteristic of a "housekeeping" gene.

Mutation analysis
After the determination of the structure of the ASA gene, Southern blot analysis of DNA from MLD patients was carried out but neither deletions nor rearrangements could be detected. This indicates that in the majority of cases small deletions, insertions, or point mutations are likely to cause the molecular defect associated with the disease. Since then, several groups have analysed affected patients and characterised MLD associated mutations in the ASA gene. Most of these mutations have been proven to be disease causing by expression of mutant protein. A summary of mutations reported to date are listed in the table. Two of these mutations, a single base change (G→A) which destroys the splice donor site at the start of intron 2 (459 + 1G→A) and a single base change (C→T) causing a proline to leucine substitution (P426L) in exon 8; together account for about 50% of mutations in the MLD populations studied to date in northern Europe. The former is associated with more severe forms of the disease while the latter is associated with later onset forms of the disease. The remaining 50% of MLD associated mutations seem to be heterogeneous. Amino acid substitutions are the most frequent group of mutations responsible for the molecular defect associated with MLD; however, two small deletions and splice site mutations involving the splice donor site of introns 2, 4, and 7 have been reported. About half of the mutations reported so far are located in exons 2 and 3, which indicates that they are clustered in the 5' region of the gene. The location of this cluster of mutations at the 5' end of the ASA gene is not surprising since sulfatases share a higher degree of homology in this region, suggesting functional importance. The continuous characterisation of mutations in MLD patients will help to elucidate the role of different domains in the ASA protein.

Population screening and prenatal diagnosis
Screening for known mutations can be performed by several molecular methods. The detection of the two common mutations can be carried out by hybridisation with allele specific oligonucleotides (ASO). When the mutation alters a restriction site, the detection can be performed by digestion with appropriate restriction endonucleases. If the mutation itself does not alter a restriction site, the introduction of a mismatched PCR primer in the amplification reaction can generate a product which will enable the use of a restriction enzyme to detect a mutant allele.

Before the isolation of the ASA gene, prenatal diagnosis was based only on the measurement of ASA activity. Now, molecular methods can be used as additional tests, which enhances the specificity of the diagnosis. However, previous characterisation of mutations responsible for the genetic defects in MLD families is required. The detection can be made by the same methods mentioned previously or even by direct sequencing of the region of interest.

Genotype-phenotype relationship
Availability of molecular information on a patient may allow some predictions of the genotype-phenotype relationship. In the case of the two common mutations associated with MLD it has been observed that the 459 + 1G→A mutation is associated with the late infantile form of the disease, while the P426L mutation is associated with the late onset forms of MLD. However, the P426L mutation has been described in a late infantile case who has the same genotype as a juvenile patient. Hence, it is likely that genotype-phenotype predictions will not be entirely straightforward.


Mutations associated with MLD

<table>
<thead>
<tr>
<th>Name</th>
<th>cDNA nucleotide change</th>
<th>Effect on coding sequence</th>
<th>Exon</th>
<th>Patient's clinical form</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>R84Q</td>
<td>G→A at 251</td>
<td>Arg→Gln at 84</td>
<td>2</td>
<td>Adult</td>
<td>35</td>
</tr>
<tr>
<td>S96F</td>
<td>C→T at 287</td>
<td>Ser→Phe at 96</td>
<td>2</td>
<td>Late infantile</td>
<td>20</td>
</tr>
<tr>
<td>G99D</td>
<td>G→A at 296</td>
<td>Gly→Asp at 99</td>
<td>2</td>
<td>Adult</td>
<td>36</td>
</tr>
<tr>
<td>298delC</td>
<td>Deletion of C at 298</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G122S</td>
<td>G→A at 364</td>
<td></td>
<td>2</td>
<td>Adult</td>
<td>38</td>
</tr>
<tr>
<td>459 + 1G→A</td>
<td>G→A at 459 + 1</td>
<td>Destroys splice donor site of exon 2</td>
<td>2</td>
<td>Late infantile, juvenile, and adult</td>
<td>39</td>
</tr>
<tr>
<td>1179S</td>
<td>T→G at 536</td>
<td>Ile→Ser at 179</td>
<td>3</td>
<td>Juvenile</td>
<td>40</td>
</tr>
<tr>
<td>A212V</td>
<td>C→T at 635</td>
<td>Ala→Val at 212</td>
<td>3</td>
<td>Adult</td>
<td>22</td>
</tr>
<tr>
<td>A224V</td>
<td>C→T at 671</td>
<td>Ala→Val at 224</td>
<td>3</td>
<td>Adult</td>
<td>22</td>
</tr>
<tr>
<td>G245R</td>
<td>G→A at 731</td>
<td>Gly→Arg at 245</td>
<td>4</td>
<td>Late infantile</td>
<td>41</td>
</tr>
<tr>
<td>T274M</td>
<td>C→T at 821</td>
<td>Thr→Met at 274</td>
<td>4</td>
<td>Late infantile</td>
<td>42</td>
</tr>
<tr>
<td>S295Y</td>
<td>C→T at 884</td>
<td>Ser→Tyr at 293</td>
<td>5</td>
<td>Late infantile</td>
<td>22</td>
</tr>
<tr>
<td>G309S</td>
<td>G→A at 925</td>
<td>Gly→Ser at 309</td>
<td>5</td>
<td>Late infantile</td>
<td>37</td>
</tr>
<tr>
<td>E382K</td>
<td>G→A at 1145</td>
<td>Glu→Lys at 387</td>
<td>7</td>
<td>Juvenile</td>
<td>22</td>
</tr>
<tr>
<td>1205 + 1G→A</td>
<td>G→A at 1205 + 1</td>
<td>Destroys splice donor site of exon 7</td>
<td>7</td>
<td>Juvenile</td>
<td>40</td>
</tr>
<tr>
<td>P426L</td>
<td>C→T at 1277</td>
<td>Pro→Leu at 426</td>
<td>8</td>
<td>Juvenile and adult</td>
<td>39</td>
</tr>
<tr>
<td>1401del11</td>
<td>Deletion of 11 bp from 1401 to 1411</td>
<td>Frameshift</td>
<td>8</td>
<td>Late infantile</td>
<td>43</td>
</tr>
<tr>
<td>Q486X</td>
<td>C→T at 1458</td>
<td>Stop</td>
<td>8</td>
<td>Late infantile</td>
<td>44</td>
</tr>
</tbody>
</table>

* As defined by Stein et al. 16


