The Enzymatic Degradation of Heparan Sulfate
Laura S. Griffin*, Tracey M. Gloster*

*Biomedical Sciences Research Complex, North Haugh, University of St Andrews, St Andrews, KY16 9ST, UK.

Abstract: Glycosaminoglycans (GAGs) such as heparan sulfate (HS) interact with a number of factors in the extracellular matrix (ECM) and as a consequence play a key role in the metabolic processes occurring within the cell. The dynamic synthesis and degradation of HS (and all GAGs) are necessary for ensuring that optimal chains are present for these functions. The degradation of HS begins at the cell surface and finishes in the lysosome, after which components can be recycled. Deficiencies or mutations in the lysosomal enzymes that process GAGs result in rare Mucopolysaccharidoses disorders (MPSs). There are few treatments available for these genetically inherited diseases and those that are available often do not treat the neurological symptoms of the disease. In this review, we discuss the enzymes involved in the degradation of HS and their related diseases, with emphasis on those located in the lysosome.

Keywords: Glycosaminoglycan, heparan sulfate, degradation, lysosome, Mucopolysaccharidoses, enzymes.

INTRODUCTION

Carbohydrates known as glycosaminoglycans (GAGs) are major components of the extracellular matrix (ECM). Each tissue contains a unique, dynamic and heterogeneous ECM that infers tissue specificity via interactions with signalling molecules. Despite the physiological importance of GAGs, their synthesis is not directly encoded by the genome. There has hence been a great deal of interest in this field, not just to pinpoint the distinct functions of each GAG, but also to establish how and why particular chains are synthesised and degraded. GAGs in the ECM are typically found linked to a protein, forming a proteoglycan (PG), and is therefore a result of a protein undergoing post-translational modification. These PGs can possess a single GAG chain or many of different types. Hylanuronan is the only GAG that is not synthesised while covalently attached to a core protein. GAGs are composed of alternating disaccharides containing an amino sugar and an uronic acid or galactose (Gal) residue. With the exception of hyaluronan, GAG chains are often modified which impacts the interactions capable with other molecules. Many enzymes participate in the synthesis and modification of GAGs, often together in protein complexes. This observation has led to the proposed existence of a “GAGosome”, comprising a complex of the enzymes required to synthesise a particular GAG [1]. The exact composition of such a complex or complexes remains unknown.

Heparan sulfate (HS) is a linear GAG produced in the Golgi of all mammalian cells and is initially synthesised as alternating N-acetylgulucosamine (GlcNAc) and glucuronic acid (GlcA) repeating disaccharides. Heparin is also composed of alternating GlcNAc and GlcA residues but is only produced by mast cells. HS is known to influence pathogen entry, inflammation, and cell growth and migration, lending roles in disease and tissue development [2]. HS and the proteins which interact with and/or synthesise and degrade HS are therefore emerging targets in disease treatments. The ability to isolate or synthesise specific chains of HS or heparin allows the examination of the interactions between HS and the proteins involved in these biological processes. This has resulted in the development of HS and heparin libraries for probing these interactions [3]. Fully understanding the enzymes responsible for synthesising and degrading HS and heparin is therefore not just vital for designing therapeutics for their associated diseases, but also in the design and creation of larger optimized HS libraries.

The enzymes responsible for degrading HS and their associated diseases will be discussed here. In addition, there will be a brief introduction into the enzymes responsible for the synthesis and modification of HS. For recent detailed reviews on the biological synthesis and modification of HS (and other GAGs), please see Ref. 4 and 5.

HS SYNTHESIS

The GAG-protein linkage typical of all sulfated GAGs (HS, chondroitin sulphate (CS) and dermatan sulphate (DS)) is synthesised first, which involves attachment of the sugars comprising a conserved tetrasaccharide to a specific serine residue in the protein core. This serine is part of the S[G/A][G/A] conserved signal sequence [6,7], where at least two acidic residues either precede the serine residue or surround it [8]. The initial tetrasaccharide is synthesised sequentially by a xylosyltransferase (XylT), galactosyltransferase I (GaIT-I), galactosyltransferase II (GaIT-II) and glucuronyltransferase I (GlcAT-I), forming glucuronic acid(β1-3)-galactose(β1-3)-galactose(β1-4)-xylose(β1-O)-Ser. The donor sugars are all obtained from uridine diphosphate (UDP) linked substrates which are

*Address correspondence to this author at the Biomedical Sciences Research Complex, North Haugh, University of St Andrews, St Andrews, KY16 9ST, UK; Tel/Fax: +44-1334-467245, +44-1334-462595; E-mail: tmg@st-andrews.ac.uk
synthesised in the cytosol and subsequently transported to the Golgi [9], except in the case of UDP-xylose (UDP-Xyl) which is converted from UDP-GlcA by UDP-GlcA decarboxylase in the ER and Golgi [10]. The conserved tetrasaccharide in the GAG-protein linkage can be modified by phosphorylation of the Xyl (by FAM20B) [11] and de-phosphorylation of the Xyl (by 2-phosphoxylase phosphatase (XYLP)) [12]. The tetrasaccharide of CS or DS can also be modified by 6-O sulfation of Gal (by chondroitin 6-sulfotransferase I (C6ST-I)) [13] and 4-O sulfation of Gal [14]. These modifications to the tetrasaccharide in the GAG-protein linkage are often important for the regulation of the type and length of GAG chain synthesised.

Once the tetrasaccharide linker has been synthesised, GlcNAc is added to the terminus by N-acetylgalactosaminyltransferase I (GlcNACT-I), committing the GAG to being HS or heparin [15,16]. If N-acetylgalactosamine (GalNAc) is added to the terminus by N-acetylgalactosaminyltransferase I (GalNACT-I), the GAG is committed to being CS or DS [17]. It has been proposed that the activities of these two transferases can actually be attributed to one enzyme with two functional domains, EXTL2 [18]. Extension of the HS polymer with repeating disaccharides of GlcNAc and GlcA (Figure 1) is conducted by the so-called HS polymerase, which exhibits both N-acetylgalactosaminyltransferase II (GlcNACT-II) and glucuronyltransferase II (GlcAT-II) activities. HS polymerase is a heterodimer encoded by the EXTI and EXT2 genes [19]. There are also three exostosin-like genes, EXTL1-3, which all have roles in HS polymerisation.

The HS chain is heavily modified by a number of isoenzymes (depicted in Figure 2), yielding somewhat unique chains. GlcNAc of HS can be N-deacetylated and N-sulfated by four different GlcNAc N-deacetylase/N-sulfotransferases (NDSTs), each of which have high sequence similarity [20,21]. If GlcNAc residues have been deacetylated and sulfated, neighbouring D-GlcA residues can act as substrates for the glucuronyl C5-epimerase (GLCE) and be epimerised to L-iduronic acid (1-IdoA) [22]. HS can be O-sulfated by HS 2-O-sulfotransferase (2-OST), HS 3-O-sulfotransferase (3-OST) and HS 6-O-sulfotransferase (6-OST), all of which utilise 3'-phosphohadenosine 5'-phosphosulfate PAPS as a sulfur donor. 2-OST sulfates IdoA and GlcA residues, with a preference towards IdoA [23]. In fact, the occurrence of 2-O sulfated GlcA in GAGs is rare. Seven human isoforms of 3-OST sulfate glucosamine (GlcN), N-sulfogluconosamine (GlcNS) or N-sulfogluconosamine-6-sulfate (GlcNS6S) residues of HS and heparin, and each exhibit different substrate specificities and expression patterns in tissues [24,25]. Three isoforms of 6-OST sulfate GlcNAc and GlcNS residues of HS [26,27]. 6-O sulfation on the reducing side of IdoA or GlcA prevents 2-OST binding, implying that 2-O sulfation must occur before 6-O sulfation [28].

HS can also be remodelled extracellularly by the two human endosulfatases (Sulf1, Sulf2). These sulfatases remove the 6-O sulfate group from GlcNS6S residues, particularly in regions containing the disaccharide IdoA2S-GlcNS6S [29,30]. Overexpression of a quail orthologue (QSulf1) does not affect the level of the HS proteoglycan (HSPG) Glypican1 on the cell surface, inferring a role for these Sulfs in remodelling HS but not degradation [30]. Subsequently, de-sulfation of HS by the Sulfs disrupts interactions between HS and signalling molecules [31].

**HS DEGRADATION**

Due to the number of biological processes HSPGs mediate, turnover is a tightly regulated and necessary process. HSPG degradation is conducted by a number of exo- and endo-glycosidases and sulfatases, both extracellularly and intracellularly. This degradation begins at the cell surface and finishes in the lysosome, after which components can be recycled. The mechanism whereby HSPGs are internalised by the cell is unknown but they are trafficked through the endocytic pathway thereafter.

![Figure 1. Extension of the HS polymer is conducted by HS polymerase. The repeating HS disaccharide unit is initially composed of (4)GlcA(β1-4)GlcNAc(α1-3). Sites which can be modified by HS modifying enzymes have been illustrated as R or R'. GlcA can also be epimerised to IdoA, which has not been depicted here.](image1)

![Figure 2. As the HS chain is extended, it is heavily modified by a number of isoenzymes. GlcNAc can be N-deacetylated and N-sulfated by NDSTs, GlcA can be epimerised by GLCE, IdoA and GlcA can be sulfated by 2-OST, GlcNAc and GlcNS can be sulfated by the 6-OSTs and GlcN, GlcNS or GlcNS6S can be sulfated by the 3-OSTs.](image2)
HEPARANASES

Heparanase-1 (Hep1) is an extra- and intra-cellular endo-β-glucuronidase which in general cleaves the β1-4 GlcA linkage in GlcN6S-GlcA-GlcNS regions of HS, although this sequence alone is not sufficient and additional sulfation on this or the neighbouring sequence is essential [32]. Hep1 cleaves preferentially in highly sulfated regions of HS and the minimum oligosaccharide length it requires is a trisaccharide. The expression of Hep1 is tightly regulated. The Hep1 promoter is a “housekeeping” gene [33] with three Sp1 binding sites and four ETS-relevant elements (EREs), which are vital for basal activity. In fact, two additional EREs were discovered in the Hep1 promoter of a metastatic breast cancer cell line, suggesting that upregulation of Hep1 (and therefore ECM remodeling) is a factor in metastasis [34]. Hep1 expression levels are highest in the placenta, spleen, lymph node, peripheral blood leukocytes, bone marrow and foetal liver [35]. The high expression levels in lymphoid tissues suggests Hep1 typically acts at sites of inflammation or injury, by degrading HS and attracting leukocytes to these sites [36]. Overexpression of Hep1 has been observed in the tumours of various tissues and laboratory metastatic cell lines, and mice inoculated with Hep1 overexpressing cells developed lung and liver metastases [35,37,38].

The human Hep1 cDNA was first reported independently by three groups, all of whom purified Hep1 from human tissues and used derived sequences to screen EST databases [35,37,39]. Identification of the Hep1 cDNA was confirmed when cell lysates and/or media from overexpressing insect, COS-7 or NIH/3T3 cells were able to degrade HS substrates at a low pH optimum. The cDNA encodes a 65 kDa precursor containing a secretion signal sequence and six predicted N-glycosylation sites. The 65 kDa precursor is proteolytically cleaved to 8 kDa and 50 kDa forms which heterodimerise and produce active enzyme [40]. The linker region between the 8 kDa and 50 kDa forms has been predicted to block the active site of Hep1, rendering the precursor inactive [41]. This was confirmed in the X-ray crystal structure of mature Hep1 [42]. The protease responsible for this processing is thought to be a cathepsin-L-like protease [43]. Site-directed mutagenesis studies of the precursor cleavage sites suggested a non-specific protease acted, and a hydrophobic residue was required at one of the sites. Cathepsin inhibitors prevented processing of the Hep1 precursor and cathepsin-L was able to process the recombinant Hep1 precursor to the active form.

Hep1 appears to localise both in the lysosome and extracellularly. Immunofluorescence and GFP fusion studies indicated a localisation of Hep1 to acidic compartments within the cell [44]. These studies also demonstrated localisation to the Golgi, although it was suggested that transport to the Golgi was necessary for the proteolytic processing of Hep1 to the active enzyme. Hep1 can be endocytosed by primary human fibroblasts via an actin cytoskeleton pathway [45]. After endocytosis Hep1 doesn’t translocate to the lysosome, but instead resides in the cytoplasm and endosomes. This internalisation of Hep1 is mediated by HSGLPs [46]. It has been suggested that extracellular Hep1 functions more in binding HS for endocytosis, whereas intracellular Hep1 is the catalytically active form, as indicated by localisation studies and optimal activity at low pH [44,47].

The existence of a Hep1 related protein, Heparanase-2 (Hep2), was initially discovered by searching a then proprietary EST database with the Hep1 protein sequence [48]. Hep2 is alternatively spliced yielding three transcripts, which encode proteins of varying lengths. Hep1 and Hep2 share 35% sequence identity but high expression levels of each are detected in different tissues. Hep2 expression is highest in the brain, mammary gland, prostate, small intestine, testis and uterus, and significant expression is observed in the tumours of tissues which normally have low expression [48]. Hep2 doesn’t appear to have enzymatic activity towards HS but has been shown to inhibit Hep1 through its strong affinity to HS and Hep1 [49]. Hep2 is upregulated in some tumours, indicating a possible role in inhibiting Hep1-related cancers in this instance [48,49].

Connective tissue activating peptide-III (CTAP-III), a processing product of platelet basic protein (PBP), is an 8-10 kDa protein with optimum heparanase activity at pH 6.8 [50]. Inflammatory sites characteristically have a low pH, lending a role for CTAP-III in wound healing. The expression level in neutrophils of an 80 kDa CTAP-III isoform and Hep1 is comparable, but the activity of Hep1 is much higher [51]. It is not clear why neutrophils have multiple heparanase-enzyme active enzymes and little is known about CTAP-III.

LYSOSOMAL DEGRADATION

All of the enzymes which degrade HS in the lysosome are exo-acting, meaning they must act sequentially to fully break down the oligosaccharide chains. An example of the sequential breakdown of a HS chain by lysosomal enzymes is depicted in Figure 3. These enzymes contain phosphorylated mannos (Man) residues for targeting to the lysosome via the mannose-6-phosphate receptor (MPR). Genetically inherited deficiencies or mutations in the lysosomal enzymes that process GAGs results in the lysosomal storage disorders called Mucopolysaccharidoses (MPSs), where accumulation of incomplete degradation products in the lysosome is observed. Mutations located in the active site of these enzymes, or those which cause the enzyme to misfold (meaning it is targeted for proteosomal degradation rather than trafficked to the lysosome), can impair the enzyme function. The number or location of mutations in these lysosomal enzymes directly correlates with the severity of disease phenotype, with pathologies including respiratory problems, skeletal deformities and/or central nervous system (CNS) abnormalities [52]. The MPSs are therefore multisystem diseases which require input from a variety of different medical disciplines for diagnosis and treatment. The thirteen types of MPS are inherited and each one is related to a specific enzyme (Table 1). Nine of these enzymes are involved in the degradation of HS, either exclusively or in addition to other GAGs.

There are currently a number of treatments available for some types of MPS, including enzyme replacement therapy (ERT) and haematopoietic stem cell transplantation (HSCT). The idea of molecular chaperones has been explored, where small molecules bind to the enzyme and promote the correct folding of mutants and/or aid delivery across the blood brain barriers.
Table 1. The lysosomal enzymes involved in the degradation of GAG and their associated MPS disorders. Reported incidence rates are from the Registry for Mucopolysaccharide and Related Diseases. The MPS V and MPS VIII nomenclature is no longer in use. MPS V was originally assigned to Scheie syndrome before it was realised that Hurler and Scheie are caused by mutations in the same enzyme. MPS VIII originated from a misdiagnosis and therefore misassignment to a new disease type. KS, keratan sulfate.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>GAG</th>
<th>Disease</th>
<th>Incidence</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-L-iduronidase</td>
<td>HS and DS</td>
<td>MPS I (Hurler, Hurler-Scheie or Scheie syndrome)</td>
<td>1 in 100,000</td>
<td>ERT and HSCT</td>
</tr>
<tr>
<td>Iduronate-2-sulfatase</td>
<td>HS and DS</td>
<td>MPS II (Hunter syndrome)</td>
<td>1 in 100,000</td>
<td>ERT</td>
</tr>
<tr>
<td>Sulfamidase</td>
<td>HS</td>
<td>MPS IIIA (Sanfilippo A syndrome)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-N-acetylglicosaminidase</td>
<td>HS</td>
<td>MPS IIIB (Sanfilippo B syndrome)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acetyl-CoA:α-glucosamide N-acetyltransferase</td>
<td>HS</td>
<td>MPS IIIC (Sanfilippo C syndrome)</td>
<td>1 in 85,000 (for all MPS III combined)</td>
<td>No treatment available</td>
</tr>
<tr>
<td>N-acetylglucosamine-6-sulfatase</td>
<td>HS</td>
<td>MPS IIID (Sanfilippo D syndrome)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-sulfoglicosamine-3-sulfatase</td>
<td>HS</td>
<td>MPS IIIE (Sanfilippo E syndrome)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-acetylgalactosamine-6-sulfatase</td>
<td>KS and CS</td>
<td>MPS IVA (Morquio A syndrome)</td>
<td>Unknown</td>
<td>ERT</td>
</tr>
<tr>
<td>β-galactosamine</td>
<td>KS</td>
<td>MPS IVB (Morquio B syndrome)</td>
<td>Unknown</td>
<td>No treatment available</td>
</tr>
<tr>
<td>N-acetylgalactosamine-4-sulfatase</td>
<td>CS and DS</td>
<td>MPS VI (Maroteaux-Lamy Disease)</td>
<td>Unknown</td>
<td>ERT and HSCT</td>
</tr>
<tr>
<td>β-D-glucuronidase</td>
<td>HS, CS and DS</td>
<td>MPS VII (Sly syndrome)</td>
<td>&gt;1 in 250,000</td>
<td>HSCT</td>
</tr>
<tr>
<td>Glucuronate-2-sulfatase</td>
<td>HS, CS and DS</td>
<td>Unknown</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Hyaluronidase</td>
<td>Hyaluronan</td>
<td>MPS IX (Natowicz syndrome)</td>
<td>Unknown</td>
<td>No treatment available</td>
</tr>
</tbody>
</table>

As previously mentioned, mutations which cause misfolding of lysosomal enzymes often also prevent localisation to the lysosome. Schueler et al. noted that only 11-15% residual activity of glucosylceramide β-glucosidase (mutations of which cause Gaucher disease) in the lysosome is sufficient to prevent glucosylceramide accumulation [54]. Therefore, even if the folding of a small amount of a particular lysosomal GAG degrading enzyme is aided by a molecular chaperone and hence targeted to the lysosome, accumulation of the GAG can be prevented. An alternative treatment currently being explored is substrate reduction therapy (SRT). This approach uses small molecules that inhibit the enzymes involved early in the synthesis of the GAG of interest, thereby reducing the level of GAG requiring degradation. Gene therapy is also being investigated whereby the gene encoding the wild type enzyme is incorporated into the patient’s genome. Gene therapy methods may potentially allow permanent restoration of active enzyme expression, which would be an improvement over HSCT and ERT methods that require regular transfusions.

**GLYCOSIDASES**

Glycosidases (or glycoside hydrolyses) are a large family of enzymes responsible for cleavage of the glycosidic bond.
linking a sugar to another sugar, protein or lipid. Hydrolysis can occur with either inversion or retention of stereochemistry at the anomeric position, and typically involves two carboxylate-containing residues in the enzyme active site. Glycosidases are classified into 135 families based on sequence similarity (which is catalogued in the Carbohydrate Active enZyme (CAZy) database [55]) and display a variety of structural folds. There are three lysosomal glycosidases which degrade HS oligosaccharides to monosaccharides: α-L-iduronidase (IDUA), β-d-glucuronidase (GUSB) and α-N-acetylgalcosaminidase (NAGLU).

**IDUA**

IDUA (CAZy family GH39) cleaves the α-L-idooA glycosidic bond in HS and DS and has been purified from a number of human tissues and recombinant cell lines [56,57]. The cDNA of IDUA encodes a 653 amino acid polypeptide chain containing a 26 amino acid signal sequence and six potential N-glycosylation sites [58]. IDUA isolated from liver displayed a low pH optimum, an increase in activity upon the addition of bovine serum albumin (BSA) and a preference for C6 sulfation on the neighbouring residue at the reducing end [59]. The activity of recombinant IDUA also increased upon addition of the reducing agent dithioerythritol (DTE) and the non-ionic detergent Triton X-100 [57]. It was thought that the addition of Triton X-100 reduced aggregation of IDUA in the activity assays.

The crystal structure of human IDUA was reported independently by two groups [60,61]. Both groups observed a Man residue of the N-glycan on amino acid Asn372 interacting with the IDUA active site residues and, through de-glycosylation studies and inhibitor complexes, realised that this interaction aids substrate binding and therefore activity.

A deficiency or mutations in IDUA has been linked to MPS I, also known as Hurler, Hurler-Scheie or Scheie syndrome [62]. Hurler, Hurler-Scheie and Scheie syndromes refer to the more severe, intermediate and mild forms of the disease respectively. Recombinant IDUA has been shown to be endocytosed by fibroblasts from MPS I patients via the MPR pathway and restore degradation of HS and DS [57]. MPS I is currently treated with HSCT or ERT (Laronidase; Aldurazyme®) or a combination thereof, with a combination being the recommended course of treatment [63,64]. It has also been shown that there is scope for gene therapy for treating this disorder [65].

**GUSB**

GUSB (CAZy family GH2) cleaves the β-d-glucuronic acid glycosidic bond in HS, CS and DS and has been purified from a number of human tissues and recombinant cell lines [66,67]. GUSB localises both in the ER and lysosomes in a tissue-dependent manner. The level of murine GUSB in the ER of different tissues was shown to be dependent on the level of the ER membrane protein egasyn [68]. If GUSB is in excess of egasyn, it is transported to the lysosome. Egasyn binds to the C-terminal end of precursor GUSB, a region which is not present in lysosomal localised GUSB [69]. GUSB contains four N-glycosylation sites which are required for tetramerisation [70]. The X-ray crystal structure of tetrameric GUSB has been solved, which displays four separate active sites [71]. The active site of each monomer lies within a cavity formed between two monomers; molecular modelling indicates that each active site can independently bind substrates at the same time.

A deficiency or mutations in GUSB has been linked to MPS VII, also known as Sly syndrome, a rare but fatal disease [72]. MPS VII is currently treated with HSCT, and the first example of ERT for MPS VII improved the symptoms of a severely affected patient [73]. It has also been shown that there is scope for gene therapy for treating this disorder [74,75].

**NAGLU**

NAGLU (CAZy family GH89) cleaves the α-GlcNAc glycosidic bond in HS and has been purified from a number of human tissues and urine [76–80]. Northern blot analysis revealed expression of NAGLU resulted in the production of a single species across all tissues tested [79]. The structure of the Clostridium perfringens NAGLU homolog (with 30% amino acid sequence identity to human NAGLU) has been solved and used to determine a model structure for human NAGLU [81], although an experimentally determined structure is not yet solved. A deficiency or mutations in NAGLU has been linked to MPS IIIB (also known as Sanfilippo B syndrome) [82,83]. There are currently no treatments available for MPS IIIB, although a number of methods are being explored. The uptake of recombinant NAGLU has been observed in NAGLU-deficient cell lines and mice [84]. Uptake was observed in tissues expressing MPR, such as Kupffer and endothelial cells of the liver and macrophages of the spleen. Not surprisingly, uptake was not seen in hepatocytes, lymphocytes and epithelial cells of the kidney, which are all deficient in MPR. To avoid lack of cellular uptake of NAGLU due to the absence of recognition of mannose-6-phosphate, an insulin-like growth factor II (IGFII) fusion was designed [85,86]. MPRs bind IGFII at an alternative binding site allowing uptake in a variety of tissues. The NAGLU-IGFII fusion was injected directly into mice brain (thus bypassing the BBB) and endocytosis was observed in neurons and liver cells. Gene therapy has also been explored as a treatment for MPS IIIB, where an adeno-associated virus (AAV) vector system was used to deliver the NAGLU gene to the CNS and peripheral nervous system (PNS) of mice [87] and primates [88].

**SULFATASES**

The sequences of sulfatases are highly conserved and arise from a common ancestral gene, which is most likely the same ancestral gene that encodes phosphatases as they share a common phosphatase-like structural fold [89–91]. The overall structure and active site residues of sulfatases are highly conserved in all organisms [92]. The N-terminus is typically composed of a ten stranded β-sheet flanked by α-helices and contains the active site. The C-terminus of sulfatases has a conserved structure but not sequence; the structure is typically composed of a four stranded β-sheet flanked by α-helices. Although these evolutionary related sulfatases fundamentally catalyse the same reaction, they
Griffin et al.

Figure 3. An example of how the sequential degradation from the non-reducing end of the tetrasaccharide IdoA2S(α1-4)GlcNS3S(α1-4)GlcA2S(β1-4)GlcNAc6S would occur in the lysosome by HS degrading enzymes. The arrows illustrate which moiety of the sugar chain is cleaved by the corresponding enzyme, with exception of HGSNAT which catalyses the acetylation of GlcN to GlcNAc.

IDS

IDS cleaves the C2 sulfate group from iduronate-2-sulfate (Ido2S) at the non-reducing end of HS and DS. The 2.3 kb full length IDS gene encodes a 550 amino acid precursor polypeptide containing a predicted 25 amino acid signal sequence and eight potential N-glycosylation sites [97]. Cleavage of the first eight amino acids after the signal sequence cleavage site is also thought to occur, which removes one of these predicted N-glycosylation sites. To characterise IDS, it has previously been purified from human urine, plasma and various tissues, and from the overexpression of recombinant protein in mammalian cell lines [98–105]. Purification from tissues or urine typically yielded multiple IDS isoforms. Two of these isoforms from liver were studied further [104]; both were formed from 42-45 kDa and 14 kDa polypeptides and it was proposed that together they produced the active enzyme. Activity of liver IDS was dependent on the identity of the carbohydrate on the reducing side of Ido2S and was more active when this residue was sulfated at the C6 position.

Hunter syndrome (also known as MPS II) is an X-linked recessive inherited disorder, the symptoms of which were first recorded in 1917 [106] before the disease was linked to IDS [98,107]. The severity of the disease is dependent on the type of genetic mutation the patients carry; these mutations include single or multiple base deletions or substitutions, gene rearrangements and complete deletions of the IDS gene [108]. The X-ray crystal structure of IDS has recently been solved [109] providing a rationalization to the pathologies of known mutations. Rearrangements and deletions of the IDS gene can also impact neighbouring genes, giving rise to Hunter syndrome patients with additional disease phenotypes. An example of this is a deletion spanning to and including the FMR2 locus [110,111]. Mutations in the FMR2 gene are associated with Fragile X E (FRAXE) syndrome and together with Hunter syndrome may cause more severe mental retardation. The IDS gene is located within 80 kb of an IDS-related or ‘pseudogene’ (IDS2) [112]. This IDS2 gene has sequence similarity to IDS and recombination of the two genes is present in some Hunter syndrome patients [113,114].

Hunter syndrome is currently treated with ERT where ELAPRASE® (idursulfase, Shire Human Genetic Therapies Inc., Lexington, USA) or Hunterase® (idursulfase beta, each exhibit substrate specificity. All eukaryotic sulfatases (and some prokaryotic) contain a conserved C[X]PSR motif and a partially conserved AALLTGR auxiliary motif, where [X] is either Thr, Cys, Ser or Ala [93]. The conserved cysteine residue is oxidised by the formylglycine generating enzyme (FGE) in the ER, forming the formylglycine (FGly) residue required for enzymatic activity. Mutations in the FGE encoding SUMF1 gene cause the fatal lysosomal storage disorder Multiple Sulfatase Deficiency (MSD), where the activity of all sulfatases is diminished due to lack of FGly conversion [94–96]. There are five sulfatases which are involved in the degradation of HS in the lysosome: iduronate-2-sulfatase (IDS), glucuronate-2-sulfatase (GDS), N-sulfoglucosamine-3-sulfatase (GN3S), N-acetylgalactosamine-6-sulfatase (GN6S) and heparan N-sulfatase (also known as sulfamidase).
Green Cross Crop., Yongin, Korea) is administered intravenously on a weekly basis. High doses of these drugs are potentially dangerous to patients and they do not appear to improve neurological impairments due to the inability of the drugs to cross the BBB [115,116]. Fusion of a monoclonal antibody raised against the human insulin receptor to recombinant IDS (HIRMAb-IDS) has been shown to improve the ability of IDS to cross the BBB in Rhesus monkeys [116,117]. Gene therapy has also been explored as a treatment for Hunter syndrome, where an AAV vector system was used to deliver the IDS gene to the CNS and liver of mice, with promising outcomes [118]. HSCT or bone marrow transplant (BMT) in Hunter syndrome patients has been relatively unsuccessful, with complications arising from the procedure and little improvement of symptoms observed in surviving patients [119].

GDS

GDS catalyses the removal of the C6 sulfate group from the β-D-GlcA residues of HS, CS and DS [120]. For characterisation, GDS has been purified from chick embryo chondrocytes, human fibroblast cell lines and human tissues [120–122]. Liver isolated GDS displayed a native molecular mass of 63 kDa by gel filtration and 47 kDa and 19.5 kDa forms by SDS-PAGE [121]. This indicated a possible proteolytic processing event where the mature protein is eventually formed of a heterodimer. GDS demonstrated a low pH optimum and inhibition by NaCl and EDTA, typical characteristics of lysosomal sulfatases. Until recently the existence of the GDS gene remained elusive but it has since been established that arylsulfatase K (ARSK) and GDS are synonymous [123]. The ARSK gene encodes a 536 amino acid protein of 68-70 kDa when recombinantly expressed [124,125]. A deficiency in this enzyme has not been reported in any patients to date, although a mild phenotype is expected due to the rarity of sulfated GlcA residues.

GN3S

GN3S catalyses the removal of the C3 sulfate group from the rare N-sulfoglucosamine-3-sulfate (GlcNS3S) residues of HS. This sulfatase was first isolated from human urine, where it was shown to specifically cleave the C3 sulfate group from glucosamine-3-sulfate, but only when the amino group was also sulfated, and at an optimum pH of 6.3 [126]. What had previously been annotated as arylsulfatase G (ARSG) has now been assigned the enzymatic role of GN3S after ARSG-knockout mice accumulated HS chains with a terminal N-sulfoglucosamine-3-sulfate residue [127]. The ARSG sequence encodes a 525 amino acid and 62 kDa protein with 87% homology to the murine orthologue [128]. Murine ARSG is present in most tissues except the heart and lungs, and a processed product of 34 kDa is also present in high quantities in the liver and brain [129]. Subcellular fractionation of mouse liver tissue suggested there was prelysosomal processing of ARSG. It has been proposed that the 63 kDa isoform of ARSG is tightly membrane associated in the ER/Golgi where it is first cleaved to 44 kDa and 18 kDa forms, and then the 44 kDa form is cleaved again to 34 kDa and 10 kDa forms, all in-transit to the lysosome [129]. It was suggested that the proteases cathepsin B and/or L were involved in this processing, but not exclusively, and additional proteases were also likely to be involved.

A deficiency in GN3S/ARSG has provisionally been assigned as MPS IIIE syndrome, although there have been no recorded patients suffering from MPS IIIE to date [127]. Early onset degeneration of the retina was observed in ARSG knockout mice [130] and mutations in the ARSG gene in American Staffordshire terrier (AST) dogs has been shown to cause ataxia and inherited late-onset neuronal ceroid lipofuscinosis (NCL) [131], indicating potential phenotypes of any yet unidentified MPS IIIE sufferers.

GN6S

GN6S catalyses the removal of the C6 sulfate group from glucosamine-6-sulfate (GlcN6S) and N-acetylglucosamine-6-sulfatate (GlcNAc6S) residues of HS and KS [132–134]. However, β-N-acetylgalactosaminidase A is able to remove GlcNAc6S from the non-reducing end of KS without the removal of the C6 sulfate by GN6S, thus bypassing the need for GN6S in KS degradation in vivo [135]. GN6S is also able to catalyse the removal of the C6 sulfate group from glucose-6-sulfate in vivo [133]. For characterisation, GN6S has previously been purified from human urine and liver, rat skin, bovine kidney and from overexpression of recombinant protein using mammalian cell lines [133,136–140]. GN6S isoforms isolated from liver demonstrated a substrate-dependent optimum activity at pH 3.9-5.7, enhanced activity with the addition of BSA and DTT, and inhibition by sulfate and phosphate ions and NaCl above 25 mM [141]. These liver isoforms were more active on GlcNAc6S than GlcN6S, and most active when GlcNS6S was at the non-reducing end of the trisaccharide GlcNS6S-IdoA2S-anhydromannitol-6-sulfate. It has been postulated that the precursor GN6S polypeptide is cleaved yielding 32 kDa and 48 kDa polypeptides which interact to form mature GN6S [142].

GN6S has been linked to Sanfilippo D syndrome, also known as MPS IID [143]. This rare disease has only 31 reported sufferers and 23 different GNS (encoding GN6S) mutations reported thus far [144], explaining why this enzyme has been so little studied and why there is currently no treatment available for this disease. However, a colony of severe MPS IID suffering goats has been established to study GN6S and design therapeutics [145]. Recombinant caprine GN6S expressed in CHO cells was able to restore the degradation of HS in MPS IID fibroblasts and the degradation of HS in the organs of MPS IID suffering goats [139,146,147].

Sulfamidase

Sulfamidase cleaves the sulfate from the amino group of GlcNS residues of HS and has been purified from a number of human tissues and recombinant cell lines [148–150]. Sulfamidase is an N-sulfatase which has structural similarities to O-sulfatases, but there is a clear difference in the vicinity of the active site to accommodate the different substrate [151]. Mutational studies of the five predicted N-glycosylation sites of sulfamidase indicated that the N-glycans on at least two of the sites (Asn41 and Asn151) are required for protein folding and/or enzyme activity [149]. A deficiency or mutations in sulfamidase has been linked to ceroid lipofuscinosis (NCL) [131], indicating potential phenotypes of any yet unidentified MPS IIIE sufferers.
Sanfilippo A syndrome, also known as MPS IIIA [152]. Sulfamidase is a functional homodimer and some of the known MPS IIIA causing mutations disrupt the dimer interface and/or substrate access to the active site [151]. MPS IIIA is the most common of the Sanfilippo syndromes and there are a number of treatment options currently being investigated. There is evidence that injecting recombinant sulfamidase into the cerebrospinal fluid of MPS IIIA suffering mice reduces neurological symptoms, although this required regular injections and a considerable humoral response was detected [153]. A sulfamidase-IGFII fusion has also been designed with uptake in the brain of Rhesus monkeys looking promising [154]. Gene therapy treatment has also been explored with MPS IIIA. An adenovirus vector system was used to deliver both SGSH (encoding sulfamidase) and SUMF1 (encoding FGE) genes to the CNS [155] and a reduction in GAG storage in the lysosome was observed in the brain. A humoral response is often seen with human adenovirus gene transfer, so a canine adenovirus transfer has also been tested with promising outcomes [156]. Similarly, intravenous lentiviral-mediated SGSH gene delivery increased sulfamidase levels in many tissues in mice [157–159]. HSCT is an unpopular treatment for MPS IIIA and is not thought to improve neurological symptoms [160].

ACETYL-COA:α-GLUCOSAMIDE N-ACETYLTRANSFERASE

Acetyl-CoA:α-glucosamidase N-acetyltransferase (HGSNAT) catalyses the acetylation of GlcN residues in HS, after desulfation by sulfamidase. The product of this acetylation, GlcNAc, can subsequently be cleaved from the HS chain by NAGLU. HGSNAT is the only HS lysosomal enzyme that isn’t directly involved in degradation. For characterisation, HGSNAT has been partially purified from human tissues and recombinant cell lines [161–164]. The HGSNAT gene product is not homologous to any other proteins, and is predicted to contain 11 transmembrane domains and five N-glycosylation sites [165]. The N-terminus of the protein is predicted to be lysosomal and the C-terminus cytosolic. HGSNAT is ubiquitously expressed in all tissues with higher expression levels in the heart, skeletal muscle, liver, leukocytes, placenta and kidneys [165,166]. HGSNAT transfers acetyl groups from the cytoplasm to the inside of the lysosome and onto GlcN using acetyl-coenzyme A (acetyl-CoA) as the donor [167]. Acetylation of HGSNAT has a neutral pH optimum and transfer to GlcN has a low pH optimum.

The mechanism of action of HGSNAT is under debate and there are two proposed models. One model suggests a ‘ping-pong’ double displacement mechanism where an active site His residue on the cytosolic side of HGSNAT is acetylated and then a conformational change takes place to present the acetylated-His on the lysosomal side for acetyl transfer to GlcN [168]. In support of this model it was proposed that HGSNAT is proteolytically processed to two domains (29 kDa and 48 kDa) in the lysosome, which are held together with disulfide bonds and eventually form a ~440 kDa active oligomer [164]. It was suggested that this proteolytic processing occurs within a loop containing the critical His making it more flexible and susceptible to the proposed conformational change. Another model proposed a ‘random-order ternary-complex’ mechanism which doesn’t involve an acetylated-His intermediate [169]. A group who support this mechanism agrees that HGSNAT is cleaved in the lysosome to 27 kDa and 44 kDa products, but concluded that this processing event is not necessary for HGSNAT activity [163]. Their kinetic studies suggested a random sequential mechanism and they were unable to detect an acetylated intermediate by radiolabelling, which would have confirmed the ping-pong mechanism.

A deficiency of or mutations in HGSNAT has been linked to Sanfilippo C syndrome, also known as MPS IIIC [170]. It has been suggested that there are two classes of MPS IIIC patients, those that possess HGSNAT which cannot be acetylated and those that possess HGSNAT which can be acetylated but cannot transfer the acetyl group to GlcN [171]. Many of the missense HGSNAT mutations result in misfolded and/or mis-glycosylated protein, making these MPS IIIC patients good candidates for molecular chaperone therapy [172]. It has also been reported recently that MPS IIIC suffering mice not only present with lysosomal storage accumulation but with mitochondrial structural abnormalities in neurons [173].

CONCLUSION

Knowledge of the enzymes which synthesise and degrade HS, and the ability to overexpress and purify them, has not only been critical for understanding the related genetic diseases, but also for probing the structure of HS and heparin itself. In recent years the research in this field has mostly focussed on those enzymes that synthesise and modify HS, and molecules which interact with HS to manipulate biological processes. However, the lysosomal enzymes which degrade HS (and additional GAGs) are fundamental to the recycling of HS and therefore the life of the cell. Deficiencies in these enzymes may be related to rare MPS disorders but these are multisystem, fatal diseases. There are few approved treatment methods for the majority of these diseases, and those that are available often produce detrimental side effects and are unable to improve neurological symptoms. A better understanding of the mode of action of these lysosomal enzymes will allow for the design of more optimal therapeutics such as molecular chaperones.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

ACKNOWLEDGEMENTS

The authors would like to acknowledge financial support from the Wellcome Trust (Research Career Development Fellowship to TMG) and the BBSRC (studentship to LSG).

REFERENCES

The Enzymatic Degradation of Heparan Sulfate

Rosen, S. D.; Morimoto, W.; Liu, J.; Pedersen, L. C.


Sheldrick, G. M.; Gärtner, J.; Krätzner, R.; Steinfeld, R. Acta


Pharm. 2014, 11 (8), 2928–2934.

Sutherland, L.; Auricchio, A.; Ballabio, A.; Hopwood, J. J. Hum.
Mol. Genet. 2007, 16 (22), 2693–2702.

Genet. Metab. 2010, 100 (2), 168–175.

418.

717–728.


[160] Lau, A. A.; Hannouche, H.; Rozaklis, T.; Hassiotis, S.; Hopwood,

(2), 99–106.

[162] Freeman, C.; Clements, P. R.; Hopwood, J. J. Biochem. Int. 1983,
6 (5), 663–671.

One 2011, 6 (9), e24951.


738–744.

[166] Hrebícek, M.; Mrázová, L.; Seyrantepe, V.; Durand, S.; Roslin, N.
Poupotová, H.; Sikora, J.; Urinovská, J.; Stranecký, V.; Zeman, J.;
Lepage, P.; Roquis, D.; Verner, A.; Beesley, C. E.;
Maire, I.; Poorthuis, B. J. H. M.; van de Kamp, J.; van Diggelen,
O. P.; Wevers, R. A.; Hudson, T. J.; Fujiwara, T. M.; Majewski,

11299.

10132.

308 (Pt 1), 327–333.

A. 1978, 75 (10), 5185–5189.


4 (10), e7434.

Grigoryeva, L.; Choi, Y.; Langford-Smith, A.; Wilkinson, F. L.;
Moreau, A.; Svebozavá, E.; Hájková, Z.; Tesařová, M.;
Hansiková, H.; Bigger, B. W.; Hrebícek, M.; Pshezhetsky, A. V.