Glycoconjugate Journal 17, 597–605, 2000

'Heparin' – from anticoagulant drug into the new biology

Ulf Lindahl

Department of Medical Biochemistry and Microbiology, University of Uppsala, The Biomedical Center, Box 582, SE-751 23 Uppsala, Sweden

This overview attempts to cover, from a personal viewpoint, the development of the 'heparin' field during the last four decades. In particular, it emphasizes the metamorphosis of heparan sulfate (HS), from a disturbing contaminant in heparin production to the present-day key player in cell and developmental biology. Our understanding of the structural properties of the polysaccharides has been greatly promoted by studies of their biosynthesis. We now have a fairly detailed view of the various enzymatic reactions, that convert the initial [4GlcAβ1-4GlcNAcα1-]$_n$ polymer into sulfated products with highly variable proportions of GlcA/IdoA and of N-acetyl, N-sulfate and O-sulfate substituents. It is also recognized that the variously substituted domains of the polysaccharide serve to interact, in more or less specific fashion, with a multitude of proteins, and that these interactions are essential to the biological functions of the proteins. Molecular genetics has unravelled the gene structures for almost all of the enzymes required to synthesize a heparin or HS chain, and has shown that several of these proteins exhibit genetic polymorphism. While differences in substrate specificity between enzyme isoforms may help to explain the structural variability of, in particular, HS chains, we still only partly understand the key features of heparin/HS biosynthesis and its regulation.

Keywords: **heparin, heparan sulfate, antithrombin, growth factors, iduronic acid, sulfotransferases**

My first experimental encounter with heparin, in fact with research on the whole, was in 1962 when I began my graduate training at the LaRabida Children's Hospital in Chicago, at the connective-tissue research center headed by Albert Dorfman. I have continued the study of 'heparin' ever since, and thus have been able to watch at close range its remarkable development from a poorly characterized antithrombotic drug to a most intriguing player in cell and developmental biology. In the present paper I will attempt to outline this development, with some bias toward studies in our own laboratory.

Dull as it may seem, a matter of nomenclature is unavoidable. *Heparin* ('the stuff kept in bottles and given to patients') may be defined as the glucosaminoglycan produced by connective-tissue type mast cells, as part of the serglycin proteoglycan; all related polysaccharides, generated by other cells, are *heparan sulfates*. Generally, the two species may also be distinguished by structural criteria alone, regardless of tissue or cell source, heparin being more extensively sulfated than heparan sulfate (HS) [1]. While objections may be raised to either of these definitions, they should be kept in mind by readers confronted with indiscriminate reference to 'heparins' or even 'heparans' in current literature. For simplicity I will use current nomenclature for all glycosaminoglycans (GAGs) throughout this account.

To whom correspondence should be addressed. Ulf Lindahl.
Tel.: +46-18-471 4196; Fax: +46-18-471 4209;
E-mail: ulf.lindahl@imbim.uu.se

Structural aspects

In the early 60's the proteoglycan (a term not yet in use at the time) concept was recognized in that chondroitin sulfate chains were known to be bound to serine residues in polypeptide core structures. Lennart Rodén and his coworkers isolated saccharide substituted polypeptide ('millipede') following digestion of the major cartilage proteoglycan with testicular hyaluronidase, and proceeded to generate small glycopeptides ('millipede legs') through proteolytic degradation of the core. Analysis of the product unexpectedly showed the presence of two neutral monosaccharides, one of which was identified as galactose, the other tentatively as xylose ('the pink spot' observed on paper chromatograms stained with aniline hydrogen phthalate). Since residual amino acid residues had been found also in some heparin preparations I was assigned the task, as part of my thesis project at LaRabida, to analyse this polysaccharide for the occurrence of neutral sugars. The final result of this work was the identification of the same GlcAβ1,

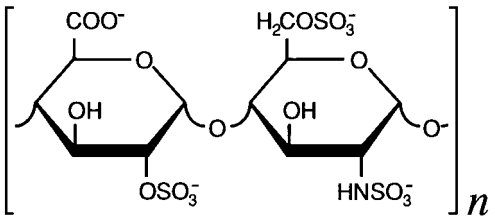

Figure 1. 'Heparin' in 1960.

3Galβ1,3Galβ1,4Xylβ-O-serine 'linkage region' attached to both heparin and chondroitin sulfate chains (reviewed in [2]). Rodén supervised the project, an outstanding privilege for me who was then his only (and first) graduate student. I still benefit from his mentorship.

After graduation, in 1966, I started to build up an independent research group in Uppsala, still with focus on heparin. This became possible due to the generous support of Torvard Laurent, then head of polysaccharide/polymer research at the Department of Medical and Physiological Chemistry, University of Uppsala. The current notion of heparin at the time was that of a polymer composed largely of repeating -4GlcA(2-OSO_3)α1-4GlcNSO_3(6-OSO_3)α1-disaccharide units (Figure 1). The finding by Cifonelli and Dorfman of L-iduronic acid (IdoA) in heparin [3] had been dismissed, by and large, as being due to contamination with dermatan sulfate (where IdoA was first detected, by Karl Meyer, the great pioneer of the GAG field). Nevertheless, our analysis, mainly based on paper chromatography of products obtained by deamination with nitrous acid and acid hydrolysis revealed IdoA as the major hexuronic acid component of heparin, and further showed that most of the hexuronyl 2-O-sulfate groups are bound to IdoA and not to GlcA units [4,5]. The occurrence of IdoA was verified by other groups using different techniques, including NMR [6] (see also [7]). Torsten Helting, then a graduate student in our group, found that GlcA residues could be added to the nonreducing termini of the appropriate -GlcNAc-GlcA- sequences, using a microsomal fraction from a heparin-producing mouse mastocytoma as enzyme source [8]. Unexpectedly, the GlcA units thus incorporated could be released by treatment with β-D-glucuronidase, suggesting that the GlcA residues in heparin might be linked in β-D-rather than α-D-configuration as previously assumed (mainly based on the optical properties of heparin solutions). This finding, along with NMR studies [9] (see also [7]), established the β-D-gluco and α-L-ido configurations for the hexuronic acid components in heparin, with inverted structures at C5 but identical disposition at C1 (Figure 2A). Analogous copolymeric arrangement of hexuronic acid units was found in dermatan sulfate by Lennart Rodén and Lars-Åke Fransson [10]. Notably, at this time (early 70's) rare ('unique') structural components of heparin, such as the GlcN 3-O-sulfate group (see below) were yet to be discovered.

Polysaccharides structurally related to but less sulfated than heparin were early recognized in heparin manufacture, as 'side fractions' that needed to be eliminated during the production process. While HS was defined as a distinct molecular entity already in 1948, by Erik Jorpes and Sven Gardell [11], the elucidation of its structure, more complex than that of heparin, has been a slow, painstaking process. The single most important contribution to this process, in my opinion, was the demonstration of the basic domain structure of the polysaccharide by Joseph A. ('Tony') Cifonelli, who applied nitrous acid to distinguish between N-sulfated and N-acetylated regions of the chain [12] (Figure 2A). The resultant deaminative cleavage was exclusively restricted to N-sulfated GlcN units, whereas N-acetylated sequences were left intact, such that the N-substitution pattern could be deduced from the amounts of variously sized oligosaccharide degradation products. Cifonelli, an outstanding scientist with a unique sense of humor, also worked at LaRabida, introduced me to the nitrous acid technique and thus enabled the demonstration of scarce N-acetylated GlcN residues also in heparin [4,13] (Figure 2B). We found, through serendipity, that the target specificity of the nitrous acid reaction could be manipulated through changes of pH conditions [14], and this observation was refined by Edward Conrad into procedures for selective attack of either N-sulfated or N-unsubstituted GlcN residues, with concomitant cleavage of the saccharide chain [15]. We further noted that N-acetyl groups in heparin could be released by hydrazinolysis, without any apparent loss of sulfate groups [16]. Conrad and coworkers optimized the conditions for N-deacetylation of GAGs, and showed that N-deacetylated HS could be largely converted into disaccharides by nitrous acid treatment at the appropriate pH [17]. The procedures for N-deacetylation and selective nitrous acid cleavage elaborated by Conrad et al. remain invaluable tools for compositional analysis of HS and for selective dissection of HS chains, to generate fragments derived from various types of saccharide domains (see [7]).

The detailed characterization of HS domain structure and organization is still underway. Most of the information available has been obtained in our laboratory, using the chemical methods described above in conjunction with NaB^3H_4 labeling of cleavage products, and in that of John Gallagher and coworkers, who employed enzymatic degradation of metabolically radiolabeled HS [7,18–20]. Some key features are incorporated in Figure 2A, that depicts a hypothetical HS sequence composed of contiguous N-acetylated and N-sulfated stretches (NA- and NS-domains, respectively) and also mixed structures (NA/NS-domains). The NA-domains contain essentially -GlcA-GlcNAc- disaccharide units, whereas the NS- and NA/NS-domains both contain IdoA and GlcA residues as well as GlcN 6-O-sulfate groups. Remarkably, 2-O-sulfation of the IdoA units appears to be restricted to the NS-domains. Comparison of the final polymeric structures generated through the biosynthetic processes illustrated in Figures 2A and 2B suggests that heparin, from a strictly structural angle, may be considered essentially an extended and highly O-sulfated NS-domain.

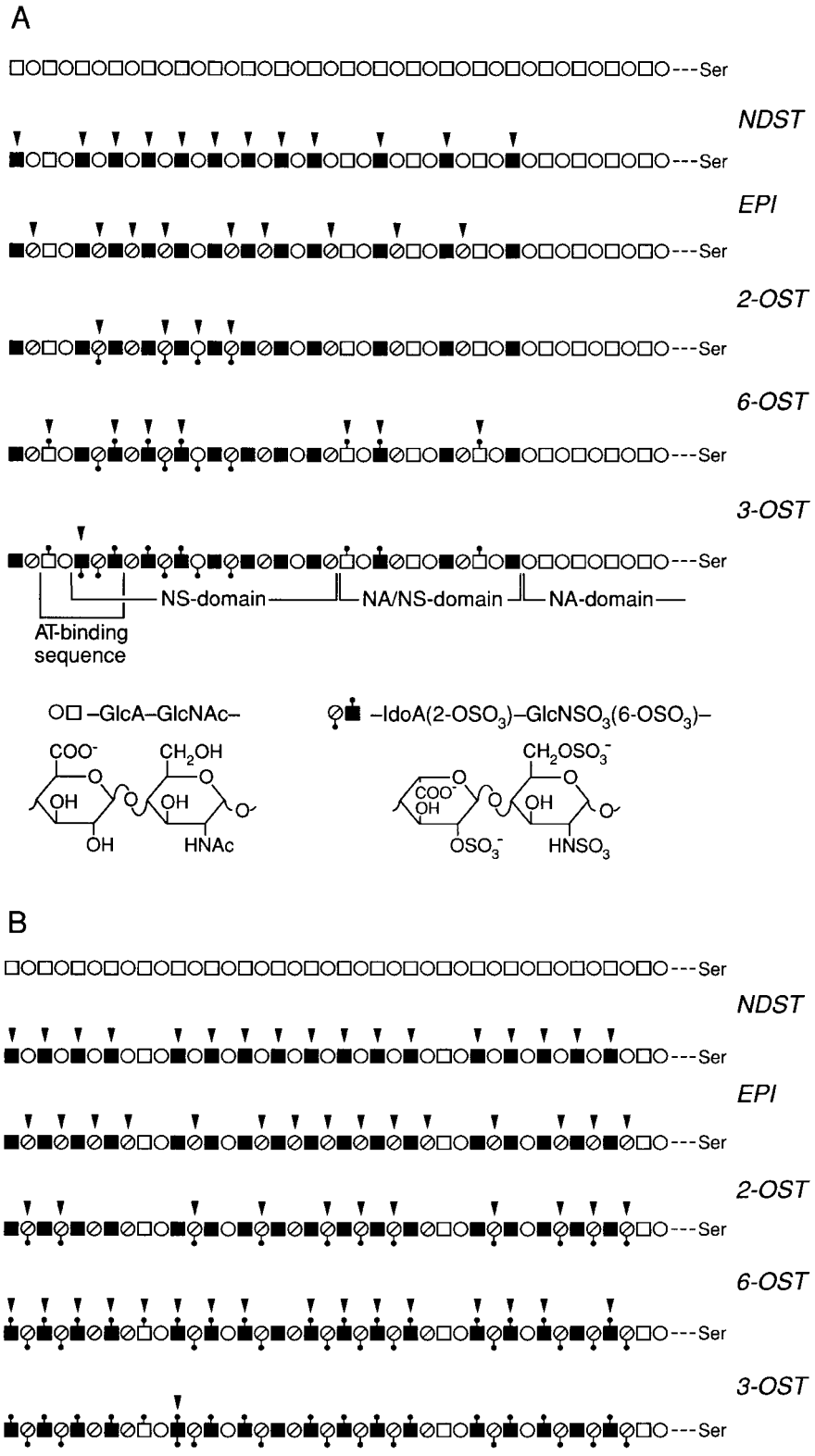

Figure 2. Schematic illustration of polymer-modification reactions in the biosynthesis of HS (A) and heparin (B). The disaccharide structures shown under the scheme in panel A define the various symbols used. The target residues for the various enzymes (NDST, GlcNAc N-deacetylase/N-sulfotransferase; EPI, GlcA C5-epimerase; 2-OST, hexuronyl 2-O-sulfotransferase; 6-OST, GlcNR 6-O-sulfotransferase; 3-OST, GlcNR 3-O-sulfotransferase) are indicated by *arrowheads*. Panel A emphasizes the generation of the typical domain structure of a HS chain. Note that both the HS and the heparin structure illustrated contain the AT-binding pentasaccharide sequence (indicated in panel A only).

Recent observations point to strictly regulated distribution of 2-O- and 6-O-sulfate groups in HS. We thus found that whereas 2-O-sulfation of human aortic HS remains essentially constant, the proportions of 6-O-sulfate groups increase steadily with increasing age [21]. HS preparations from different organs differ in composition, in apparently similar manner from one individual to another [18,22,23], and subject to distinct changes in certain disease conditions [23]. The proportions of 2-O- and 6-O-sulfate groups vary independently with the length of the NS-domains [24]. Attempts at sequence analysis, in Gallagher's group, suggested preferential 6-O-sulfation at distinct sites of metabolically radiolabeled NS-domains [20].

In addition to the potential sequence variability conferred by the major structural components recognized in heparin and HS, i.e. the three monosaccharide units (GlcN, GlcA, IdoA), the N-acetyl and N-sulfate groups, and the IdoA 2-O-sulfate and GlcN 6-O-sulfate groups, we also need to consider the 'rare' components that generally occur much less frequently. To date, such components include the GlcN 3-O-sulfate group (required for antithrombin binding; see below), the N-unsubstituted GlcN residue, and the 2-O-sulfated GlcA unit [7]. Maybe there are more; for instance, it is notable that no 3-O-sulfated hexuronic acid residue has yet been found in heparin or HS.

Given the multitude of building blocks so far identified in the polysaccharides, the potential for sequence variability is obviously enormous. To what extent is this variability exploited in the body, in regulated fashion? This key issue in 'heparin' research was highlighted by the finding of Jaap van den Born, that different anti-HS monoclonal antibodies recognized distinct cell types in rat kidney. While none of the cognate HS epitopes has yet been defined, van den Born identified, during a visit as guest scientist in our laboratory, N-unsubstituted GlcN as an essential component for recognition by one of these antibodies [25]. Similar distinct staining patterns were reported by Toin van Kuppefelt using phage-display antibodies [26]. Do all of these distinct epitopes depend on the occurrence of a 'rare' component, such as N-unsubstituted GlcN, or can they also be formed through unique combinations of the 'major' structural building blocks described above? The ultimate answer to this problem will require access to micro-scale HS sequencing techniques, that are currently being approached through different strategies (masspectrometry, exo-enzyme digestion) by several groups, including our own [27–29].

Our laboratory was also involved in some of the early characterization of heparin and HS proteoglycan structures. We proposed, albeit based on rather crude experiments (gel chromatography before and after alkali treatment, followed by colorimetric assay for hexuronic acid), that native HS may indeed occur in proteoglycan molecules composed of several polysaccharide chains linked to a common polypeptide core [30]. Magnus Höök and his graduate student, Lena Kjellén, subsequently found that HS proteoglycans may be anchored at the cell surface through intercalation of core proteins in the plasma membrane [31], a concept that was further developed through cloning of the various members of the syndecan family [32]. The large number of HS-proteoglycans known today includes also the glypicans [32], a group of phosphatidyl-inositol-anchored core proteins, and the extracellular species, perlecan and agrin [33]. A 'macromolecular heparin' isolated from proteolytically degraded rat skin by Alan Horner was shown by Clem Robinson, who spent a sabbatical in our lab, to be a proteoglycan susceptible to alkaline β-elimination, with serine and glycine as the only amino-acid residues present in significant amounts. We proposed, based on these findings, that the polysaccharide-substituted portion of the heparin proteoglycan consists of alternating serine and glycine units [34], and this proposal was verified by cloning of the serglycine cDNA from heparin-producing mast cells [35]. Notably, all of these heparin and HS proteoglycan species, as well as most (or all?) of the chondroitin and dermatan sulfate proteoglycans characterized to date, appear to share the same basic -GlcA-Gal-Gal-Xyl-O-Ser polysaccharide-protein linkage region that we initially identified in commercial heparin.

Function

Early studies of the functional properties of heparin were essentially restricted to two types of biological activities, the blood anticoagulant activity and the antilipaemic activity (associated with the 'clearing factor' released into the blood upon administration of heparin *in vivo*). These activities were both found to be due to the interaction of heparin with proteins, antithrombin and lipoprotein lipase, respectively. The former activity has been exploited since the 1930's in prevention and treatment of thromboembolic disease, for most of this time with poor understanding of the molecular mechanism behind the effect. Early clues to this mechanism were the identification by Ulrich Abildgaard of the major 'heparin cofactor' as antithrombin (AT) [36], and the elucidation, by Robert Rosenberg, of the catalytic mode of action of heparin in promoting the inhibition, by AT, of the serine proteinases of the coagulation cascade [37]. The first indication for the requirement of a specific heparin structure in the interaction with AT was the finding by three different groups, that of Rosenberg, our own, and Lars-Olov Andersson and his coworkers at KabiVitrum in Stockholm (70 km south of Uppsala!) that only a portion of the molecules in standard heparin preparations was capable of such interaction [38–40]. We next established that the AT-binding sequence was contained within an octasaccharide fraction, recovered after partial deaminative cleavage with nitrous acid. Careful dissection of such AT-binding octamer identified the pentasaccharide structure shown in Figure 3 as the actual binding sequence, and implicated a 'rare' component, the 3-O-sulfate group on the internal GlcN unit, as being essential for the interaction [41]. Other, more ubiquitous, components, such as the two

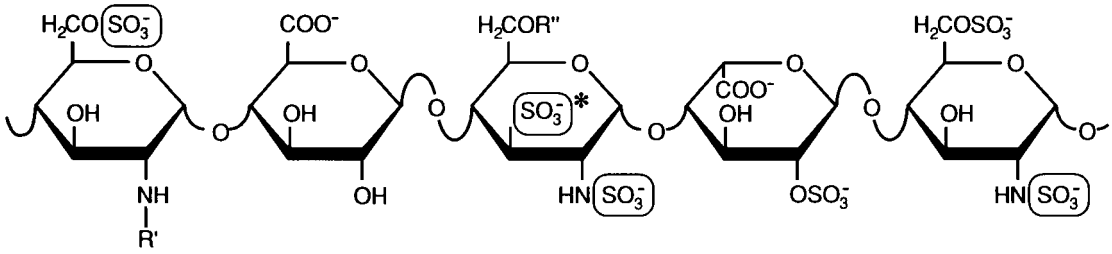

Figure 3. Structure of the pentasaccharide sequence in heparin and HS required to bind and activate AT. The circled sulfate groups are particularly important to the interaction. The 'rare' 3-O-sulfate group is indicated by an asterisk (*).

N-sulfate groups and the 6-O-sulfate group indicated in the figure, were also found to be important for the interaction [42,43]. These structure–function relationships were confirmed and extended through impressive chemical synthesis of the active pentasaccharide along with numerous derivaties, by Pierre Sinaÿ, Maurice Petitou and their collaborators [44,45]. The applicability of chemical saccharide synthesis as a tool in studies of structure–function relations for GAGs is likely to extend far beyond the heparin-AT interaction.

The elucidation of the AT-binding sequence in heparin initiated several new lines of research. In a series of elegant studies Rosenberg and coworkers showed that the AT-binding sequence, including the 'unique' 3-O-sulfate group, is not restricted to heparin, but is also expressed, albeit in highly selective fashion, in HS [46]. We as well as other groups found that the AT-binding pentasaccharide sequence alone was sufficient to promote the inactivation by AT of certain coagulation enzymes, such as Factor Xa, but was relatively inefficient with other factors, such as thrombin. These observations led to the proposal of the 'template' mechanism, by which certain coagulation enzymes are efficiently inactivated only if both the proteinase and AT are bound to the same polysaccharide chain (reviewed in [43,47]). Importantly, the new, detailed understanding of the action mechanism of heparin was prerequisite to the development of 'low-molecular weight heparin', a major advance in clinical routine prophylaxis against thromboembolic disease (see [7]). Still a number of functional problems relating to this particular structure remain to be resolved; for instance, why is it abundant in the heparin-like GAG of certain marine mussels, that have no blood to anticoagulate [48]?

It is now generally agreed that the biological function of HSs is, by and large, to interact with proteins. Is the structural diversity of HSs from different cells and tissues indeed specifically tailored to provide selective binding sites for different proteins? Notably, most of the recognized protein ligands for HS also bind heparin, in seeming contradiction of the notion of selectivity. Selectivity based on limited numbers of N- and O-sulfate groups in defined positions may, however, still be at hand, since most such combinations would actually be presented, albeit in concealed form, also by the fully sulfated heparin molecule [49]. Indeed, recent studies by our group and others indicate that many proteins bind NS type oligosaccharide sequences that contain limited numbers of 2-O- and 6-O-sulfate groups. Such mode of interaction has now been inferred for a variety of proteins, including lipoprotein lipase [50], platelet-derived growth factor A [51], basic fibroblast growth factor (FGF-2) [52–55], acidic fibroblast growth factor (FGF-1) [54,56], fibronectin [57], the *Herpes simplex* gC glycoprotein [58], and the angiogenesis inhibitor endostatin [59]. Functional implications are apparent for several of these interactions. Thus, binding of FGFs to cell-surface HS greatly promotes FGF-receptor activation (reviewed in [7]); binding of the viral glycoprotein to HS at the target cell surface is prerequisite to infectivity; and binding of endostatin to HS at the vascular endothelium prevents development of capillaries. While attempts have been made to define the minimal HS sequences for binding of some of these proteins, this area of research remains largely unexplored. However, compositional analysis alone demonstrated distinct structural requirements for binding of HS to FGF-2 (a single 2-O-sulfate but no 6-O-sulfate groups) vs. FGF-1 (both 2-O- and 6-O-sulfate residues) [53,56,60], and this difference has been associated with distinct cellular responses to the growth factors [61] and to selective interactions of the growth factors with vascular HS from individuals of different age [21,56]. While these isolated findings seem compelling, the specificity of HS-protein interactions in general is not readily assessed. In theory, a very large number of different protein-binding domains could be generated by regulated, site-specific incorporation of the 'common' sugar and sulfate residues – but to what extent does this actually happen? Given the current development of HS sequencing techniques, this issue will most likely be resolved within the next few years.

Another intriguing problem, but one that now appears to be settled, is the functional role of IdoA units. For all our interest in this constituent, its distribution, biosynthesis, constant location in the AT-binding region, we have been disturbingly unable to explain why IdoA is needed in the polysaccharide chain. Notably, it was well established that the IdoA-containing GAGs, dermatan sulfate and HS, are more prone to protein binding than chondroitin sulfate, that lacks IdoA but has similar overall charge density. The clue to the functional properties of IdoA was unravelled by Benito Casu and his colleagues, who proposed that the unique conformational

flexibility of IdoA serves to facilitate the approach of the anionic groups of the GAG to appropriate basic groups of a protein ligand (see [7]).

Demonstrating binding between HS species and proteins *in vitro* is not the same as implicating these interactions in biological functions. There has always been a nagging question – Is this important? The demonstration by, in particular, Alan Rapraeger in the early 90's, that HS-deficient cells were unable to respond to the mitogenic, and other, effects of FGF-2 [62] was therefore received with enthusiasm by the GAG-oriented community. Several groups, including our own, became involved in the growth factor area. We found that whereas the minimal ~5-mer saccharide domain sufficient for FGF-2 binding was unable to restore the mitogenic response to HS-deficient cells, a 12-mer was almost as efficient as a full-sized polysaccharide chain. Moreover, the structure of the additional piece appeared to differ from that needed to merely bind the FGF-2 molccule [63]. These findings, along with those of Gallagher and his coworkers [64,65], suggest that in order to promote the functional interaction between the growth factor and its receptor (ultimately leading to receptor dimerization), the HS chain needs to interact with both the receptor and the growth factor molecules. Alternative hypotheses by other groups favor dimerization of growth factor molecules along the HS chain. Results compatible with either view have been recently obtained through crystallography – a future key technique in the area? – of growth factors, their receptors and oligosaccharides [7]. An exciting development can be foreseen.

The final boost to shatter any doubts on the biological significance of HS-protein interactions came from unexpectcd direction. Developmental biologists have uncovered an array of genes, in flies, worms and mammals, with essential roles in embryonic development. It is only recently, when molecular genetics conquered also the proteoglycan field, that the functional roles of many of the corresponding proteins became apparent. Information from quite diverse fields of research thus has revealed that an appreciable proportion of developmentally important genes encode either core proteins of HS-proteoglycans, enzymes involved in the biosynthesis of HS chains, growth factors/cytokines that depend on HS for their proper presentation/function, or the receptors of such factors [66,67]. The interaction between the fields of 'heparin' research and developmental biology is likely to expand.

Biosynthesis

Having worked for some time with the structural aspects of heparin I became interested in the mechanism of its biosynthesis. These studies were initiated around 1970 and have developed ever since, with continually new questions relating to the increasingly complex structural and functional aspects. The foundation for these studies had been established through the pioneering efforts of Jeremiah Silbert, who showed that incubation of a microsomal fraction from the Dunn-Potter mouse mastocytoma with UDP-GlcA and UDP-GlcNAc yielded a nonsulfated polysaccharide, that was resistant to digestion with testicular hyaluronidase but susceptible to degradation by a 'heparinase' from *Flavobacterium heparinum*. In the presence of the sulfate donor, 3′-phosphoadenosine-5′-phosphosulfate (PAPS), a sulfated polysaccharide was formed, which was shown to contain N-sulfate groups [68,69]. Nils Ringertz, at the Karolinska Institute in Stockholm, who had previously used the Furth mastocytoma to demonstrate the incorporation of radiolabeled sulfate into a heparin-like polysaccharide, gave me a sample of this tumor that had been kept frozen for more than 5 years. The tumor turned out to remain viable upon inoculation into the appropriate inbred mice, and it has now been maintained in the lab for more than 30 years. Contrary to other mastocytomas, 'our' tumor maintained its ability to synthesize heparin, and thus has been invaluable as a source of cells, microsomal preparations, enzymes, and graduate thesis projects.

Most of the key experiments were conducted with a microsomal fraction from the tumor. In brief, it was demonstrated that the initial assembly of a nonsulfated structure, composed of alternating GlcA and GlcNAc units, is followed by a series of polymer modification reactions that occur in ordered sequence, regulated in large part by the substrate specificities of the enzymes involved (Figure 2) (reviewed in [7,49,70–72]. While the nonsulfated polysaccharide formed in the absence of PAPS is partially N-deacetylated, the free amino groups become rapidly sulfated upon addition of the nucleotide sulfate. The N-deacetylation and N-sulfation reactions are catalyzed by the same, bifunctional enzyme [73,74], and presumably occur in concerted fashion in the intact cell. Following the introduction of N-sulfate groups, the polysaccharide is recognized as a substrate by the GlcA C5-epimerase that converts GlcA to IdoA units at the polymer level, and also by O-sulfotransferases that catalyze sulfation at C2 of IdoA, C6 and C3 of GlcNR units, essentially in the order mentioned. The process is amazingly efficient, a heparin chain, M_r 60–100×10^3, being formed, as part of the serglycin proteoglycan, in less than 30 s. Various hypothetical models of the biosynthetic machinery have been advanced, and modified as more information has become available. Our current view features a unidirectional process, the nascent polysaccharide chain moving through a complex of membrane-bound Golgi enzymes while still growing through addition of alternating GlcA and GlcNAc units at the nonreducing end [49]. While the models may be speculative (as has been pointed out by some colleagues) I still feel that they are warranted, not only as a basis for future experimentation, but also as a convenient means of visualizing the current information relating to a complex process.

Studies of GAG biosynthesis, as many other fields, have been revolutionised through the introduction of molecular genetics techniques. All enzymes required to generate a heparin/HS chain (with some reservation for the linkage region to protein) have now been cloned and expressed, and

information regarding their molecular structures, substrate specificities, catalytic properties and expression patterns is rapidly accumulating. Some characteristics and current problems relating to the individual reactions will be briefly considered. The initial polymerization reaction, by addition of GlcA and GlcNAc units, was found to be catalyzed by a bifunctional enzyme. Molecular cloning showed that it belongs to the EXT family of 'tumor suppressor' proteins, defects of which lead to hereditary multiple exostoses and other tumor forms. Intriguingly, two isoforms, EXT1 and EXT2, need to be associated to generate a functional 'polymerase' [75]. Also the first polymer-modification reaction, the N-deacetylation/N-sulfation of GlcNAc units, is catalyzed by a bifunctional enzyme that occurs in several isoforms [76,77]. By contrast, both the GlcA C5-epimerase (partial clone published [78]; full-length expressed form now available with much higher activity (Li et al., in preparation)) and the IdoA 2-O-sulfotransferase [79] have so far been found as single forms only. The same enzyme thus would seem to catalyze the formation of IdoA units in both heparin and HS. Even more striking, the same 2-O-sulfotransferase acts on both GlcA and IdoA units [80]. Conversely, both the GlcN 6-O- [81] and 3-O-sulfotransferases [82] occur as multiple isoforms, that show subtle differences in substrate preference. Two recent, exciting findings demonstrate the functional importance of enzyme polymorphism in heparin/HS biosynthesis and biology. A subset of the 3-O-sulfotransferase was found to 3-O-sulfate rare N-unsubstituted GlcN units in HS to create specific binding sites for the *Herpes simplex* gD glycoprotein, essential for viral entry into target cells [83]. Further, targeted disruption of the N-deacetylase/N-sulfotransferase gene 1 in mice lead to major derangement of HS structures and a severe, lethal phenotype [84], whereas elimination of gene 2 selectively abolished heparin formation in mast cells but left the animals otherwise seemingly unaffected [85]. The functional significance of the enzyme isoforms presumably relates to a variety of factors, including substrate specificity, expression patterns, and interactions with other (enzyme or nonenzyme) Golgi proteins.

Perspectives

It has been exciting to watch, and participate in the amazing transformation of HS, from an annoying side fraction in heparin manufacture to its present-day stature in molecular, cell, and developmental biology. The future development will probably be equally dramatic, and much faster. On the technical side we will see further refinement of HS structural analysis, with regard to composition, sequence, conformation and modelling, further chemical and enzymatic saccharide synthesis, use of defined antibodies for HS epitope mapping, and improved methods to study HS-protein interactions. On the conceptual side, an array of questions remain to be solved. How many proteins bind selectively to distinct sites in HS chains? What is the functional importance of nonspecific interactions? Overall role of HS in embryonic development, growth factor action, inflammation, tumor biology – and, indeed, hemostasis? What is the blueprint for the assembly line that generates HS; does it come anywhere near the (last) proposed model? What other factors, in addition to substrate specificity for the various enzymes involved, control polymer modification? The answers to these questions, surely to be provided, will depend on technical progress, but also on influences from many other research areas.

Will it be possible to construct a biosynthetic machinery, composed of recombinant proteins assembled in an artificial membrane system, that can be instructed to produce a HS chain of predetermined structure, tailored for interaction with a defined protein ligand? When this happens, somewhere, it would be fun to still be around.

References

1 Gallagher JT, Walker A, Molecular distinctions between heparan sulphate and heparin. Analysis of sulphation patterns indicates that heparan sulphate and heparin are separate families of N-sulphated polysaccharides, *Biochem J* **230**, 665–74 (1985).

2 Lindahl U, Rodén L, In *Glycoproteins – their composition, structure and function*, edited by Gottschalk A, Vol. A (Elsevier, Amsterdam, 1972), pp. 491–517.

3 Cifonelli JA, Dorfman A, The uronic acid of heparin, *Biochem Biophys Res Commun* **7**, 41 (1962).

4 Lindahl U, Further characterization of the heparin-protein linkage region, *Biochim Biophys Acta* **130**, 368–82 (1966).

5 Lindahl U, Axelsson O, Identification of iduronic acid as the major sulfated uronic acid of heparin, *J Biol Chem* **246**, 74–82 (1970).

6 Perlin AS, Mazurek M, Jaques LB, Kavanagh LW, A proton magnetic resonance spectral study of heparin. L-iduronic acid residues in commercial heparins, *Carbohyd Res* **7**, 369–79 (1968).

7 Casu B, Lindahl U, Structure and biological interactions of heparin and heparan sulfate. *Adv Carbhohyd Chem Biochem* (2000) in press.

8 Helting T, Lindahl U, Occurrence and biosynthesis of β-glucuronidic linkages in heparin, *J Biol Chem* **246**, 5442–7 (1971).

9 Huckerby TN, Nieduszynski IA, Proton chemical-shift assignments in the n.m.r. spectra of heparan and heparin. *Carbohyd Res* **103**, 141–5 (1982).

10 Fransson L-Å, Rodén L, Structure of dermatan sulfate II. Characterization of products obtained by hyaluronidase digestion of dermatan sulfate, *J Biol Chem* **242**, 4170–5 (1967).

11 Jorpes JE, Gardell S, On heparin monosulfuric acid, *J Biol Chem* **176**, 267–75 (1948).

12 Cifonelli A, Reaction of heparitin sulfate with nitrous acid, *Carbohyd Res* **8**, 233–42 (1968).

13 Cifonelli JA, King J, Variable content of 2-acetoamido-2-deoxy-D-glucose and neutral sugars in mammalian heparins, *Carbohyd Res* **12**, 391–402 (1970).

14 Lindahl U, Bäckström G, Jansson L, Hallén A, Biosynthesis of heparin. Formation of sulfamino groups, *J Biol Chem* **248**, 7234–41 (1973).

15 Shively JE, Conrad HE, Formation of anhydrosugars in the chemical depolymerization of heparin, *Biochemistry* **15**, 3932–42 (1976).
16 Thunberg L, Bäckström G, Lindahl U, Further characterization of the antithrombin-binding sequence in heparin, *Carbohyd Res* **100**, 393–410 (1982).
17 Guo Y, Conrad HE, The disaccharide composition of heparins and heparan sulfates, *Anal Biochem* **176**, 96–104 (1989).
18 Maccarana M, Sakura Y, Tawada A, Yoshida K, Lindahl U, Domain structure of heparan sulfates from bovine organs, *J Biol Chem* **271**, 17804–10 (1996).
19 Gallagher JT, Structure–activity relationship of heparan sulphate, *Biochem Soc Trans* **25**, 1206–9 (1997).
20 Merry CL, Lyon M, Deakin JA, Hopwood JJ, Gallagher JT, Highly sensitive sequencing of the sulfated domains of heparan sulfate, *J Biol Chem* **274**, 18455–62 (1999).
21 Feyzi E, Saldeen T, Larsson E, Lindahl U, Salmivirta M, Age-dependent modulation of heparan sulfate structure and function, *J Biol Chem* **273**, 13395–8 (1998).
22 Lindahl B, Eriksson L, Lindahl U, Structure of heparan sulphate from human brain, with special regard to Alzheimer's disease, *Biochem J* **306**, 177–84 (1995).
23 Lindahl B, Lindahl U, Amyloid-specific heparan sulfate from human liver and spleen, *J Biol Chem* **272**, 26091–4 (1997).
24 Safaiyan F, Lindahl U, Salmivirta M, Structural diversity of N-sulfated heparan sulfate domains, *Biochemistry* **39**, 10823–30 (2000).
25 van den Born J, Gunnarsson K, Bakker MAH, Kjellén L, Kusche-Gullberg M, Maccarana M, Berden JHM, Lindahl U, Presence of N-unsubstituted glucosamine units in native heparan sulfate revealed by a monoclonal antibody, *J Biol Chem* **270**, 31303–9 (1995).
26 van Kuppevelt TH, Dennissen MA, van Venrooij WJ, Hoet RM, Veerkamp JH, Generation and application of type-specific anti-heparan sulfate antibodies using phage display technology. Further evidence for heparan sulfate heterogeneity in the kidney, *J Biol Chem* **273**, 12960–6 (1998).
27 Venkataraman G, Shriver Z, Raman R, Sasisekharan R, Sequencing complex polysaccharides, *Science* **286**, 537–42 (1999).
28 Turnbull JE, Hopwood JJ, Gallagher JT, A strategy for rapid sequencing of heparan sulfate and heparin saccharides, *Proc Natl Acad Sci USA* **96**, 2698–703 (1999).
29 Vivès RR, Pye DA, Salmivirta M, Hopwood JJ, Lindahl U, Gallagher JT, Sequence analysis of heparan sulphate and heparin oligosaccharides, *Biochem J* **339**, 767–73 (1999).
30 Jansson L, Lindahl U, Evidence for the existence of a multichain proteoglycan of heparan sulphate, *Biochem J* **117**, 699–702 (1970).
31 Kjellén L, Pettersson I, Höök M, Cell-surface heparan sulfate: an intercalated membrane proteoglycan, *Proc Natl Acad Sci USA* **78**, 5371–5 (1981).
32 Bernfield M, Götte M, Park PW, Reizes O, Fitzgerald ML, Lincecum J, Zako M, Functions of cell surface heparan sulfate proteoglycans, *Annu Rev Biochem* **68**, 729–77 (1999).
33 Iozzo RV, Matrix proteoglycans: from molecular design to cellular function, *Annu Rev Biochem* **67**, 609–52 (1998).
34 Robinson HC, Horner AA, Höök M, Ögren S, Lindahl U, A proteoglycan form of heparin and its degradation to single-chain molecules, *J Biol Chem* **253**, 6687–93 (1978).
35 Kjellén L, Pettersson I, Lillhager P, Steen M-L, Pettersson U, Lehtonen P, Karlsson T, Ruoslahti E, Hellman L, Primary structure of a mouse mastocytoma proteoglycan core protein, *Biochem J* **263**, 105–13 (1989).
36 Abildgaard U, Highly purified antithrombin III with heparin cofactor activity prepared by disc electrophoresis. *Scand J Clin Lab Invest* **21**, 89–91 (1968).
37 Damus PS, Hicks M, Rosenberg RD, Anticoagulant action of heparin, *Nature* **246**, 355–7 (1973).
38 Lam LH, Silbert JE, Rosenberg RD, The separation of active and inactive forms of heparin, *Biochem Biophys Res Commun* **69**, 570–7 (1976).
39 Höök M, Björk I, Hopwood J, Lindahl U, Anticoagulant activity of heparin: separation of high-activity and low-activity heparin species by affinity chromatography on immobilized antithrombin, *FEBS Lett* **66**, 90–3 (1976).
40 Andersson L-O, Barrowcliffe TW, Holmer E, Johnson EA, Sims GEC, Anticoagulant properties of heparin fractionated by affinity chromatography on matrix-bound antithrombin III and by gel filtration, *Thromb Res* **9**, 575–83 (1976).
41 Lindahl U, Bäckström G, Thunberg L, Leder IG, Evidence for a 3-O-sulfated D-glucosamine residue in the antithrombin-binding sequence of heparin, *Proc Natl Acad Sci USA* **77**, 6551–5 (1980).
42 Lindahl U, Thunberg L, Bäckström G, Riesenfeld J, Nordling K, Björk I, Extension and structural variability of the antithrombin-binding sequence in heparin, *J Biol Chem* **259**, 12368–76 (1984).
43 Bourin M-C, Lindahl U, Glycosaminoglycans and the regulation of blood coagulation, *Biochem J* **289**, 313–30 (1993).
44 Choay J, Petitou M, Lormeau JC, Sinaÿ P, Casu B, Gatti G, Structure–activity relationship in heparin: a synthetic pentasaccharide with high affinity for antithrombin III and eliciting high anti-factor X_a activity, *Biochem Biophys Res Commun* **116**, 492–9 (1983).
45 Petitou M, Herault JP, Bernat A, Driguez PA, Duchaussoy P, Lormeau JC, Herbert JM, Synthesis of thrombin-inhibiting heparin mimetics without side effects, *Nature* **398**, 417–22 (1999).
46 Rosenberg RD, Shworak NW, Liu J, Schwartz JJ, Zhang L, Heparan sulfate proteoglycans of the cardiovascular system. Specific structures emerge but how is synthesis regulated? *J Clin Invest* **100**, S67–75 (1997).
47 Bjork I, Olson ST, Antithrombin. A bloody important serpin, *Adv Exp Med Biol* **425**, 17–33 (1997).
48 Pejler G, Danielsson Å, Björk I, Lindahl U, Nader HB, Dietrich CP, Structure and antithrombin-binding properties of heparin isolated from the clams Anomalocardia brasiliana and Tivela mactroides, *J Biol Chem* **262**, 11413–21 (1987).
49 Salmivirta M, Lidholt K, Lindahl U, Heparan sulfate – a piece of information, *FASEB J* **10**, 1270–9 (1996).
50 Parthasarathy N, Goldberg IJ, Sivaram P, Mulloy B, Flory DM, Wagner WD, Oligosaccharide sequences of endothelial cell surface heparan sulfate proteoglycan with affinity for lipoprotein lipase, *J Biol Chem* **269**, 22391–6 (1994).
51 Feyzi E, Lustig F, Fager G, Spillmann D, Lindahl U, Salmivirta M, Characterization of heparin and heparan sulfate domains binding to the long splice variant of platelet-derived growth factor A chain, *J Biol Chem* **272**, 5518–24 (1997).
52 Habuchi H, Suzuki S, Saito T, Tamura T, Harada T, Yoshida K, Kimata K, Structure of a heparan sulphate oligosaccharide that binds to basic fibroblast growth factor, *Biochem J* **285**, 805–13 (1992).

53 Maccarana M, Casu B, Lindahl U, Minimal sequence in heparin/heparan sulfate required for binding of basic fibroblast growth factor, *J Biol Chem* **268**, 23898–905 (1993).

54 Ishihara M, Structural requirements in heparin for binding and activation of FGF-1 and FGF-4 are different from that for FGF-2, *Glycobiology* **4**, 817–24 (1994).

55 Turnbull JE, Fernig DG, Ke Y, Wilkinson MC, Gallagher JT, Identification of the basic fibroblast growth factor binding sequence in fibroblast heparan sulfate, *J Biol Chem* **267**, 10337–41 (1992).

56 Kreuger J, Prydz K, Pettersson RF, Lindahl U, Salmivirta M, Characterization of fibroblast growth factor 1 binding heparan sulfate domain, *Glycobiology* **9**, 723–9 (1999).

57 Lyon M, Rushton G, Askari JA, Humphries MJ, Gallagher JT, Elucidation of the structural features of heparan sulfate important for interaction with the Hep-2 domain of fibronectin, *J Biol Chem* **275**, 4599–606 (2000).

58 Feyzi E, Trybala E, Bergström T, Lindahl U, Spillmann D, Structural requirement of heparan sulfate for interaction with Herpes simplex virus type 1 virions and isolated glycoprotein C, *J Biol Chem* **272**, 24850–7 (1997).

59 Sasaki T, Larsson H, Kreuger J, Salmivirta M, Claesson-Welsh L, Lindahl U, Hohenester E, Timpl R, Structural basis and potential role of heparin/heparan sulfate binding to the angiogenesis inhibitor endostatin, *Embo J* **18**, 6240–8 (1999).

60 Faham S, Hileman RE, Fromm JR, Linhardt RJ, Rees DC, Heparin structure and interactions with basic fibroblast growth factors, *Science* **271**, 1116–20 (1996).

61 Nurcombe V, Smart CE, Chipperfield H, Cool SM, Boilly B, Hondermarck H, The proliferative and migratory activities of breast cancer cells can be differentially regulated by heparan sulfates, *J Biol Chem* **275**, 30009–18 (2000).

62 Rapraeger AC, Krufka A, Olwin BB, Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation, *Science* **252**, 1705–8 (1991).

63 Guimond S, Maccarana M, Olwin BB, Lindahl U, Rapraeger AC, Activating and inhibitory heparin sequences for FGF-2 (basic FGF). Distinct requirements for FGF-1, FGF-2, and FGF-4, *J Biol Chem* **268**, 23906–14 (1993).

64 Pye DA, Vives RR, Turnbull JE, Hyde P, Gallagher JT, Heparan sulfate oligosaccharides require 6-O-sulfation for promotion of basic fibroblast growth factor mitogenic activity, *J Biol Chem* **273**, 22936–42 (1998).

65 Pye DA, Gallagher JT, Monomer complexes of basic fibroblast growth factor and heparan sulfate oligosaccharides are the minimal functional unit for cell activation, *J Biol Chem* **274**, 13456–61 (1999).

66 Perrimon N, Bernfield M, Specificities of heparan sulphate proteoglycans in developmental processes, *Nature* **404**, 725–8 (2000).

67 Selleck SB, Proteoglycans and pattern formation: sugar biochemistry meets developmental genetics, *Trends Genet* 206–12 (2000).

68 Silbert JE, Biosynthesis of heparin. IV. N-deacetylation of a precursor glycosaminoglycan, *J Biol Chem* **242**, 5153–7 (1967).

69 Silbert JE, Biosynthesis of heparin. III. Formation of a sulfated glycosaminoglycan with a microsomal preparation from mast cell tumors, *J Biol Chem* **242**, 5146–52 (1967).

70 Lindahl U, Feingold DS, Rodén L, Biosynthesis of heparin, *Trends Biochem Sci* **11**, 221–5 (1986).

71 Lindahl U, In *Heparin – chemical and biological properties, clinical applications*, edited by Lane DA, Lindahl U, (Edward Arnold, London, 1989), pp. 159–89.

72 Lindahl U, Kusche-Gullberg M, Kjellen L, Regulated diversity of heparan sulfate, *J Biol Chem* **273**, 24979–82 (1998).

73 Wei Z, Swiedler SJ, Ishihara M, Orellana A, Hirschberg CB, A single protein catalyzes both N-deacetylation and N-sulfation during the biosynthesis of heparan sulfate, *Proc Natl Acad Sci USA* **90**, 3885–8 (1993).

74 Eriksson I, Sandbäck D, Ek B, Lindahl U, Kjellén L, cDNA cloning and sequencing of mouse mastocytoma glucosaminyl N-deacetylase/N-sulfotransferase, an enzyme involved in the biosynthesis of heparin, *J Biol Chem* **269**, 10438–43 (1994).

75 Senay C, Lind T, Muguruma K, Tone Y, Kitagawa H, Sugahara K, Lidholt K, Lindahl U, Kusche-Gullberg M, The EXT1/EXT2 tumor suppressors; catalytic activities and role in heparan sulfate biosynthesis, *EMBO Reports* **1**, 282–6 (2000).

76 Pikas DS, Eriksson I, Kjellén L, Overexpression of different isoforms of glucosaminyl N-deacetylase/N-sulfotransferase results in distinct heparan sulfate N-sulfation patterns, *Biochemistry* **39**, 4552–8 (2000).

77 Aikawa J, Esko JD, Molecular cloning and expression of a third member of the heparan sulfate/heparin G1cNAc N-deacetylase/N-sulfotransferase family, *J Biol Chem* **274**, 2690–5 (1999).

78 Li J-p, Hagner-McWhirter Å, Kjellén L, Palgi J, Jalkanen M, Lindahl U, Biosynthesis of heparin/heparan sulfate. cDNA cloning and expression of D-glucuronyl C5-epimerase from bovine lung, *J Biol Chem* **272**, 28158–63 (1997).

79 Kobayashi M, Habuchi H, Yoneda M, Habuchi O, Kimata K, Molecular cloning and expression of Chinese hamster ovary cell heparansulfate 2-sulfotransferase, *J Biol Chem* **272**, 13980–5 (1997).

80 Rong J, Habuchi H, Kimata K, Lindahl U, Kusche-Gullberg M, Expression of heparan sulphate L-iduronyl 2-O-sulphotransferase in human kidney 293 cells results in increased D-glucuronyl 2-O-sulphation, *Biochem J* **346**, 463–8 (2000).

81 Habuchi H, Tanaka M, Habuchi O, Yoshida K, Suzuki H, Ban K, Kimata K, The occurrence of three isoforms of heparan sulfate 6-O-sulfotransferase having different specificities for hexuronic acid adjacent to the targeted N-sulfoglucosamine, *J Biol Chem* **275**, 2859–68 (2000).

82 Liu J, Shworak NW, Sinay P, Schwartz JJ, Zhang L, Fritze LM, Rosenberg RD, Expression of heparan sulfate D-glucosaminyl 3-O-sulfotransferase isoforms reveals novel substrate specificities, *J Biol Chem* **274**, 5185–92 (1999).

83 Shukla D, Liu J, Blaiklock P, Shworak NW, Bai X, Esko JD, Cohen GH, Eisenberg RJ, Rosenberg RD, Spear PG, A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry, *Cell* **99**, 13–22 (1999).

84 Ringvall M, Ledin J, Holmborn K, van Kuppevelt T, Ellin F, Eriksson I, Olofsson AM, Kjellén L, Defective heparan sulfate biosynthesis and neonatal lethality in mice lacking N-deacetylase/N-sulfotransferase-1, *J Biol Chem* **275**, 25926–30 (2000).

85 Forsberg E, Pejler G, Ringvall M, Lunderius C, Tomasini-Johansson B, Kusche-Gullberg M, Eriksson I, Ledin J, Hellman L, Kjellén L, Abnormal mast cells in mice deficient in a heparin-synthesizing enzyme, *Nature* **400**, 773–6 (1999).