Mini Reviews in Medicinal Chemistry, 2003, *3,* 659-667

The Potential of the Molecular Diversity of Heparin and Heparan Sulfate for Drug Development

Péter Fügedi*

Chemical Research Center, Hungarian Academy of Sciences, H-1525 Budapest, P.O.Box 17, Hungary

Abstract: Heparin and heparan sulfate have been shown to interact with a large number of biologically important proteins regulating important physiological processes. Specific oligosaccharide structures within the heterogeneous polysaccharide chains are responsible for the binding to individual proteins. Identification of specific protein-binding oligosaccharides provides lead compounds in pharmaceutical development and in one case has already resulted in an approved drug. The chemical and biosynthetic basis of the molecular diversity of heparin and heparan sulfate, its manifestation in heparin-protein interactions, and recent progress for drug development offered by this diversity are reviewed.

Keywords: Heparin, heparan sulfate, carbohydrate-protein interactions, diversity, lead compounds, drug development.

INTRODUCTION

Heparin is a well-known blood anticoagulant, which is extensively employed in medical practice [1, 2]. Since its introduction in 1937, it has been widely used as an aid in the surgery of the heart and blood vessels and for control and prevention of postoperative thrombosis. Over the years a range of diverse biological activities not related to blood coagulation have been ascribed to heparin. Physiological effects, such as anti-inflammatory activity, antiatherosclerotic activity, inhibition of smooth muscle cell proliferation, antiangiogenic and antiviral activities [3], have raised the possibility of using heparin (or heparin related molecules) in the treatment of diseases, such as asthma, cancer or Alzheimer's disease [4-6].

For a long time after its introduction to the clinic, the molecular basis for the anticoagulant activity of heparin remained unclear. It was only in 1973 that Rosenberg showed [7] that heparin exerted its anticoagulant effect by binding to a protein, antithrombin III (AT III). The conformational change induced by heparin binding in antithrombin III results in an accelerated reaction of antithrombin with thrombin, leading to an inactive complex of the two proteins. A specific protein binding oligosaccharide sequence responsible for the interaction of heparin with antithrombin III was proposed [8, 9], and that the binding motif is a specific pentasaccharide was proved by chemical synthesis [10, 11]. Surprisingly, this pentasaccharide sequence turned out to be rare in heparin, occurring only in about one-third of the heparin chains.

In the past 1-2 decades, it became obvious that heparin and the closely related molecule, heparan sulfate, interact not only with antithrombin III, but also with a large number of other proteins of diverse biological functions [3, 12-14], and a growing number of biological activities are regulated by these interactions. It was proposed as early as 1979 that in the classic lock-and-key model, heparin can be viewed as "a bag of skeleton keys that can fit many locks" [15]. As our information about heparin's heterogeneity, finer structural details and biosynthesis increases, it is commonly assumed that different structural motifs might be responsible for the different interactions of heparin and heparan sulfate with individual proteins. This opens up a new possibility for drug development of specific pharmaceuticals based on a well-known therapeutic agent.

THE STRUCTURES OF HEPARIN AND HEPARAN SULFATE

Heparin and heparan sulfate belong to the glycosaminoglycan (GAG) family of polysaccharides. They are built up of linear chains of alternating aminosugar and uronic acid units. The aminosugar is D-glucosamine (D-GlcN), while the uronic acid is either D-glucuronic acid (D-GlcA) or its C-5 epimer, L-iduronic acid (L-IdoA). All monosaccharides occur in the pyranose form. The D-glucosamine is linked α-(1→4) to the uronic acid, the uronic acids are linked to O-4 of glucosamine, D-glucuronic acid with a β-(1→4), L-iduronic acid with an α-(1→4) linkage. The typical disaccharide backbones building up the polymer are →4)-β-D-Glc*p*A-(1→4)-α-D-Glc*p*N-(1→ (**1**), and →4)-α-L-Ido*p*A-(1→4)-α-D-Glc*p*N-(1→ (**2**) shown in Fig. (**1**). The alternation of amino sugar and uronic acids gives the structure a great deal of regularity; the change in the uronic acid, however, introduces irregularity. There are characteristic differences in the composition of heparin and heparan sulfate. In heparan sulfate, D-glucuronic acid is the predominant uronic acid; whereas in heparin, L-iduronic acid. The L-iduronic acid content can be as high as 90 % of the total uronic acids in heparin. The composition of both heparin and heparan sulfate, however, varies to a significant extent depending on the source of the glycosaminoglycan.

The carbohydrate chains of heparin and heparan sulfate are substituted at various positions. The nitrogen of D-glucosamine is typically *N*-acetylated or *N*-sulfated, though the occurrence of free amino groups has been reported recently [16, 17]. The hydroxyl groups of both the glucosamine and the uronic acids can carry sulfate

*Address correspondence to this author at the Chemical Research Center, Hungarian Academy of Sciences, H-1525 Budapest, P.O.Box 17, Hungary; Fax: +36-1-325-7863; E-mail: pfugedi@chemres.hu

1389-5575/03 $41.00+.00 © 2003 Bentham Science Publishers Ltd.

substituents. D-Glucosamine can be O-sulfated at O-6 and O-3, whereas the uronic acids (both glucuronic acid and iduronic acid) can be sulfated at O-2. The carbohydrate backbone in heparin is more highly sulfated (as an average about 2.5 sulfate groups per disaccharide unit) than in heparan sulfate (about 1 sulfate per disaccharide unit). The most abundant disaccharide in heparin is the *N*-sulfated, di-*O*-sulfated α-L-Ido*p*A2S-(1→4)-α-D-Glc*p*NS6S (**3**), whereas in heparan sulfate it is the *N*-acetylated, not *O*-sulfated β-D-Glc*p*A-(1→4)-α-D-Glc*p*NAc (**4**) [Fig. (**1**)].

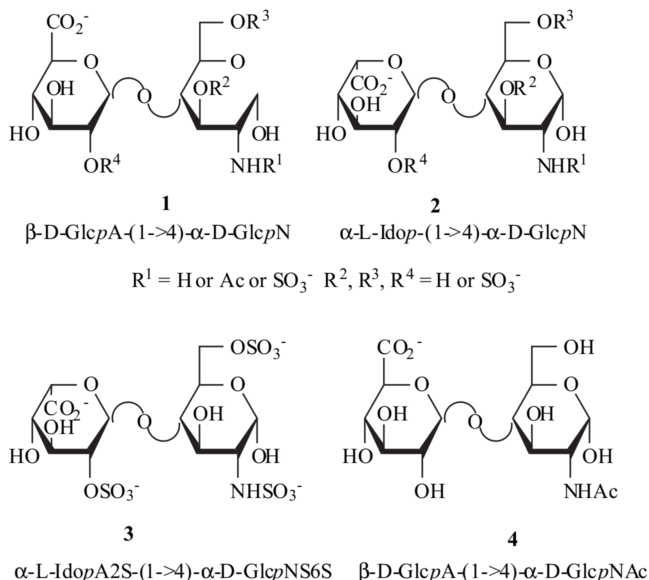


Fig. (1). The disaccharide units of heparin and heparan sulfate. The → 4) -β-D-Glc*p*A- (1 → 4)-α-D-Glc*p*N-(1 → (**1**) and → 4)-α-L-Ido*p*A-(1 → 4)-α-D-Glc*p*N-(1 → (**2**) backbones can be substituted at positions marked by R groups. The most abundant disaccharide component in heparin is **3**, in heparan sulfate it is **4**.

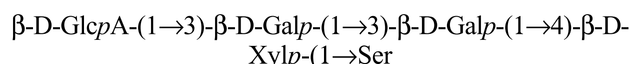
The substitution of the amino and hydroxyl groups results in an enormous structural diversity. Taking the possible permutations and biosynthetic restrictions into account, 24 different oligosaccharide units are considered as possible building blocks of heparin and heparan sulfate [12], and 19 of them have been identified as actual components [Table (**1**)]. From 24 disaccharides 576 tetrasaccharides and 13824 hexasaccharides of different structures can be derived, the number of possible structures reaching astronomical numbers on the polysaccharide level.

Generally in carbohydrate-protein interactions relatively small (from tri- to hexasaccharide) oligosaccharide epitopes of a larger carbohydrate chain are responsible for the interaction. The great structural diversity of heparin and heparan sulfate means that in the structure of these polysaccharides, oligosaccharide epitopes for specific interactions with a multitude of proteins can be encoded. In this regard heparin and heparan sulfate can be looked at as a library of oligosaccharide epitopes. The situation is reversed compared to common combinatorial chemical libraries, where libraries exist as mixtures of individual compounds. In the case of heparin and heparan sulfate, a single molecule can be regarded as a library, which involves a lot of protein-specific structures.

The possible structures are not randomly occurring within the polysaccharide chain. Heparan sulfate is known to have a domain structure in which regions possessing a low degree of sulfation are alternating with domains of higher degree of sulfation [18-20].

BIOSYNTHESIS OF HEPARIN AND HEPARAN SULFATE

Heparin is biosynthesized as a proteoglycan, the carbohydrate chains attached to a protein. The core protein, serglycin, contains a large number of serine and glycine repeats. On the core protein, a tetrasaccharide linkage region is synthesized first by stepwise addition of the individual sugars by four glycosyl transferases. The linkage region tetrasaccharide has the sequence:



Onto the linkage region the first amino sugar, which can be *N*-acetyl-D-glucosamine or *N*-acetyl-D-galactosamine, is added. This initial aminosugar addition determines whether

Table 1. Disaccharide Constituents of Heparin and Heparan Sulfate

GlcA→GlcNAc	IdoA→GlcNAc
GlcA→GlcNAc-6-SO_3^-	IdoA→GlcNAc-6-SO_3^-
GlcA→GlcNSO_3^-	IdoA→GlcNSO_3^-
GlcA→GlcNSO_3^- -6-SO_3^-	IdoA→GlcNSO_3^- -6-SO_3^-
GlcA→GlcNSO_3^- -3-SO_3^-	IdoA→GlcNSO_3^- -3-SO_3^-
GlcA→GlcNSO_3^- -3,6-di-SO_3^-	IdoA→GlcNSO_3^- -3,6-di-SO_3^-
GlcA-2-SO_3^-→ GlcNAc	IdoA-2-SO_3^-→GlcNAc
GlcA-2-SO_3^-→GlcNAc-6-SO_3^-	IdoA-2-SO_3^-→GlcNAc-6-SO_3^-
GlcA-2-SO_3^-→GlcNSO_3^-	IdoA-2-SO_3^-→GlcNSO_3^-
GlcA-2-SO_3^-→GlcNSO_3^- -6-SO_3^-	IdoA-2-SO_3^-→GlcNSO_3^- -6-SO_3^-
GlcA-2-SO_3^-→GlcNSO_3^- -3-SO_3^-	IdoA-2-SO_3^-→GlcNSO_3^- -3-SO_3^-
GlcA-2-SO_3^-→GlcNSO_3^- -3,6-di-SO_3^-	IdoA-2-SO_3^-→GlcNSO_3^- -3,6-di-SO_3^-

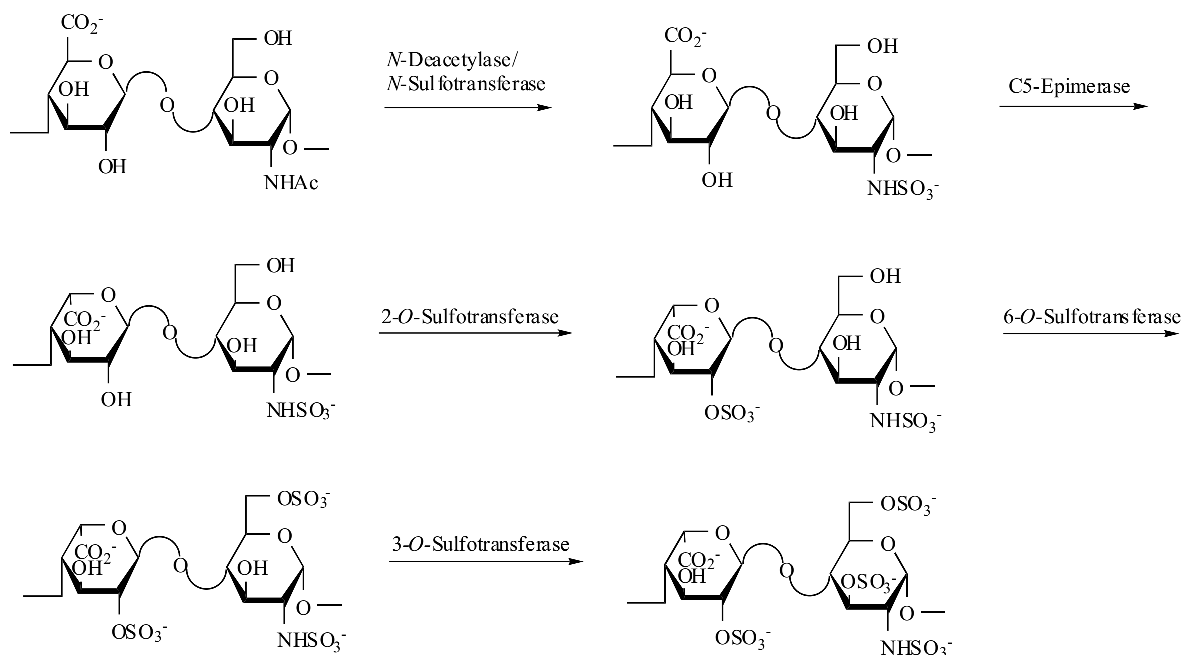


Fig. (2). The biosynthesis of heparin and heparan sulfate.

the synthesized compound will be a glucosamino- (heparin, heparan sulfate), or a galactosamino-glycan (chondroitin sulfates, dermatan sulfate); addition of *N*-acetyl-D-galactosamine results in the latter. After addition of the first aminosugar alternating transfers of D-Glc*p*A and D-Glc*p*NAc residues from the corresponding UDP-sugar nucleotides result in fast growth of the polysaccharide chain. The protein responsible for the addition of both sugars is a hetero-oligomeric complex. Parallel to chain elongation, the carbohydrate backbone also undergoes a series of modification reactions. These modifications start with an *N*-deacetylation - *N*-sulfation reaction, which is performed by an enzyme having dual activity. *N*-Deacetylation/*N*-sulfation is a signal for further chemical modifications [Fig. (**2**)], which occur on the *N*-sulfated glucosamine unit or on the adjacent sugar residues. D-Glucuronic acid is transformed to L-iduronic acid by a C-5 epimerase. The uronic acids are then *O*-sulfated at the *O*-2 position by a sulfotransferase using PAPS as the sulfate donor. The D-glucosamine unit is then 6-*O*-sulfated by a sulfotransferase, that transfers sulfate to both D-Glc*p*NAc and D-Glc*p*NS units. Finally a 3-*O*-sulfotransferase modifies some of the glucosamine residues.

The synthesized proteoglycan molecule undergoes further processing. Proteases release the carbohydrate chain linked to a small peptide; the long carbohydrate chain (100 kDa) is processed by a β-endoglucuronidase releasing a number of smaller (10-20 kDa) polysaccharide chains.

The biosynthesis of heparan sulfate follows essentially the same steps. However, the core proteins (syndecans and glypicans) are different, and the carbohydrate chains remain connected to the core protein; heparan sulfate exists as a proteoglycan. While heparin is synthesized by connective tissue-type mast cells only and is stored in granules, most cell types produce heparan sulfate and it is ubiquitous on cell surfaces and occurs in the extracellular matrix. From the biosynthetic point of view, heparin is more mature than heparan sulfate, its biosynthesis is progressed further; heparin has a higher iduronic acid content and a higher degree of sulfation.

The biosynthetic steps described above do not go to completion, thus resulting in structural heterogeneity of the polysaccharide chains. The seemingly random and incomplete nature of the initial *N*-deacetylation/*N*-sulfation step is responsible for the domain structure of heparan sulfate. Incomplete sulfation results in a varied sulfation pattern. The regulation of the biosynthetic steps is not understood at present. Several of the enzymes in the biosynthetic process exist in different isoforms having different specificities, and the isozymes appear to be expressed in a tissue-specific and developmentally regulated manner [21]. Despite the lack of knowledge on the exact mechanisms in control of generating specific saccharide sequences, several lines of evidence suggest that the diversity of heparan sulfate is regulated [21, 22]. Molecular diversity of heparan sulfate proteoglycans is the key in their specific interactions with proteins.

HEPARIN/HEPARAN SULFATE – PROTEIN INTERACTIONS

Heparin and heparan sulfate interact with a wide range of diverse, biologically important proteins. The number of currently known heparin-binding proteins exceeds 100 and it is increasing rapidly [3, 12-14, 23]. The heparin binding proteins are quite diverse both in structure and in biological functions. Among the proteins are enzymes, such as superoxide dismutase; protease inhibitors, such as antithrombin III; growth factors, such as FGF-1 and FGF-2; chemokines, such as IL-4; extracellular matrix proteins, such as collagens and laminin; selectins (L- and P-selectin); viral coat proteins (HIV-1 gp120); and several others. A detailed discussion of all the heparin binding proteins and the details of the interactions are beyond the scope of this mini-review, some selected examples are presented below.

Interaction of Heparin with Antithrombin III

The best understood of heparin-protein interactions is the binding of heparin with antithrombin III. Heparin exerts its anticoagulant activity primarily through this interaction. Antithrombin III belongs to the serpin (serine protease inhibitor) family of proteins; serpins react with serine proteases to form inactive complexes. All the coagulation factors are inhibited by AT III, forming 1:1 complexes. Heparin binds to AT III and thrombin in a ternary complex, which increases the inhibition of thrombin 2000 fold. The binding of heparin to antithrombin III induces a conformational change [24-26], which facilitates the binding of the protease inhibitor to thrombin. A specific pentasaccharide sequence (**5**) [8, 9] within the heparin chain is responsible for the binding [Fig. (**3**)]. This pentasaccharide has a unique monosaccharide component, GlcNS3S6S, previously not found in heparin, and the pentasaccharide sequence is rare in heparin.

The antithrombin-binding pentasaccharide [10, 11, 27] and a series of analogs have been synthesized. Structure-activity relationship studies of the pentasaccharide analogs [28] revealed that the pentasaccharide possesses two essential *O*-sulfate groups, two essential *N*-sulfate groups, and two essential carboxylate groups. Removal of one of the key sulfate groups leads to significant loss of activity. The type of charge is also important; replacement of an essential sulfate group by phosphate, or a carboxyl group by a $CH_2OSO_3^-$ residue is detrimental to activity. Additional sulfate groups at certain positions are tolerated, whereas in other positions they reduce the activity. The persulfated pentasaccharide is practically inactive. The orientation of the key charged groups in space is important since epimers display reduced activity. Conformational flexibility of the L-iduronic acid is necessary.

A synthetic pentasaccharide (**6**) [Fig. (**3**)] was approved by the FDA for the prophylaxis of deep vein thrombosis in 2001, and marketed under the name Arixtra® (Fondaparinux sodium) [29-31].

The original chemical synthesis [10, 11] of the pentasaccharide is quite demanding, involves a large number of steps, and affords the target compound in low overall yield. Though the synthesis was improved later on, a series of analogs much easier to synthesize were prepared. In these derivatives, frequently referred to as "non-glycosamino"glycans the *N*-sulfate groups are replaced by *O*-sulfates, and the free hydroxyl groups are methylated, such as in **7** on Fig. (**3**) [28, 32]. Some of these compounds showed increased anticoagulant activity compared to the natural pentasaccharide, and might have promise for the development of second-generation synthetic antithrombotics.

For the formation of a ternary complex between antithrombin, thrombin and heparin, oligosaccharides of the size of 16-mers are required, though only the pentasaccharide is needed for binding to antithrombin III. Oligosaccharides

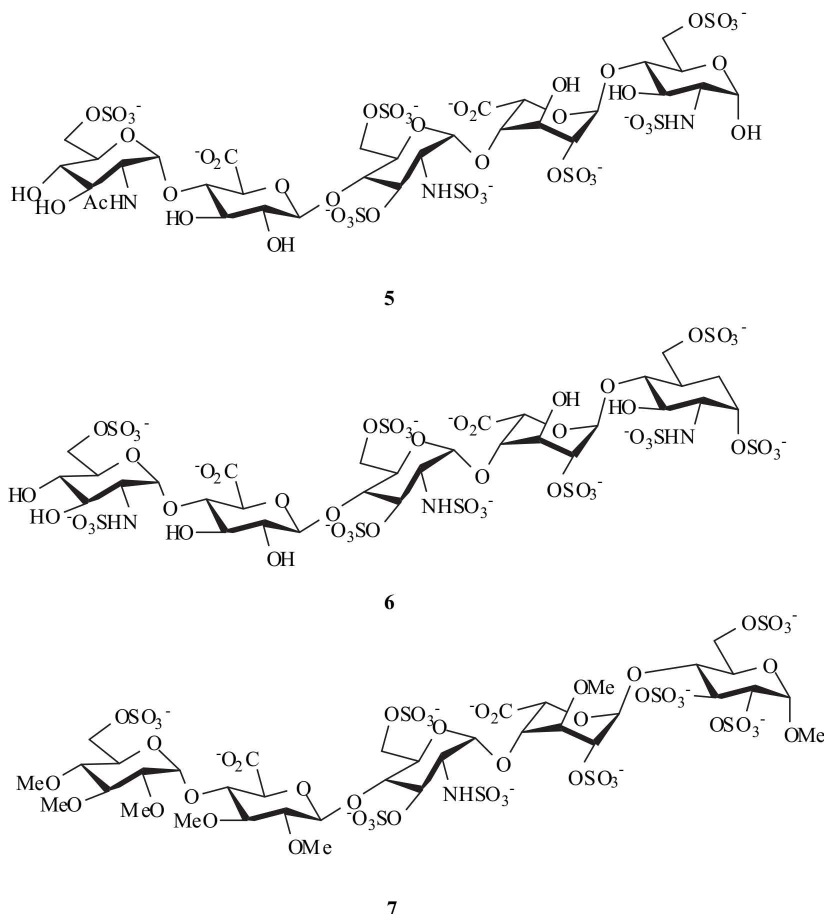


Fig. (3). Antithrombin III binding pentasaccharides. The natural oligosaccharide (**5**), the synthetic drug Arixtra ® (**6**) and a high activity "non-glycosamino" glycan pentasaccharide (**7**).

of this size, incorporating both the antithrombin binding sequence and a sulfated region necessary for binding to thrombin, has been prepared, and are proved to be 5- to 10-fold more potent than standard heparin in models of venous and artherial thrombosis [33]. Importantly, these compounds show promise to be devoid of undesirable side effects, such as interaction with platelet factor 4.

Interaction of Heparin with Growth Factors

Fibroblast growth factors (FGFs) are a family of proteins involved in a wide variety of biological activities, including cell differentiation, proliferation, morphogenesis and angiogenesis [34, 35]. The interaction of heparin with acidic fibroblast growth factor (FGF-1) and basic fibroblast growth factor (FGF-2) is best studied. FGFs exert their biological effects by binding to their cell surface receptors (fibroblast growth factor receptors, FGFRs). Acidic and basic FGFs also bind to heparan sulfate, and heparan sulfate proteoglycans regulate FGF signaling by the formation of a ternary complex between FGF, FGFR and heparan sulfate [36, 37]. High resolution X-ray structures of some of these complexes have been published [38, 39].

A great deal of information on heparan sulfate oligosaccharide sequences required for binding to FGF-1 or FGF-2, and also for the formation of ternary complexes is known. The hexasaccharide **8** on Fig. (**4**) has been identified as a heparin fragment that binds strongly to both FGF-1 and FGF-2 [40-42], but it does not represent the minimum structure required.

Other studies suggested that the GlcN 6-O-sulfates are not essential [42-44] for FGF-2 binding, and oligosaccharides as small as di- and trisaccharides, some of them not even sulfated, have also shown significant FGF-related activities [45]. More recently, from a series of heparan sulfate oligosaccharides having high affinity for FGF-1 and FGF-2, a common trisaccharide motif (**9**) [Fig. (**4**)] has been identified for FGF-1, and a similar motif lacking the 6-O-sulfate group (**10**) seems necessary for high affinity binding to FGF-2 [46].

A great deal of information regarding molecular details of the interactions is provided by crystallographic analysis of growth factors complexed with different heparin oligosaccharides with or without FGFRs [38, 39, 47-49].

Although the optimum and minimum structures required for FGF-binding and activation have not yet been firmly established, it is clear that oligosaccharides significantly differing from the antithrombin binding pentasaccharide are involved. It is noteworthy that the hexasaccharide **8**, Fig. (**4**), and related structures possess high FGF-related and weak blood anti coagulant activities, whereas the situation is the opposite with **5**. This means that the different biological activities of heparin and heparan sulfate can be separated to a large extent, opening the way of drug development for specific activities.

Interactions with Other Proteins

Apolipoprotein E (ApoE) is a lipid transport protein, which plays an important role in lipoprotein metabolism and in the transport and redistribution of lipids in human plasma and brain [50]. An isoform of this protein, ApoE4, is a major genetic risk factor in a number of diseases, including heart disease and Alzheimer's disease [51-54]. ApoE binds to low density lipoprotein (LDL) receptors, and plays a central role in lipoprotein metabolism and cholesterol transport. ApoE has heparin binding sites in both the *N*-terminal and *C*-terminal domains. The high-affinity heparin binding site of the *N*-terminal domain overlaps with the LDL receptor-binding region. It is proposed that in the brain interaction of ApoE with heparan sulfate represents the recognition step that localizes ApoE to the cell surface. An octasaccharide was proposed as the minimum structural element required for ApoE binding [55]. Heparin and heparin oligosaccharides are considered to have potential in the treatment of Alzheimer's, and possibly also of prion diseases [13, 56].

Just like with apolipoprotein E, sequence and size requirements necessary for interactions with several other proteins or for biological activities are also known. For

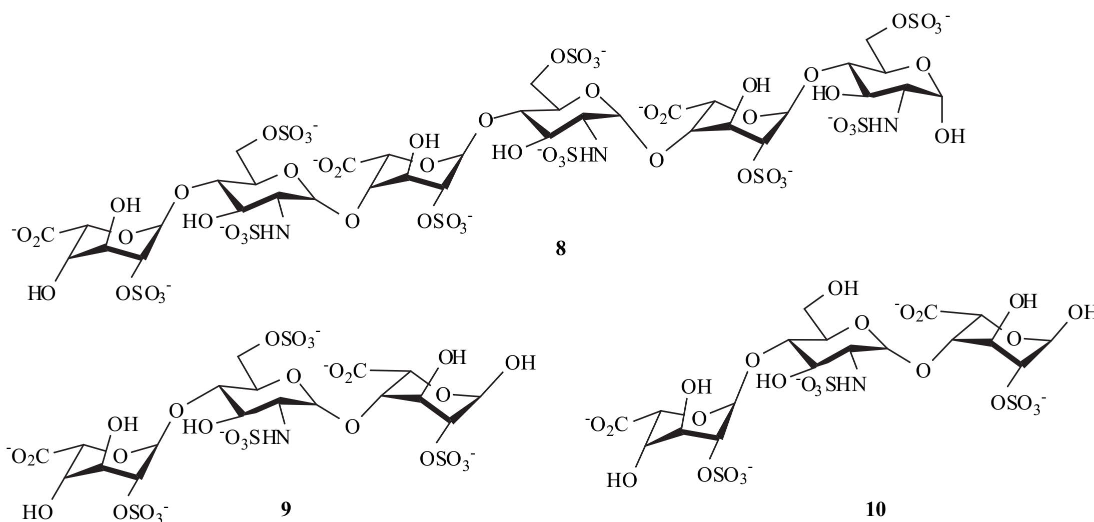


Fig. (4). FGF-binding heparin oligosaccharide structures. The hexasaccharide (**8**) was identified to have high activity both for FGF-1 and FGF-2. The trisaccharide **9** is reported as the minimum binding motif for FGF-1 and trisaccharide **10** for FGF-2.

example, 3-*O*-sulfation of glucosamine has been reported necessary for herpes simplex virus 1 entry into cells [57, 58], and free amino groups of glucosamine are reported to be important in calcium dependent L-selectin binding [16]. Oligosaccharides of the size of tetra- to hexasaccharides are required for the binding to annexins [59, 60], a chain length of 14 monosaccharide residues was reported to be the minimum size needed for inhibition of smooth muscle cell proliferation [61], and the size of a 20-mer is essential for binding to interleukin 8 [62]. The biological significance and details of these, and several other heparin-protein interactions have been recently reviewed [3, 12-14].

Types and Specificity of Heparin-Protein Interactions

For potential therapeutic applications of heparin-protein interactions, two aspects of specificity are of particular interest:

Whether a specific oligosaccharide epitope within the polysaccharide chain is responsible for the binding to a given protein.

How different the oligosaccharide epitopes for different proteins are, and how specific their activities are?

As it could be seen on the examples described in the previous section and also on numerous other ones described in the literature, not the whole polysaccharide, but smaller size fragments of defined structure suffice for binding to a given protein. The fact that highly anionic compounds bind selectively to proteins might, prima facie, seem surprising, as the highly charged nature of the HSPG molecules led people to believe that they bind through non-specific ionic interactions.

Heparin/heparan sulfate-protein interactions definitely have a strong ionic component. Basic amino acids on proteins form ion pairs with negatively charged sulfate and carboxyl groups on the glycosaminoglycan molecule, complementary disposition of the charges are critical for binding. The carbohydrate backbone in this regard can be considered as a scaffold for appropriately positioning the charged groups in space. Heparin adopts a helical conformation [63] exposing the charged groups. It should be noted, however, that the carbohydrate scaffold is not rigid. Though D-glucosamine and D-glucuronic acid prefer the $^{4}C_{1}$ conformation, L-iduronic acid is conformationally mobile and it can adopt $^{4}C_{1}$, $^{1}C_{4}$, and $^{2}S_{o}$ conformations [Fig. (**5**)]. The actual conformation is strongly dependent on the substitution pattern of the L-iduronic acid itself, and the neighboring sugar units in the chain [64]. This conformational flexibility contributes significantly to protein binding in the case of ATIII [28]. The delicate balance how iduronic acid conformation is influenced by its environment is well-illustrated by a crystal structure of an FGF-2 heparin hexasaccharide complex, in which one Ido*p*A2S residue is in the $^{1}C_{4}$ conformation, while another one within the same molecule adopts the $^{2}S_{o}$ conformation [47].

However, besides ionic interactions, other interactions contribute significantly to the binding. The importance of hydrogen bonding between polar amino acids of the protein and functional groups in the GAG chain cannot be neglected [65]. The contribution from van der Waals type hydrophobic interactions for these highly charged molecules might seem surprising first, but several studies reported that in fact ionic interactions contribute to a smaller extent to the binding energy than hydrophobic and hydrogen-bonding interactions [66, 67].

It is generally assumed [46, 68], though not rigorously proven, that different proteins recognize distinct oligosaccharide epitopes of heparan sulfate. Independent regulation of different proteins by heparan sulfate proteoglycans would require uniqueness of oligosaccharide binding epitopes for individual proteins. This aspect of specificity is of significant interest [69], but less understood at present. In the case of blood coagulation, specificity is achieved by having a rare structural component, the 3,6-di-*O*-sulfated D-glucosamine unit in the binding epitope. Though it is appealing to assume other rare monosaccharide components, it seems more likely that most proteins interact with combinations of more common monosaccharide units. It can be hypothesized that addition of one or two additional sulfate groups to specific positions of an otherwise not rare oligosaccharide sequence might be an alternate way to assure uniqueness for protein binding epitopes.

ADVANCES IN LEAD IDENTIFICATION

Determination of the oligosaccharide epitope required for binding provides the lead compound for pharmaceutical development. The major obstacle of drug development at present is the lack of information on protein-binding epitopes on heparan sulfate chains. Identification of the binding epitope is generally attempted by affinity separation of depolymerized heparin oligosaccharides, followed by structure determination by disaccharide analysis and physicochemical methods. This approach is not without problems and in the last 1-2 years other approaches have emerged which represent a shift from this paradigm, and hold promise to facilitate achieving the same goal.

Heparin is readily available and can be depolymerized by chemical (nitrous acid) and enzymatic methods making size-fractionated heparin oligosaccharides relatively easily accessible. The affinity fractionation methods are biased in the sense that more abundant oligosaccharides will be

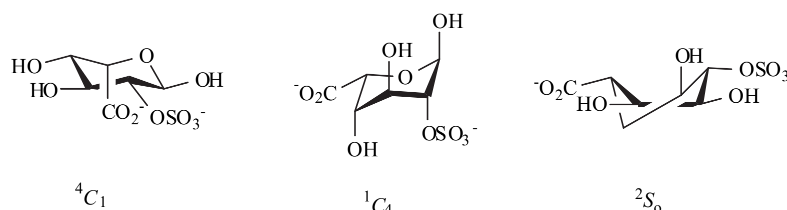


Fig. (5). The conformers of L-iduronic acid 2-O-sulfate.

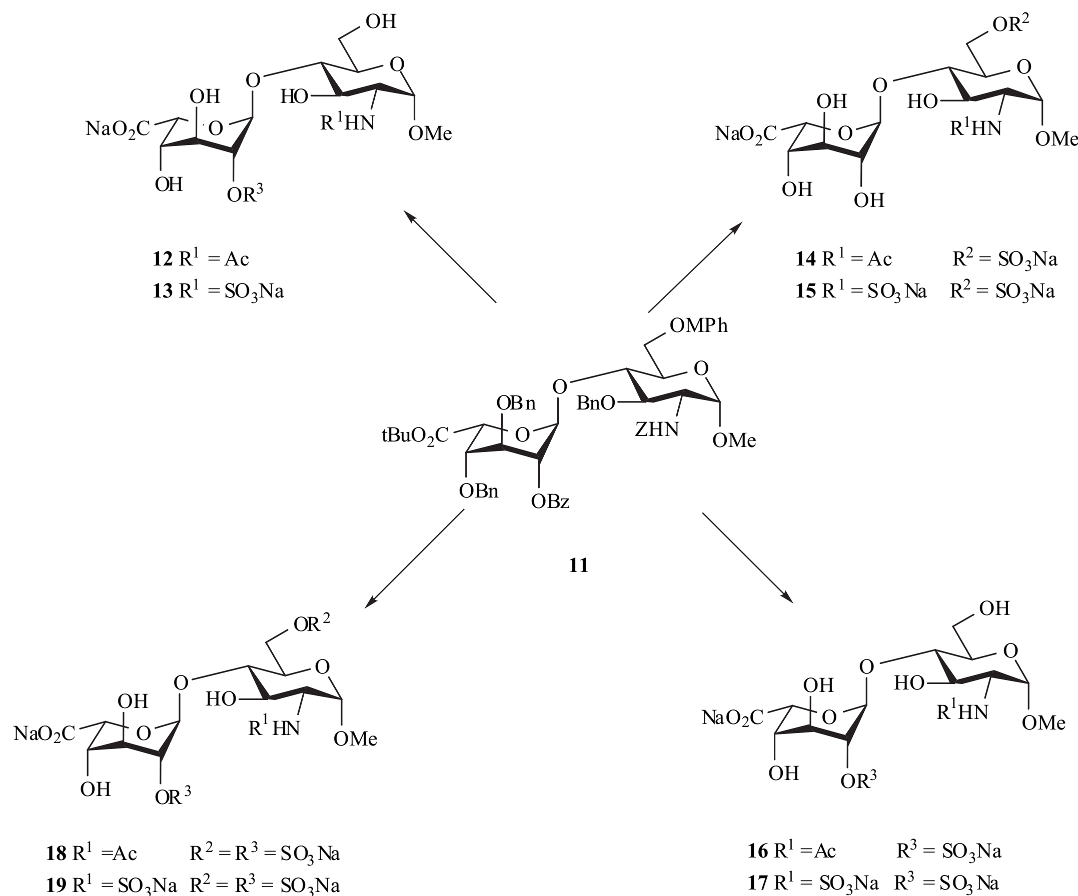


Fig. (6). Orthogonal protection-based synthesis strategy for heparin oligosaccharides. By a sequence of selective removal of orthogonal protecting group(s), *O*-sulfation, deprotection, and N-acetylation or N-sulfation all possible combinations of the sulfated compounds can be prepared from a single protected oligosaccharide.

isolated with higher probability than minor components of similar activity. Also, in nature it is not heparin, but heparan sulfate, which serves as receptor for most proteins. As heparin is more highly sulfated and structurally less diverse than heparan sulfate, the use of heparin oligosaccharides might result in the isolation of tight-binding oligosaccharides in which the minimum epitopes are obscured by additional sulfate groups. This seems to be the case comparing the previously identified FGF-binding hexasaccharide **8** [40, 41] with the recently determined binding motifs (**9** and **10**) [46]. A problem using heparan sulfate oligosaccharides instead of heparin fractions is that they are available only in minute quantities, making structural determinations very difficult. The recently developed sequence analysis methods [70-72] afford more complete structural information on small amounts (picomole) of material.

A strategy different from the one above could be the screening of individual proteins against a library of heparan sulfate oligosaccharides. Use of this strategy was prevented so far by the inaccessibility of the oligosaccharides. Though heparin oligosaccharide library generation is still in its infancy, important developments have emerged. Biosynthetic libraries in the form of mixtures were generated from a chemically *O*-desulfated *N*-sulfated heparin octasaccharide fraction using 6-*O*- and 2-*O*-sulfotransferases [73]. A similar approach was reported starting from a synthetic undersulfated antithrombin III binding pentasaccharide [74].

Chemical synthesis of heparan sulfate oligosaccharide libraries is an approach, which can provide the oligosaccharides as individual entities. A number of heparin oligosaccharides have been synthesized so far, but the majority of these syntheses targeted the antithrombin binding pentasaccharide and its cognate structures [28, 75]. Though chemical synthesis of this kind of oligosaccharides is complex and laborious (the original synthesis of the antithrombin binding pentasaccharide involved some 60 steps!), several groups have taken up the challenge. In one approach [76], synthesis of heparin oligosaccharides is designed by using 19 protected oligosaccharide building blocks corresponding to the sulfated disaccharide units found to occur in the natural polysaccharide. In the building blocks the positions to be sulfated in the target compound are protected by a temporary levulinoyl group [77]. Another approach is designed to take advantage of solid-phase oligosaccharide synthesis [78].

Our synthesis strategy is based on orthogonal protection of the positions, which are optionally sulfated in the target compounds. The strategy is illustrated on Fig. (**6**).

The central protected disaccharide (**11**) has benzoyl and *p*-methoxyphenyl orthogonal protecting groups on *O*-2 of the L-iduronic acid unit and on *O*-6 of the glucosamine,

respectively. Selective removal of the *p*-methoxyphenyl group, followed by *O*-sulfation, deprotection and *N*-acetylation or *N*-sulfation affords two target compounds, **14** and **15**. Alternatively, selective removal of the benzoyl group followed by the same sequence of steps affords the O-2' sulfated derivatives, **16** and **17**. By removal of both orthogonal protecting groups both the di-*O*-sulfated (**18** and **19**) and the non-*O*-sulfated derivatives (**12** and **13**) are accessible. In this way 8 target compounds from a single protected disaccharide can easily be accessed. Using a larger set of orthogonal protecting groups for longer oligosaccharides, any sulfation pattern, any combination of sulfated derivatives of a given oligosaccharide backbone can be prepared from a single protected derivative saving a large number of steps that would be required in a traditional one-by-one synthesis.

SUMMARY AND OUTLOOK

A large number of proteins interact with heparin and heparan sulfate. Different oligosaccharide epitopes within the heterogeneous polysaccharide chain are responsible for the specificity of these interactions. The biological importance of these interactions offers a number of different potential therapeutic applications for heparin and heparin-derived compounds. Identification of specific oligosaccharide sequences that affect a particular biological process enables the development of novel molecular therapeutics.

With increasing realization of the importance of the biological role of heparin and heparan sulfate beyond their anticoagulant activities, more and more research is devoted to understand their other biological functions. The growing interest combined with technical advances is expected to provide not only a better understanding of the details and significance of heparin/heparan sulfate-protein interactions, but also lead structures for pharmaceutical development. The potential of the molecular diversity of these compounds for therapeutical development will be better explored and also better exploited. Recent progress in polysaccharide sequencing and oligosaccharide library generation will contribute significantly to lead identification. Progress in the chemical synthesis of heparin oligosaccharides holds promise for practical applications in the future.

ABBREVIATIONS

ApoE	=	Apolipoprotein E
AT III	=	Antithrombin III
FGF	=	Fibroblast growth factor
FGFR	=	Fibroblast growth factor receptor
GAG	=	Glycosaminoglycan
GlcA	=	Glucuronic acid
GlcN	=	Glucosamine
HSPG	=	Heparan sulfate proteoglycan
IdoA	=	Iduronic acid
LDL	=	Low density lipoprotein
PAPS	=	3'-Phosphoadenosine-5'-phosphosulfate

ACKNOWLEDGEMENTS

The Hungarian Science Fund (OTKA T 32740*),* and the Hungarian National Research and Development Program (1/047 NKFP MediChem) supported this research.

REFERENCES

[1] Lane, D. A.; Lindahl, U. (Eds.), *Heparin, Chemical and Biological Properties, Clinical Applications*, CRC Press: Boca Raton, Florida, **1989.**

[2] Casu, B. *Adv. Carbohydr. Chem. Biochem.*, **1985**, *43*, 51-134.

[3] Linhardt, R. J.; Toida, T. Heparin Oligosaccharides: New Analogues Development and Applications. In *Carbohydrates in Drug Design*; Witczak, Z. J.; Nieforth, K. A., Eds.; Marcel Dekker: New York, **1997**; pp. 277-341.

[4] Tyrrell, D. J.; Kilfeather, S.; Page, C. P. *Trends Pharmacol. Sci.,* **1995**, *16*, 198-204.

[5] Ragazzi, E.; Chinellato, A. *Gen. Pharmac.,* **1995**, *26*, 697-701.

[6] Jaques, L. B. *Pharmacol. Rev.,* **1980**, *31*, 99-167.

[7] Rosenberg, R. D.; Damus, P. S. *J. Biol. Chem.*, **1973**, *248*, 6490-6505.

[8] Lindahl, U.; Bäckström, G.; Thunberg, L.; Leder, I. G. *Proc. Natl. Acad. Sci. USA*, **1980**, *77*, 6551-6555.

[9] Thunberg, L.; Bäckström, G.; Lindahl, U. *Carbohydr. Res.,* **1982**, *100*, 393-410.

[10] Sinaÿ, P.; Jacquinet, J.-C.; Petitou, M.; Duchaussoy, P.; Lederman, I.; Choay, J.; Torri, G. *Carbohydr. Res.*, **1984**, *132*, C5-C9.

[11] Petitou, M.; Duchaussoy, P.; Lederman, I.; Choay, J.; Sinaÿ, P.; Jacquinet, J.-C.; Torri, G. *Carbohydr. Res.*, **1986**, *147*, 221-236.

[12] Conrad, H. E. *Heparin-Binding Proteins*, Academic Press; San Diego, **1998**.

[13] Capila, I.; Linhardt, R. J. *Angew. Chem. Int. Ed.*, **2002**, *41*, 390-412. .

[14] Casu, B.; Lindahl, U. *Adv. Carbohydr. Chem. Biochem.,* **2001**, *57*, 159-206.

[15] Jaques, L. B. *Science,* **1979**, *206*, 528-533.

[16] Norgard-Sumnicht, K.; Varki, A. *J. Biol. Chem.*, **1995**, *270*, 12012-12024.

[17] van den Born, J.; Gunnarsson, K.; Bakker, M. A. H.; Kjellén, L.; Kusche-Gullberg, M.; Maccarana, M.; Berden, J. H. M.; Lindahl, U. *J. Biol. Chem.,* **1995**, *270*, 31303-31309.

[18] Maccarana, M.; Sakura, Y.; Tawada, A.; Yoshida, K.; Lindahl, U. *J. Biol. Chem.*, **1996**, *271*, 17804-17810.

[19] Gallagher, J. T.; Turnbull, J. E.; Lyon, M. *Int. J. Biochem.,* **1992**, *24*, 553-560.

[20] Turnbull, J. E.; Gallagher, J. T. *Biochem. J.,* **1991**, *277*, 297-303.

[21] Esko, J. D.; Lindahl, U. *J. Clin. Invest.,* **2001**, *108*, 169-173.

[22] Lindahl, U.; Kusche-Gullberg, M.; Kjellén, L. *J. Biol. Chem.*, **1998**, *273*, 24979-24982.

[23] Mulloy, B.; Linhardt, R. J. *Curr. Opin. Struct. Biol.,* **2001**, *11*, 623-628.

[24] Villaneuva, G. B.; Danishefsky, I. *Biochem. Biophys. Res. Commun.,* **1977**, *74*, 803-809.

[25] Jin, L.; Abrahams, J. P.; Skinner, R.; Petitou, M.; Pike, R. N. *Proc. Natl. Acad. Sci. USA,* **1997**, *94*, 14683-14688.

[26] Whisstock, J. C.; Pike, R. N.; Jin, L.; Skinner, R.; Pei, X. Y.; Carrell, R. W.; Lesk, A. M. *J. Mol. Biol.,* **2000**, *301*, 1287-1305.

[27] van Boeckel, C. A. A.; Beetz, T.; Vos, J. N.; de Jong, A. J. M.; van Aelst, S. F.; van den Bosch, R. H.; Mertens, J. M. R.; van der Vlugt, F. A. *J. Carbohydr. Chem.,* **1985**, *4*, 293-321.

[28] van Boeckel, C. A. A.; Petitou, M. *Angew. Chem. Int. Ed.,* **1993**, *32*, 1671-1690.

[29] Walenga, J. M.; Jeske, W. P.; Samama, M. M.; Frapaise, F. X.; Bick, R. L.; Fareed, J. *Expert Opinion Investigational Drugs*, **2002**, *11*, 397-407.

[30] Cheng, J. M. W. *Clin. Therapeutics*, **2002**, *24*, 1757-1769.

[31] Keam, S. J.; Goa, K. L. *Drugs*, **2002**, *62*, 1673-1685.

[32] Basten, J. E. M.; Jaurand, G.; Olde-Hanter, B.; Duchaussoy, P.; Petitou, M.; van Boeckel, C. A. A. *Bioorg. Med. Chem. Lett.,* **1992**, *2*, 905-908.

[33] Petitou, M.; Hérault, J.-P.; Bernat, A.; Driguez, P.-A.; Duchaussoy, P.; Lormeau, J.-C.; Herbert, J.-M. *Nature*, **1999**, *398*, 417-422.

[34] Gospodarowicz, D.; Neufeld, G.; Schweigerer, L. *J. Cell. Physiol. Suppl. 5*, **1987**, 15-26.
[35] Klagsbrun, M. *Progress Growth Factor. Res.,* **1989**, *1*, 207-235.
[36] Yayon, A.; Klagsbrun, M.; Esko, J. D.; Leder, P.; Ornitz, D. M. *Cell,* **1991**, *64*, 841-848.
[37] Guimond, S. E.; Turnbull, J. E. *Curr. Biol.,* **1999**, *9*, 1343-1346.
[38] Schlessinger, J.; Plotnikov, A. N.; Ibrahimi, O. A.; Eliseenkova, A. V.; Yeh, B. K.; Yayon, A.; Linhardt, R. J.; Mohammadi, M. *Mol. Cell*, **2000**, *6*, 743-750.
[39] Pellegrini, L.; Burke, D. F.; von Delft, F.; Mulloy, B.; Blundell, T. L. *Nature*, **2000**, *407*, 1029-1034.
[40] Barzu, T.; Lormeau, J.-C.; Petitou, M.; Michelson, S.; Choay, J. *J. Cell. Physiol.* **1989**, *140*, 538-548.
[41] Tyrrell, D. J.; Ishihara, M.; Rao, N.; Horne, A.; Kiefer, M. C.; Stauber, G. B.; Lam, L. H.; Stack, R. J. *J. Biol. Chem.*, **1993**, *268*, 4684-4689.
[42] Maccarana, M.; Casu, B.; Lindahl, U. *J. Biol. Chem.* **1993**, *268*, 23898-23905.
[43] Habuchi H.; Suzuki, S.; Saito, T.; Tamura, T.; Harada, T.; Yoshida, K.; Kimata, K. *Biochem. J.*, **1992**, *285*, 805-813.
[44] Turnbull, J. E.; Fernig, D. G.; Ke, Y.; Wilkinson, M. C.; Gallagher, J. T. *J. Biol. Chem.*, **1992**, *267*, 10337-10341.
[45] Ornitz, D. M.; Herr, A. B.; Nilsson, M.; Westman, J.; Svahn, C.-M.; Waksman, G. *Science*, **1995**, *268*, 432-436.
[46] Kreuger, J.; Salmivirta, M.; Sturiale, L.; Giménez-Gallego, G.; Lindahl, U. *J. Biol. Chem.,* **2001**, *276*, 30744-30752.
[47] Faham, S.; Hileman, R. E.; Fromm, J. R.; Linhardt, R. J.; Rees, D. C. *Science*, **1996**, *271*, 1116-1120.
[48] Pellegrini, L. *Curr. Opin. Struct. Biol.,* **2001**, *11*, 629-634.
[49] Faham, S.; Linhardt, R. J.; Rees, D. C. *Curr. Opin. Struct. Biol.,* **1998**, *8*, 578-586.
[50] Weisgraber, K. H. *Adv. Protein Chem.*, **1994**, *45*, 249-302.
[51] Weisgraber, K. H.; Mahley, R. W. *FASEB J.*, **1996**, *10*, 1485-1494.
[52] Davignon, J.; Cohn, J. S.; Mabile, L.; Bernier, L. *Clin. Chim. Acta*, **1999**, *286*, 115-143.
[53] Corder, E. H.; Saunders, A. M.; Risch, N. J.; Strittmatter, W. J.; Schmechel, D. E.; Gaskell, Jr., P. C.; Rimmler, J. B.; Locke, P. A.; Conneally, P. M.; Schmader, K. E.; Small, G. W.; Roses, A. D.; Haines, J. L.; Pericak-Vance, M. A. *Nat. Genet.*, **1994**, *7*, 180-184.
[54] Strittmatter, W. J.; Roses, A. D. *Annu. Rev. Neurosci.*, **1996**, *19*, 53-77.
[55] Dong, J.; Peters-Libeu, C. A.; Weisgraber, K. H.; Segelke, B. W.; Rupp, B.; Capila, I.; Hernáiz, M. J.; LeBrun, L. A.; Linhardt, R. J. *Biochem.,* **2001**, *40*, 2826-2834.
[56] Kisilevsky, R.; Lemieux, L. J.; Fraser, P. E.; Kong, X.; Hultin, P. G.; Szarek, W. A. *Nat. Med.,* **1995**, *1*, 143-148.
[57] Shukla, D.; Liu, J.; Blaiklock, P.; Shworak, N. W.; Bai, X. M.; Esko, J. D.; Cohen, G. H.; Eisenberg, G. H.; Rosenberg, R. D.; Spear, P. G. *Cell,* **1999**, *99*, 13-22.
[58] Shukla, D.; Spear, P. G. *J. Clin. Invest.,* **2001**, *108*, 503-510.
[59] Capila, I.; van der Noot, V. A.; Mealy, T. R.; Seaton, B. A.; Linhardt, R. J. *FEBS Lett.,* **1999**, *446*, 327-330.
[60] Capila, I.; Hernaiz, M. J.; Mo, Y. D.; Mealy, T. R.; Campos, B.; Dedman, J. R.; Linhardt, R. J.; Seaton, B. A. *Structure*, **2001**, *9*, 57-64.
[61] Garg, H. G.; Cindhuchao, N.; Quinn, D. A.; Hales, C. A.; Thanawiroon, C.; Capila, I.; Linhardt, R. J. *Carbohydr. Res.,* **2002**, *337*, 2359-2364.
[62] Spillman, D.; Witt, D.; Lindahl, U. *J. Biol. Chem.,* **1998**, *273,* 15487-15493.
[63] Mulloy, B.; Forster, M. J. *Glycobiology,* **2000**, *10*, 1147-1156.
[64] Ferro, D. R.; Provasoli, A.; Ragazzi, M.; Torri, G.; Casu, B.; Gatti, G.; Jacquinet, J.-C.; Sinaÿ, P.; Petitou, M.; Choay, J. *J. Am. Chem. Soc.,* **1986**, *108*, 6773-6778.
[65] Hileman, R. E.; Jennings, R. N.; Linhardt, R. J. *Biochemistry*, **1998**, *37*, 15231-15237.
[66] Thompson, L. D.; Pantoliano, M. W.; Springer, B. A. *Biochemistry*, **1994**, *33*, 3831-3840.
[67] Olson, S. T.; Björk, I. *J. Biol. Chem.*, **1991**, *266*, 6353-6354.
[68] Sasisekharan, R.; Vekataraman, G. *Curr. Opin. Chem. Biol.,* **2000**, *4*, 626-631.
[69] Spillman, D.; Lindahl, U. *Curr. Opin. Struct. Biol.,* **1994**, *4*, 677-682.
[70] Turnbull, J. E.; Hopwood, J. J.; Gallagher, J. T. *Proc. Natl. Acad. Sci. USA,* **1999**, *96*, 26908-2703.
[71] Venkataraman, G.; Shriver, Z.; Raman, R.; Sasisekharan, R. *Science,* **1999**, *286*, 537-542.
[72] Vives, R. R.; Pye, D. A.; Salmivirta, M.; Hopwood, J. J.; Lindahl, U.; Gallagher, J. T. *Biochem. J.,* **1999**, *339*, 767-773.
[73] Jemth, P.; Kreuger, J.; Kusche-Gullberg, M.; Sturiale, L.; Giménez-Gallego, G.; Lindahl, U. *J. Biol. Chem.,* **2002**, *277*, 30567-30573.
[74] Wu, Z. L.; Zhang, L.; Beeler, D. L.; Kuberan, B.; Rosenberg, R. D. *FASEB J.,* **2002**, *16*, 539-545.
[75] Yeung, B. K. S.; Chong, P. Y. C.; Petillo, P. A. *J. Carbohydr. Chem.,* **2002**, *21*, 799-865.
[76] Haller, M.; Boons, G.-J. *J. Chem. Soc., Perkin Trans. 1*, **2001**, 814-822.
[77] Haller, M. F.; Boons, G.-J. *Eur. J. Org. Chem.,* **2002**, 2033-2038.
[78] Seeberger, P. H.; Orgueira, H.; Schell, P. Solid- and solution-phase synthesis of heparin and other glycosaminoglycans. *US Patent Application,* PCT WO 02/058633 A2.

Copyright © 2003 EBSCO Publishing