Introduction

Heparin, a sulfated polysaccharide, is a major clinical anticoagulant. It is exploited to overcome the natural propensity for blood to clot. Surgical procedures often require anticoagulation as do extracorporeal therapies, such as heart lung oxygenation and kidney dialysis.1 Whenever blood contacts damaged endothelium in a surgical wound or the synthetic surface of a medical device, the coagulation cascade is activated and a clot is formed (Figure 1).2

Despite heparin's widespread medical use, its precise chemical structure, the range of its biological activities, and the structure–activity relationship (SAR) for these activities is not yet well understood. This paper will review some of the important chemical and biological aspects of heparin and the related polysaccharide, heparan sulfate, with a focus on major research accomplishments taking place in our laboratory and others over the past 25 years. Finally, future directions in heparin research will be discussed.

Early History

The heparin story1,3,4 began during World War I in 1916 at Johns Hopkins University in Baltimore. Jay McLean, a second-year medical student working under the direction of William Howell was searching for substances within the blood that caused it to clot. Instead, McLean isolated fractions from mammalian tissues that inhibited blood coagulation. After some initial skepticism, Howell, recognizing the importance of the discovery of his student, suggested using this isolate to treat coagulation disorders. The discovery of this anticoagulant was controversial because McLean misnamed the extract “hepar (liver) phosphatide”. Heparin, later renamed by Howell, was instead a sulfated polysaccharide.

An understanding of heparin’s structure developed slowly. Howell determined that heparin had no phosphate and was a carbohydrate. Sune Bergstrom, winner of the Nobel Prize for research on prostaglandins, correctly identified glucosamine as a sugar component in heparin while working as a student working with Eric Jorpes in Sweden. Jorpes and later Arthur Charles established that heparin contained a high content of...
structures,7 trisulfated disaccharide repeating unit (Figure 2a), it in italics are important (25–50% loss in binding energy on removal), and those undersulfated structural variants; (c) antithrombin III pen-
saccharide (DEFGH) binding site. The anionic groups in bold
are critical (95% loss in binding energy on removal), and those
in italics are important (25–50% loss in binding energy on removal) for interaction with ATIII.

Heparin is prepared by extraction from tissues of
animals suitable for food (i.e., porcine intestine, bovine
lung). Like all other polysaccharides, heparin is a
polydisperse mixture containing chains having different
molecular weights.5,6 While heparin consists of a major
trisulfated disaccharide repeating unit (Figure 2a), it
also contains a number of additional disaccharides
structures,7–10 making heparin’s structure complex.

Heparin is the most highly sulfated member of the
heparan sulfate family of glycosaminoglycans (GAGs).1,11
Heparin differs from other members of this family in
that it is biosynthesized as a proteoglycan (PG) attached
to the serglycin protein core, uniquely found intracel-
larly, in the granules of mast cells. In contrast,
heparan sulfate PGs are ubiquitous and found extra-
cellulary in virtually all types of animal tissues.12
Tissues, from which heparin is prepared, are rich in
mast cells but also contain heparan sulfate PGs. All the
disaccharide repeating units found within heparin are
also found within heparan sulfate but in different
proportions.7,13 Their relative proportions of these
disaccharides explain why heparin and heparan sulfate
differ substantially in their level of sulfation and in their
anticoagulant activity.13 Extraction methods that focus
on recovery of heparin with high specific anticoagulant
activity, required to meet United States Pharmacopeal
(USP) specifications, eliminate much (but not all) of the
heparan sulfate from pharmaceutical heparin.

Pharmaceutical heparin is a purified polydisperse
(Mw/Mn ≈ 1.1–1.6)6 mixture of chains11 with an average
molecular weight of ~12 kDa and containing significant
sequence heterogeneity. Fully sulfated heparin chains
have three sulfo groups per disaccharide unit, while
undersulfated heparin chains (or contaminating hepa-
ran sulfate) contain less than two sulfo groups per
disaccharide unit (Figure 2b). Only about one-third of
the chains comprising pharmaceutical grade heparin
contain a binding site for an important protein in the
cogulation cascade (Figure 1), called antithrombin III
(ATIII), and are termed “high-affinity heparin”.14

Heparin is found primarily in mast cells, granulated
cells that are found in organs such as liver, intestine,
and lung that are commonly exposed to parasites.15,16
While the precise physiological function of the endog-
neutral heparin in mast cells is not known,17 it is clearly
not involved in blood coagulation (a role most likely
relegated to anticoagulant heparan sulfate on the lu-
minal endothelium).18,19 Heparins obtained from differ-
ent tissues and various species also differ structur-
ally.8,15,16 While porcine intestine is the most widely
used tissue for the preparation of pharmaceutical hepa-
rin, bovine lung affords a pharmaceutical heparin that
differs slightly in structure. Porcine intestinal heparin
contains an ATIII binding site primarily having an
N-acetyl group in residue D (Figure 2c), while bovine
lung heparin primarily has an N-sulfo group at residue
D, resulting in their slightly different affinities for
ATIII.8 Human heparin, first isolated in our laboratory
from a hemangioma rich in mast cells, is structurally
similar to porcine intestinal heparin.20

Heparin is prepared from animal organ tissue col-
lected at a slaughterhouse. In the case of porcine
heparin, the whole intestine is used to prepare “hashed
pork guts” or the empty intestine is scraped to be
processed into both casings for food use and mucosa for
preparation of intestinal mucosal heparin.21 Next, hepa-
rin is separated from the tissue using elevated
temperatures and pressures and/or proteases and re-
covered by precipitation using quaternary ammonium
salts or by adding anion-exchange resin. After precipita-
tion or resin adsorption of the heparin, impurities are
removed by washing and heparin is recovered by res-
diluting with sodium chloride solution. The concen-
trated crude heparin solution is desalted by diafiltration
or precipitated with ethanol, and it is vacuum-dried.
Crude heparin is purified under cGMP conditions to remove impurities such as other GAGs, extraneous cationic counterions, heavy metals, residual proteins or nucleic acids, solvent, salts other than heparin, bacterial endotoxins, bioburden, and viruses. Generally, purification involves oxidation and heating steps to sanitize, decolor, depyrogenate, and inactivate any viruses that might be present. Finally, the purified heparin is precipitated with ethanol, dried or redissolved in purified water, sterile filtered, and freeze-dried. The yield of porcine intestinal heparin is typically 250 mg/kg of wet tissue corresponding to 30000–50000 U/animal.

### Biosynthesis

Extensive studies by Lindahl and co-workers have shown that heparin and heparan sulfate are biosynthesized through a similar pathway (Figure 3). First, the core protein is synthesized in the endoplasmic reticulum, and then as the core protein transits the Golgi, a tetrasaccharide linker is attached in regions rich in Ser-Gly repeats. From this linkage, a repeating 1→4 glycosidically linked copolymer of D-glucuronic acid and N-acetyl-D-glucosamine is synthesized through the stepwise addition of UDP-activated sugars catalyzed by EXT enzyme. During its formation, the linear homopolymer is sequentially modified through the action of N-deacetylase/N-sulfotransferases, C5 epimerase, and 2-, 6-, and 3-O-sulfotransferases (OSTs). Complete or nearly complete modification of this nascent GAG chain results in a highly N- and O-sulfot, L-iduronic acid rich GAG called heparin. Partial modification of the same chain results in an O-sulfot, N-acetyl-D-glucosamine, D-glucuronic acid rich glycosaminoglycan called heparan sulfate. During the past 5 years all of the enzymes involved in the biosynthesis of heparin and heparan sulfate have been cloned and expressed. A major finding of these studies is that many of these biosynthetic enzymes had multiple isoforms that are differentially expressed in various cell types. Despite these major advances in understanding heparin biosynthesis, it is still unclear how biosynthesis (i.e., extent, type of chain modification, and domain placement) is controlled.

Since heparin and heparan sulfate share a common biosynthetic pathway, all of the disaccharide units found in the heparin are also observed in heparan sulfate and vice versa. Moreover, a number of extended sequences, such as the ATIII pentasaccharide binding site (Figure 2c), can be found in both heparin and heparan sulfate. An understanding of the complexity of heparin biosynthesis is best exemplified by the formation of the ATIII binding site.

The biosynthesis of the ATIII binding site requires the placement of a critical 3-O-sulfonated group in its central glucosamine residue F (Figure 2C). Early work from our laboratory demonstrated that not all sites within heparin containing an ATIII binding site precursor sequence are 3-O-sulfonated. Moreover, Lindahl and Dietrich demonstrated that some heparins, particularly from invertebrates, contained 3-O-sulfoglucosamine residues in sites that were missing the other requisite structures in the ATIII binding site. The explanation for such unusual structures awaited the purification, cloning, expression, and specificity studies of the biosynthetic enzymes responsible for 3-O-sulfonation. Work-
ing on heparan sulfate having an extremely low content of 3-O-sulfoglucosamine residues.

Rosenberg and co-workers isolated multiple isoforms of 3-OST's responsible for heparan sulfate biosynthesis.24 One very important feature of the 3-OST isoforms is that they each act within different substrate precursor sequences to afford structurally distinct 3-O-sulf-containing sequences.31

Isoform 1, 3-OST-1, which is expressed in endothelial cells, mast cells, and many other tissues, specifically 3-O-sulfonates, the precursor structure found in heparan sulfate and heparin,31 form the ATIII binding site responsible for heparin's anticoagulant activity. In heparan sulfate, the ATIII binding site is believed to be responsible for the blood compatibility of the vascular endothelium.18 In contrast, modification of heparan sulfate by 3-OST-3 results in a 3-O-sulfonation of an unsubstituted 6-O-sulfoglucosamine that serves as a viral entry receptor for HSV-1.32,33 The 3-OST isoform specificity currently appears to be somewhat promiscuous, but the precise specificities are still under investigation.

The in vivo importance of heparan sulfate's 3-O-sulf group has been examined by Rosenberg and co-workers in 3-OST-1 knockout mice. Surprisingly, the 3-OST-1 knockout mouse shows no coagulation abnormalities.24 One explanation for the absence of coagulation abnormalities in the 3-OST-1 knockout mice could be a compensatory mechanism involving the up-regulation of another 3-OST isofrom in the endothelium. The 3-OST knockout mice show an unexpected phenotype. All these mice develop aseptic eye lesions and exhibit lacrimal gland pathology, interuterine growth retardation, and male infertility. It is currently unclear what missing protein binding activities account for these phenotypes.

All 12 of the enzymes involved in the biosynthesis of heparan sulfate have now been cloned.24 The genes for these enzymes have been identified from a variety of animal species ranging from Drosophila melanogaster to mammals. Recombinant enzymes have also been successfully prepared, generally in either yeast or baculovirus expression systems because many of these enzymes are glycosylated. The sulfotransferases involved in heparin/heparan sulfate biosynthesis include N-deacetylasel/N-sulfotransferase (NDST) (three isoforms), 2-O-sulfotransferase (2-OST) (one isoform), 6-O-sulfotransferase (6-OST) (three to four isoforms), and 3-O-sulfotransferase (3-OST) (four to five isoforms). Some limited information is available on the specificity of these sulfotransferase isoforms.24 NDST-1 appears to be involved in heparan sulfate biosynthesis, N-deacetylation, and N-sulfonating small domains within the chain, while NDST-2 probably is involved in heparin biosynthesis affording nearly fully modified chains. NDST-1 knockout mice die from a lung defect, while NDST-2 knockout mice have glycosaminoglycan chains with very low levels of trisulfated disaccharide, resulting in mice with eye abnormalities. The double knockout (NDST-1 and -2) is lethal at all stages of development. 2-OST knockout mice have heparan sulfate with increased 6-O-sulfo group and N-sulfo group content, presumably a compensatory effect that affords heparan sulfate still capable of activating bFGF. 6-OST isoform 3 (6-OST-3) shows differential temporal expression in developing mouse brain.

Structure

Heparin is a linear polysaccharide consisting of 1→4 linked pyranosyluronic acid (uronic acid) and 2-amino-2-deoxyglucopyranose (D-glucosamine, GlcN) repeating units.26 The uronic acid usually comprises 90% L-idopyranosyluronic acid (L-iduronid acid, IdoA) (Figure 2a) and 10% D-glucopanosyranosyluronic acid (D-glucuronic acid, GlcA). Heparin, with its high content of sulfo and carboxyl groups, is a polyelectrolyte, having the highest negative charge density of any known biological macromolecule. At the disaccharide level, a number of structural variations exist (Figure 2b), leading to sequence microheterogeneity within heparin. GAG heparin is polydisperse with a molecular weight range of 5–40 kDa, an average molecular weight of ~12 kDa, and an average negative charge of about ~75, making it an extremely challenging molecule to characterize.

Heparin's complexity extends through multiple structural levels. At the PG level, different numbers of GAG chains (possibly having different saccharide sequences) can be attached to the various serine residues present on the core protein. Heparin chains are biosynthesized attached to a unique core protein, serglycin, found primarily in mast cells. On mast cell degranulation, proteases act on the heparin core protein to release peptidoglycan heparin, which is further processed by a β-endoglucuronidase into GAG heparin.31 The chemical, physical, and biological properties of heparin are primarily ascribed to GAG structure (or sequence), saccharide conformation, chain flexibility, molecular weight, and charge density.

Most studies, notably ones from our laboratory, have focused on heparin's primary structure. Compositional analysis typically relies on enzymatic (using heparin lyases) or chemical (using nitrous acid) depolymerization of heparin into disaccharides that are then measured by strong anion exchange (SAX) or reversed-phase ion pairing (RP/IP), HPLC, or capillary electrophoresis (CE).38 Oligosaccharide mapping, utilizing the same separation methods as well as NMR, MS, and polycrylamide gel electrophoresis (PAGE), relies on partial depolymerization affording larger oligosaccharides and some sequence information.7,9 Sequencing heparin is a complex task and has yielded limited success relying on the use of specific heparin lyases, nitrous acid, periodate oxidation, exoglucosidase lysozyme enzymes.42 Recently, LC–MS has been used in our laboratory in conjunction with heparin lyases to determine heparin's sequence.43 In addition to standard separation methods, affinity chromatography and affinity electrophoresis are often utilized to confirm the presence of protein binding sites within a heparin chain. Footprinting experiments are also useful to localize protein binding sites within heparin chains.46

Heparan sulfate, while structurally related to heparin, is much less substituted with sulfo groups and has a more varied structure (or sequence) (Figure 2).47 D-glucuronic acid predominates in heparan sulfate, and it is polydisperse, having an average molecular weight of ~30 kDa ranging from 5 to 50 kDa.37 Heparan sulfate chains also often contain domains of extended sequences having low or high sulfation.48
Heparan sulfate, while biosynthesized through the same pathway as heparin, remains connected to its core protein. Two types of heparan sulfate PGs, the syndecans (an integral membrane protein) and the glypicans (a GPI-anchored protein), are ubiquitously distributed on cell surfaces and the extracellular matrix. The heparan sulfate chains on these heparan sulfate PGs bind a variety of proteins and are responsible for mediating many physiologically important processes. Although structurally similar, heparin and heparan sulfate GAGs can often be distinguished through their different sensitivity toward a family of GAG-degrading, microbial enzymes, the heparin lyases.

Heparin and heparan sulfate exist primarily as an extended helical structure and is not known to fold into any tertiary structures. The specificity of heparin’s interaction with proteins depends on how it displays its sulfo and carboxyl groups. The conformational flexibility of the L-iduronic acid residue within heparin is believed to be responsible for the wide range of specific protein interactions exhibited by this family of GAGs. This flexibility depends on substitution and its relative position in the chain. An internal L-IdoA2S residue has two nearly equal energy conformations, the 1C4 chair and the 2S0 skew boat. NMR analysis of the conformation of individual sugars within heparin indicates that the δ-glucosamine and δ-glucuronic acid residues assume the preferred 4C1 conformation. While the iduronate rings are flexible and in equilibrium between the 2S0 and 1C4 conformation in the unbound state, they can be locked into either conformation when bound to protein.

Despite the conformational flexibility of the L-iduronate residues, heparin saccharides show relatively conserved φ-ψ angles in glycosidic linkages. The helical parameters for heparin oligosaccharides bound to fibroblast growth factors have been determined by X-ray crystallography, and these values are comparable with about five residues per turn of the heparin helix. Molecular dynamic simulations suggest that while the helix backbone of heparin remains extended, the sulfo groups can rapidly reposition because of the conformational flexibility of L-iduronic acid residues and rotation around C5-C6 of the 6-sulfoglucosamine residue.

**Anticoagulant Activity**

The anticoagulant action of pharmaceutical heparin (120–180 USP U/mg) is the most thoroughly studied of its activities. Anticoagulation occurs when heparin binds to ATIII, a serine protease inhibitor (serpin). ATIII undergoes a conformational change and becomes activated as an inhibitor of thrombin and other serine proteases in the coagulation cascade. A major breakthrough in the study of heparin-catalyzed anticoagulation resulted from the separation of distinct heparin fractions differing markedly in affinity for ATIII. Low-affinity ATIII binding heparin (Kd ≈ 10^-4 M) comprises about two-thirds of porcine intestinal heparin and has low anticoagulant activity (typically <20 U/mg). In contrast, high ATIII affinity heparin (Kd ≈ 10^-8 M) comprises the remaining third of porcine intestinal heparin starting material and has high anticoagulant activity (typically ~200 U/mg). Rosenberg et al. and Lindahl et al. examined the ATIII binding site by performing a partial chemical and enzymatic depolymerization of heparin and then purified the products using affinity chromatography on immobilized ATIII. The isolation of 3-O-sulfatase from human urine, capable of desulfonating 3-O-sulfoglucosamine residues, provided the crucial clue to the structure of the ATIII binding site. NMR studies by several groups proved the presence of the 3-O-sulfoglucoaminoglycosyl residue within ATIII binding site.
pentasaccharide containing this unique 3,6-di-O-sulfo group containing residue substantiated these findings.59 The ATIII binding sequences found in certain heparan sulfates are partially responsible for the blood compatibility of the vascular endothelium.18

After the minimum binding site in heparin for ATIII was described, extensive investigation examined the structure–activity site for interaction with ATIII (Figure 2c).58 Many of these studies utilized selective, partial desulfonation desacetylation, carboxyl group reduction, and removal of specific saccharide residues of an oligosaccharide, derived from heparin, having high affinity for ATIII.63,64 Synthetic carbohydrate chemists have also prepared a number of analogues of the ATIII pentasaccharide binding site and indirectly tested these for their interaction with ATIII by measuring their anticoagulant activity. These heparin analogues include rigid and flexible congeners of uronic acids, replacement of N-sulfo groups with O-sulfo groups, and O-alkylation of hydroxyl group. In addition to a specific positional grouping of critical charge groups, the flexibility of the iduronic acid residues is essential for binding and activation of ATIII.51

The ATIII pentasaccharide is sufficient to catalyze the ATIII-mediated inhibition of factor Xa, a critical serine protease in the coagulation cascade (Figures 1 and 5).2 To catalyze the ATIII-mediated inhibition of thrombin, 16–18 saccharide units are required.63 Thus, the SAR of thrombin inhibition has been more difficult to establish because it relied on the synthesis of substantially larger oligosaccharides for pharmacological evaluation. Recent studies show that a relatively nonspecific but highly charged thrombin binding domain in heparin, on the nonreducing side the heparin's ATIII binding site, is required to form a ternary complex (Figure 5d).63 Success in understanding the SAR of heparin's inhibition of thrombin has resulted in a new class of potent, synthetic, but still experimental thrombin inhibitors.63 Other heparin binding proteins, such as tissue factor pathway inhibitor (TFPI)2 and annexins,64 can also play important roles in anticoagulation.

Low Molecular Weight Heparins

Heparin, has been one of the most effective and widely used drugs of the past century.1,13,52,62 As one of the oldest drugs currently still in widespread clinical use, heparin is unique because it is among the first biopolymeric drugs and one of only a few carbohydrate drugs. Indeed, heparin's introduction predates the establishment of the U.S. Food and Drug Administration.3

Low molecular weight heparins (LMWHs) are a group of heparin-derived anticoagulant/antithrombotic agents, and their development began approximately 25 years ago.21,65 An improved understanding of the coagulation cascade was associated with the development of LMWHs.2,21 Heparin accelerates the inhibition of these coagulation factors by ATIII, preventing the generation of a fibrin clot. In the coagulation cascade, one factor activates the next until prothrombin (factor II) is converted to thrombin (factor IIa) by factor Xa (Figure 1). It is thrombin that acts on fibrinogen to form a fibrin clot. Factor Xa sits at the convergence of the intrinsic and extrinsic activation pathways of the coagulation cascade, suggesting that its regulation might better control blood clotting. This represented a therapeutic opportunity to develop an agent that was more specific than heparin (which acts at many points in the cascade), which might provide better regulation of coagulation, reducing the major hemorrhagic side effects associated with heparin. This rationale was the basis for developing LMWHs. Various laboratories including ours66 observed that when heparin was fractionated on the basis of size or broken down chemically or enzymatically, its activity against thrombin (factor IIa) was decreased to a much greater extent than its activity against factor Xa. This separation of activities results from differences in their molecular requirements for inhibition (Figure 5). The initial rationale for their development, while naive, has led to the successful introduction of LMWHs as new effective and improved anticoagulant/antithrombotic agents.

Pharmaceutical heparin exhibits surprisingly high physical and chemical stability with a shelf life approaching a decade. Under stressed conditions, however, heparin can be decomposed.21 The oxidative instability of heparin, widely observed in heparin manufacture, suggested a method to prepare LMWHs.21 Heparin could be broken down using a variety of oxygen-containing reagents including hydrogen peroxide, hypochlorous acid, Cu+/O2, Fe2+/O2, or ionizing γ-irradiation and oxidative deamination.21 Heparin manufacturers had also observed microbial degradation of heparin. A bacterial enzyme, heparinase, is known to act on heparin through a β-eliminative cleavage mechanism.34,35 This enzymatic reaction can be mimicked chemically by esterifying the carboxyl group of the uronic acid residue and treating the resulting heparin ester with base.21

There are five LMWHs approved for clinical use in the U.S. These are prepared through the chemical or enzymatic depolymerization of heparin and have reduced average molecular weights (3000–7000), with polydispersities ranging from 1.1 to 1.5.67,68 These drugs each exhibit an enhanced antifactor Xa to antifactor IIa activity ratio of 1.5 to 4.21,67 While these LMWHs display similar physical, chemical, and biological properties, a close examination of their structures suggests significant structural and biochemical differences among these products. Detailed analyses show some differences in their disaccharide composition but more importantly reveal significant process-dependent differences in their structures, particularly their end-group residues.67,68

Furthermore, the type of heparin degraded (source and tissue) and the level of purification of the starting material also contribute to the final characteristics of an LMWH. Thus, even with a well-controlled source of heparin, LMWHs derived by a single process still have substantial inherent structural variability. Each LMWH product is defined on the basis of a certain range of specifications associated with the physical, chemical, and biological properties of the drug product set by the manufacturer. Since LMWH is a heterogeneous mixture, these specifications do less to define the final drug product than would a similar set of specifications on a well-defined homogeneous pharmaceutical such as aspirin. The ultimate issue, of course, comes down to how structural differences affect the biological and clinical
properties of LMWHs determined by the pharmacologist and clinician and the biological equivalency of these agents.65,67,68

LMWHs offer a major advantage over heparin associated with their improved pharmacokinetic/pharmacodynamic properties resulting in longer biological half-lives and their enhanced subcutaneous bioavailability.65,67 As a result, LMWH can be administered in subcutaneous injections each day, permitting its clinical use on an out-patient basis.65,67 This has reduced the hospitalization costs normally associated with the intravenous use of heparin. As a result, ~70% of heparin anticoagulant therapy in the U.S. now relies on LMWHs. The goal of reducing heparin’s side effects by using more specific LMWHs with enhanced antifactor Xa activity has only been partially realized.65 LMWHs show a slight reduction in hemorrhagic complications when compared to heparin, including heparin-induced thrombocytopenia (HIT).69 HIT is believed to be associated with antibodies generated against platelet factor

Figure 5. Anticoagulant/antithrombotic activity of heparin. (a) ATIII binds to a specific pentasaccharide sequence in heparin to form (b) a binary complex of ATIII and heparin, resulting in a conformational change in ATIII (this activated ATIII can directly inhibit Factor Xa (FXa) without FXa interaction with heparin). Thrombin (black rectangle) then binds to heparin (c) and slides down the chain to meet ATIII. (d) A ternary complex of thrombin, heparin, and ATIII forms irreversibly, inactivating thrombin. (e) Thrombin covalently linked to ATIII dissociates from heparin, allowing heparin to catalyze the inhibition of a second molecule of thrombin (steps a–d).
by Cardin and Weintraub in 1989, used modeling to structural requirements for GAG recognition (Table 1). The first look at the general increasing numbers of heparin binding proteins, there hundreds of different proteins. With the discovery of sulfate primarily result from their interaction with hydrogen-bonding amino acids bound with highest Peptides enriched in arginine and lysine and polar of spacing of basic amino acids in heparin binding. laboratory, looked at the common amino acids in hep- 4 (PF4) bound to heparin. Smaller LMWH chains with <10 saccharide units show a reduction in HIT.

Heparin–Protein Interactions

The biological activities of heparin and heparan sulfate primarily result from their interaction with hundreds of different proteins. With the discovery of increasing numbers of heparin binding proteins, there was a need to characterize the molecular properties, within the proteins and heparin, responsible for specific recognition (Table 1). The first look at the general structural requirements for GAG–protein interactions, by Cardin and Weintraub in 1989, used modeling to demonstrate that some heparin binding proteins had defined motifs corresponding to consensus sequences. Their results suggested, for example, that if the XBBBXXBXX (B is a basic and X is a hydrophobic amino acid residue) sequence was contained in an α-helical domain, then the basic amino acids would be displayed on one side of the helix with the hydrophobic residues pointing back into the protein core. Heparin binding sites, commonly observed on the external surface of proteins, correspond to shallow pockets of positive charge. Thus, the topology of the heparin binding site is also an important factor in heparin binding consensus sequences. Structural analysis by our laboratory of the heparin binding sites in acidic fibroblast growth factor (FGF-1), basic FGF (FGF-2), and transforming growth factor β-1 (TGF-β1) implicated a TXBBXXTBXXTBB motif (T defines a turn).

Studies using peptide libraries, performed in our laboratory, looked at the common amino acids in heparin binding domains in proteins and the importance of spacing of basic amino acids in heparin binding. Peptides enriched in arginine and lysine and polar hydrogen-bonding amino acids bound with highest affinity. The effect of the pattern and spacing of the basic amino acids in heparin binding sites showed that heparin interacted most tightly with peptides containing a complementary binding site of high positive charge density while the less sulfated heparan sulfate interacted most tightly with a complementary site on a peptide that had more widely spaced basic residues. The FGF family (FGF1-21) is probably the most extensively studied heparin binding proteins. X-ray crystallographic data on FGF complexes with heparin oligosaccharides have been useful in defining the precise groups within heparin that are important for these biologically significant interactions. Both structural and experimental data suggest that the 6-O-sulfate groups within heparin, while not directly important in the binding of FGF-2 to heparin, are apparently required for the mitogenic activity of FGF-2. In contrast, the interaction of FGF-1 with heparin is directly mediated through contacts with the 6-O-sulfo groups on heparin. These differences suggest a specificity of interaction for various members of the FGF family of growth factors. While the interaction between proteins and heparin is primarily ionic and based on the presence and appropriate positioning of sulfo and carboxyl groups, this may not always be the case. For example, in the interaction between secretory leucocyte protease inhibitor (SLPI) or brain natriuretic peptide (BNP) and heparin, undersulfated heparin oligosaccharide sequences are required for specific high-affinity interaction, suggesting the importance of hydrogen-bonding interactions through the hydroxyl groups present in heparin. Heparin–heparin interaction, only a small portion of the binding free energy results from ionic contributions, with the major contribution coming from hydrogen bonding. Hydrophobic forces may also play a minor role in heparin–protein interactions.

Heparin resembles DNA because both are highly charged linear helical polymers that behave as polyelectrolytes. The high repulsive energy of multiple negatively charged groups in these polyelectrolytes promotes cation (i.e., Na+) binding to minimize these forces. The binding of Na+ ions by polyelectrolyte is entropically unfavorable. When a polyelectrolyte such as heparin binds a protein, the positively charged amino acid residues on the protein interact at anionic sites, resulting in the entropically favorable release of Na+ ions. Thus, much of the free energy of interaction of heparin with proteins is derived from the entropically favorable release of Na+ ions.

Molecular modeling can provide useful information on the molecular interactions, but its predictive power still needs to be substantiated. The value of modeling is that intricate dynamic details of molecular level events can be visualized with a relatively small investment of time and cost. Docking studies between heparin oligosaccharides and their protein binding partners can afford additional structural information on the interaction. However, such studies present a challenge because of the weak surface complementarity, the high charge density of heparin and the heparin binding site, and the highly flexible nature of the heparin chain.

Lander put forward the idea that a major role of PGs was to control molecular encounters by limiting heparin binding proteins such as growth factors. Instead of freely diffusing through the three-dimensional space of the extracellular matrix, PGs limit the diffusion of heparin binding proteins to the one-dimensional domain of a GAG chain. When a heparin binding growth factor is guided from the extracellular matrix down to the cell surface, the GAG chain enhances the chance of its productive encounter with its protein-based receptor. Thus, PGs enhance the rate of such encounters without necessarily affecting the thermodynamics of these encounters. This interesting theory suggests that it is not sufficient to simply measure the Kd of binding but also the on-rate and off-rate of these interactions. Our laboratory and others are now widely using surface plasmon resonance to obtain these kinetic data to supplement the thermodynamic data obtained with methods such as isothermal titration calorimetry.
A few examples of heparin binding proteins, responsible for important physiological and pathophysiological processes, can provide insight into some general properties of these interactions.

The FGF family of heparin binding proteins is involved in developmental and physiological processes including cell proliferation, differentiation, angio genesis, and angiogenesis. These FGFs exert their biological effects by binding to different, specific cell surface receptors called FGF receptors (FGFR1-4). The FGFRs are transmembrane tyrosine kinase receptors expressed as multiple splice variants with different affinities for the different FGFs. The FGFRs are also heparin binding proteins; thus, FGF–FGFR–heparan sulfate must interact simultaneously to initiate signal transduction. FGF stimulated signal transduction is similar to other receptor-mediated pathways in that it involves the dimerization of the FGFRs. Cell membrane heparan sulfate binds multiple FGF molecules, promoting FGFR dimerization and signal transduction. High-resolution X-ray crystal structures of FGF–FGFR–heparan dexasaccharide complexes have provided an insight into the stoichiometry and structural aspects of this physiologically relevant interaction. Our results support a dimeric 2:2:2, FGF-2/FGFR-1/(heparin or heparan sulfate) complex (Figure 6) in which heparan sulfate interacts through its nonreducing end with both FGF-2 and FGFR-1. Small heparin mimetics, such as sulfated sucrose derivatives, designed on the basis of X-ray crystallographic data can interfere with the assembly of a complete, functional signal transduction complex. These derivatives act to block signaling, inhibiting cell differentiation, and might provide a novel approach to the preparation of anticancer agents.

Chemokines are a group of small, cytokine-like proteins having a variety of biological functions including selective recruitment and activation of cells during inflammation. The first member of the chemokine family to be discovered was platelet factor 4 (PF4). PF4 is released from platelets and is believed to have a number of properties associated with inflammation and wound healing that are thought to be due to its ability to neutralize the activities of heparin and heparan sulfate PGs. When therapeutically administered heparin binds to PF4, it can lead to a dangerous immunologically induced loss of platelets, HIT. PF4 exists mainly as a tetramer under physiological conditions and binds to heparin and heparan sulfate with a very high affinity in a 1:1 ratio. NMR studies on the interaction of a heparin dodecasaccharide chain with PF-4 suggest that residues in the Arg20–Arg22 loop are involved in heparin binding. These studies also indicate that in certain PF-4/heparin ratios, the heparin chain induces PF-4 to undergo a conformational transition to a partially folded, molten-globule state. The ratio at which this conformational transition occurs is the same ratio at which HIT antibody binding is observed, suggesting that the HIT antibody recognizes a less folded, lower aggregate state of the protein.

Annexins are a family of homologous proteins (with more than 30 members) that are ubiquitous in eukaryotes. A distinctive feature of annexins is their calcium-dependent binding to the surface of phospholipid membranes. Annexins have been implicated in a wide range of functions in eukaryotes, including roles in cell signaling, membrane trafficking, blood coagulation, and inflammation. Although annexins are found primarily within cells and lack signal sequences, many extracellular events are now proposed to be annexin-mediated. While there are many examples of proteins that bind heparin, annexin V is among the few of these that are calcium-dependent. Annexin V also exhibits potent in vitro anticoagulant activity through a proposed mechanism in which it coats placental endothelial cells to form a protective antithrombotic shield. A 1.9 Å crystal structure of annexin V with bound heparin-derived tetrasaccharides from our laboratory gave the first structural information on the essential role of calcium ions in a heparin–protein interaction. On the basis of these data, a model was proposed by which cell surface heparan sulfate PG wraps around the annexin V molecule, making contact with all three of its heparin binding domains. The annexin V molecule is then delivered and assembled by heparan sulfate into an antithrombotic array on the cell surface phospholipid membrane.

The initial binding of a virus to a target cell often represents a critical step in pathogenesis. Binding may result from a receptor-like interaction between a viral coat protein molecule and the GAG chain of a PG expressed on the surface of target cells. Since heparan sulfate GAGs are found on the external surface of most mammalian tissues, it is not surprising that viruses make use of these molecules as receptors to bind to and gain entry into target cells. Herpes simplex virus (HSV), for example, requires the interaction of glycoprotein gD with a specific sequence within HS chains that have been modified by 3-O-sulfonation of specific glucosamine residues to enter into the cell. Dengue virus, a mosquito-borne flavivirus responsible for yellow fever, is also believed to target cells by the interaction of a dengue envelope protein with a highly sulfated, liver heparan sulfate, suggesting an explanation for the tropism of the virus. Two putative heparin binding motifs at the carboxy terminus of the dengue envelope protein sequence have also been identified. Heparin and a number of smaller analogues are being investigated.
as potential pharmaceutical agents to prevent dengue virus infection. Heparan sulfate is also believed to play an important role in the virulence of other pathogens. The circumsporozoite (CS) protein is a sporozoite cell surface protein of Plasmodium falciparum, the parasite causing malaria. The CS protein interacts with the highly sulfated heparan sulfate present on the surface of liver cells possibly by the same mechanism involved in the apolipoprotein E mediated clearance of lipoprotein remnants from the blood by the liver.

Two types of enzymes act on heparin and heparan sulfate: the prokaryotic heparin lyases (heparinase, acting through an eliminative mechanism) and the eukaryotic glucuronol hydrolases (heparanase, acting through a hydrolytic mechanism). Heparinases eliminate heparan sulfate, affording unsaturated oligosaccharide products. Three major heparinases (I, II, and III), isolated from Flavobacterium heparinum, are capable of cleaving linkages present in heparin and heparan sulfate. Heparinases are important clinically and have been used in the monitoring of heparin levels in the blood, in the neutralization of heparin in the blood, and in the production of low molecular weight heparins for use in humans. Heparinases I and III are also potent inhibitors of neovascularization, which is associated with the regulation of tissue development, wound healing and tumor metastasis. Heparanase, an endo-β-D-glucuronidase (hydrolase), is capable of cleaving heparan sulfate and has been purified, cloned, and characterized. Heparanase has been implicated in inflammation and tumor angiogenesis and metastasis, making it an important target for the development of inhibitors.

The interaction of heparin with various proteins that play important roles in the regulation of normal physiologic processes (Table 1) as well as disease states has led to an interest in using heparin in roles outside its normal application as an anticoagulant/antithrombotic agent. Randomized trials to study the effectiveness of LMWH, compared with the effectiveness of heparin in treating venous thromboembolism in cancer patients, led to a surprising observation that treatment with LMWH affects survival of patients with malignancy. Cancer patients who had been treated with LMWH for their thrombosis had a slightly improved 3-month survival compared to cancer patients receiving heparin. LMWH can potentially exert its activity at various stages in cancer progression and malignancy-related processes. It can affect cell proliferation, interfere with the adherence of cancer cells to vascular endothelium, regulate the immune system, and have both inhibitory and stimulatory effects on angiogenesis. There is recent evidence showing that heparin treatment reduces tumor metastasis in mice by inhibiting P-selectin-mediated interactions of platelets with carcinoma cell-surface mucin ligands.

Numerous proteins of a physiologic and pathophysiologic importance bind to heparin and heparan sulfate. This offers a large number of potential therapeutic applications for heparin. The major limitation in utilizing heparin in new ways is that its high potency as an anticoagulant now becomes a side effect that can lead to hemorrhagic complications. The introduction of LMWH and the preparation of heparin oligosaccharides and synthetic analogues devoid of anticoagulant activity may open up a wide variety of new potential therapeutic applications in the treatment of cancer, viral and bacterial infection, and Alzheimer’s disease.

**Synthesis**

Heparin’s significant medical potential makes it the most studied GAG, and numerous syntheses of heparin oligosaccharides have been reported over the past 25 years. Until recently, nearly all the syntheses have focused on pentasaccharide, which is known to bind specifically to AT III. A review article by van Boeckel and Petitou in 1993 provides a thorough overview of the synthesis of various AT III binding pentasaccharide analogues. Recent studies have resulted in a better knowledge of heparin structure—activity relationships and preparation of potent, simplified heparin oligosaccharides analogues such as pentasaccharide. Where the N-sulfo groups are replaced by O-sulfo groups and hydroxyl groups are replaced by O-methyl groups. With the structure responsible for heparin’s anticoagulant activity in hand, researchers focused on the preparation of synthetic agents that displayed both anticoagulant and antithrombotic activities. Heparin’s antithrombotic activity results from the formation of a ternary complex consisting of heparin, AT III, and thrombin, and heparin oligosaccharides having at least 16 saccharide residues are necessary to form this complex. The thrombin binding domain corresponds to a repeating trisulfated heparin disaccharide sequence, IdoA2S-GlcNS6S. Since the interaction of heparin with thrombin was found to be nonspecific, other sulfated linear oligosaccharides were also synthesized for use as the thrombin binding domain. Petitou and co-workers reported the synthesis of hexa- and deca- toicosasaccharide heparin analogues. Hydroxyl groups were alkylated and N-sulfo groups were replaced with O-sulfo groups to simplify the synthesis. Van Boeckel and co-workers have reported the syntheses of various glycoconjugate mimics containing a modified AT III binding pentasaccharide linked to an oligosaccharide, or a non-carbohydrate spacer.

Suda and co-workers reported the synthesis of various sulfonated di- and trisaccharide analogues of the HP regular region for studies of HP binding to platelets using selectively protected disaccharides having IdoA–GlcN and GlcN–IdoA sequences. On the basis of their observation that the number of GlcNS6S–IdoA2S sequences within HP increased binding to platelets, Suda and co-workers synthesized clustered analogues containing more than one disaccharide unit for biological evaluation. The replacement of IdoA with GlcA has been studied to understand heparin binding with FGF by Sinay and co-workers by synthesizing the four possible FGF-2 binding pentasaccharides.

The high structural diversity of heparin and heparan sulfate represents a major challenge for all chemists undertaking their synthesis. Using tailor-made monomers, a modular approach for heparan sulfate synthesis can allow the preparation of all sequence combinations. Seeberger and co-workers designed an automated synthesizer to rapidly prepare heparin analogues from such monomers. Our laboratory is utilizing a chemoenzymatic approach to prepare oligosaccharides.
in which heparin lyase affords unsaturated sulfated disaccharides that are used as synthetic building blocks. Thus, remarkable progress in heparin and heparan sulfate synthesis is on the horizon.

Structure–Activity Relationship

The question that most often arises when discussing heparin’s SAR is whether it really demonstrates sequence specificity in its interaction with proteins. There are actually three issues here: (1) Does heparin have defined sequences? (2) Are there defined sequences required for its interaction with specific proteins? (3) With all the proteins that are capable of interacting with heparin, can a truly specific heparin-based drug ever be prepared? The discovery of the ATIII pentasaccharide binding site and the elucidation of its SAR certainly demonstrate that heparin does have defined sequences within its binding domain that can interact with high specificity and high affinity with selected proteins and that this interaction has been exploited in the development of a highly specific anti-factor Xa agent. Furthermore, other sequences have been discovered that interact with some level of specificity to growth factors, growth factor receptors, and viral envelop proteins.

On the other side of these issues is our seeming inability to completely eliminate “nonspecific” protein interactions from even very well-defined heparin disaccharides. For example, the ATIII pentasaccharide used as an anti-factor Xa agent can also interact with PF4, causing some undesired side effects. When discussing endogenous activity, however, it is important to remember that heparan sulfate, not heparin, is responsible for normal physiologic roles. Not all heparan sulfates contain highly sulfated domains, such as the repeating trisulfated disaccharide (Figure 2a) and the ATIII binding site (Figure 2c) commonly associated with protein interactions. These sequences are highly controlled by the selective expression of specific isoforms of the biosynthetic enzymes. Our current lack of understanding of the control mechanisms in biosynthesis severely limits our understanding of specificity. It is clear, however, that expression is spatially (tissue) restricted, temporally (developmentally) restricted, and responsive to environmental signals (pathological insults).

Arguments have been made that the inherent flexibility of the iduronic acid (Figure 4b) and the flexible positioning of sulfo groups greatly reduce or eliminate heparin’s (or heparan sulfate) specificity. These arguments, however, fail to consider the limited spatial and temporal expression of rare sequences in real biological systems. Thus, despite the fact that many proteins can bind to highly sulfated domains, they may not encounter these domains at the time and location where they are present. Finally, many domains within heparan sulfate may be masked in vivo by already bound proteins or protein complexes that would need to be released prior to exposing a binding site. This adds yet one more level of specificity and control to these interactions.

In conclusion, it is clear that there are specific sequences within heparin and heparan sulfate that play specific physiologic roles. Furthermore, it is also clear that this specificity is not complete nor is it well understood. Simple in vitro assessment of specificity is unlikely to provide a full understanding of specificity. We will certainly need to rely on much more complex biological systems for such an understanding.

Future Directions

The study of heparin and heparan sulfate will certainly extend well into this new century with so many questions left still unanswered. Some focal points in the near future include (1) improved preparation and synthesis of heparins, (2) new heparin anticoagulants with improved properties, (3) new therapeutic applications for heparins, (4) new heparin mimetics, (5) new biomaterials, and (5) development of an improved understanding of physiology and pathophysiology through glycomics.

The preparation of heparin from mammalian tissues obtained at slaughterhouses represents a concern particularly with the recent appearance of bovine spongiform encephalopathy in Europe. Bovine tissues are now rarely used in heparin production, and there are growing concerns about porcine tissues. Heparins prepared by defined, recombinant mammalian cell line capable of being cultured in large-scale fermentations offer an exciting alternative to tissue isolation.

In addition, studies are currently underway in a number of laboratories to use isolated recombinant enzymes to convert heparosan, an E. coli derived capsular polysaccharide (GlcNAC(1,4)GlcA), to heparin. Finally, chemical synthesis continues to improve, offering both natural and unnatural sequences for biological evaluation. Recent advances in the solid-phase and combinatorial synthesis of carbohydrates offer promise in the design of more specific agents by a better understanding of heparin’s SAR.

New anticoagulants based on heparin’s structure might offer enhanced specificity targeting one or selected groups of coagulation proteases and avoiding undesired interactions with other proteins such as PF4, thus decreasing the side effects associated with heparin use. An orally active heparin is considered to be the “holy grail” of this field because it would improve the prophylactic application of heparin in preventing many coagulation problems ranging from cardiovascular diseases and stroke to deep vein thrombosis. Synthetic heparins might incorporate hydrophobic groups, mask or eliminate charged groups, or promote active transport of heparins, providing for their oral activity.

New therapeutic applications might include the application of heparin to treatment of infectious diseases, inflammation, and control of cell growth in wound-healing and cancer. These new activities will require the elimination of heparin’s anticoagulant activity, the engineering of appropriate pharmacokinetics and pharmacodynamics, and optimally oral bioavailability. A concern about the application of heparins to promote wound healing is that they might simultaneously promote cancer. Thus, wound healing applications will probably require localization of the drug at the site of action possibly through the application of polymers or gels.

The development of heparin polymers and the immobilization of heparins will also be important for the production of improved biomaterials for artificial organs.
(both implantable and extracorporeal). Blood compatibility is particularly important in developing indwelling catheters, stents, and artificial blood vessels. All of these will require an improved understanding of endothelial heparan sulfate and its role in maintaining blood fluidity, as well as new chemistry, to develop stable blood-compatible surfaces.

Finally, one of the most important future directions of heparin research is driven by the recent sequencing of the human genome and the field of genomics. While much attention is currently focused on the proteome encoded by the genome and the rapidly developing field of proteomics, the glycome has garnered little attention. Glycomics is the study of the structure and function of carbohydrates, offering new therapeutic approaches to treating these very serious conditions. Moreover, improved knowledge of glycomics should lead to a better understanding of physiology and pathophysiology, offering new approaches to drug development. Recent suggestions that saccharide structures in mammals and microbes coevolved suggest that this improved understanding will lead to new generations of anti-infectives.

In conclusion, while heparin has been under extensive study by many laboratories since its discovery 87 years ago and in our laboratory for the past 20 years, we have only begun to understand its structure, function, and SAR. It is possible that as we enter the 21st century, there will still remain unsolved mysteries regarding this important family of polysaccharides. It is hoped that this article will serve as an enticement for young scientists to join the quest to understand heparin, its chemistry, biology, pharmacology, and therapeutic applications.

Acknowledgment. The author gratefully acknowledges the National Institutes of Health for their generous funding of his research over the past 21 years. Current grants related to heparin studies include grants GM 38060, HL 52622, HL 62244, and AI 50238.

References

Degradation of heparan sulfate by nitrous acid.

Toida, T.; Huang, Y.; Washio, Y.; Maruyama, T.; Toyoda, H.;

Integral glycan sequencing of heparan sulfate

Islam, T.; Butler, M.; Sikkander, S. A.; Toida, T.; Linhardt, R. J.

Mach, H.; Volkin, D. B.; Burke, C. J.; Middaugh, C. R.; Linhardt, R. J.

Patterns of sulphation

Jaques, L. B. Heparin: an old drug with a new paradigm.

Leder, I. G. A novel 3-sulfotransferase specificities.

Strate specificities.

simplex virus 1 entry.

Analysis of Glycoconjugates

Flavobacterium heparinum

protein D.

Chem.

Imanari, T.; Linhardt, R. J. Chemical microdetermination of

Strategy for the Sequence Analysis of Heparin.

Capillary Zone Electrophoresis.

377–74.


Rosenberg, R. D.; Lam, L. Correlation between structure and function of heparin.

Leder, I. G. A novel 3-sulfotransferase specificities.

Strate specificities.

simplex virus 1 entry.

Analysis of Glycoconjugates

Flavobacterium heparinum

protein D.

Chem.

Imanari, T.; Linhardt, R. J. Chemical microdetermination of

Strategy for the Sequence Analysis of Heparin.

Capillary Zone Electrophoresis.

377–74.


Rosenberg, R. D.; Lam, L. Correlation between structure and function of heparin.

Leder, I. G. A novel 3-sulfotransferase specificities.

Strate specificities.


J M030176M