Brief Review

Heparin-Binding Domains in Vascular Biology

Eva M. Muñoz, Robert J. Linhardt

Abstract—Heparin is a major anticoagulant with activity mediated primarily through its interaction with antithrombin (AT). Heparan sulfate (HS), structurally related to heparin, binds a wide range of proteins of different functionality, taking part in various physiological and pathological processes. The heparin–AT complex, the most well understood facet of anticoagulation, serves as a prototypical example of the important role of heparin/HS in vascular biology. Extensive studies have identified common structural features in heparin/HS–binding sites of proteins. These include elucidation of consensus sequences in proteins, patterns of clusters of basic and nonbasic residues, and common spatial arrangements of basic amino acids in the heparin-binding sites. Although these studies have provided valuable information, heparin/HS–binding proteins differ widely in structure. The prediction of heparin/HS–binding proteins from sequence information is not currently possible, and elucidation of protein-binding sites requires the individual study of each glycosaminoglycan–protein complex. Thus, x-ray crystallography and site-directed mutagenesis experiments are among the most powerful tools, providing accurate structural information, facilitating the characterization of heparin–protein complexes. (Arterioscler Thromb Vasc Biol. 2004;24:1549-1557.)

Key Words: heparin-binding domains • glycosaminoglycan • basic amino acids • consensus sequence • antithrombin

Heparin and low molecular weight heparins are the most widely used anticoagulants in treatment of thromboembolic disease. Early workers showed that a plasma protein was responsible for the anticoagulant effect of heparin. In 1973, Rosenberg and Damus suggested that heparin binds the protease inhibitor antithrombin (AT), causing a conformational change within AT, accelerating its reaction with the protease thrombin and the formation of an active complex between protease and inhibitor. The interaction of heparin with AT was the first reported case of specific heparin–protein interaction of physiological significance. During the past 2 decades, a growing number of physiological and pathological processes, such as inflammatory response, immune cell migration and inflammation, tumor cell metastasis, and smooth muscle cell (SMC) proliferation, have been shown to be regulated by the action of heparin. In most of these biological functions, interaction of heparin with protein is essential. Proteases, growth factors, chemokines, lipid-binding proteins, and pathogen proteins are among the numerous proteins that bind heparin (Table). An understanding of heparin–protein interactions at the molecular level is fundamental to elucidate the role of heparin in these biological processes and, consequently, for development of new therapeutic agents. Unfortunately, the structure activity relationship in heparin–protein interactions is still not well understood. Among the reasons that make this task so difficult is that heparin is a polydisperse polysaccharide with a heterogeneous saccharide sequence that binds a large number of proteins. Consequently, there is a wide range of possible binding sites along the heparin chain as well as a diversity of heparin-binding sites among the numerous proteins that bind this glycosaminoglycan (GAG). Additionally, most heparin-binding proteins also bind heparan sulfate (HS). Whereas heparin is primarily intracellular, the HS proteoglycan (HSPG) is a common constituent of cell surfaces and the extracellular matrix (ECM) and is also involved in a wide range of biological functions. Although heparin and HS are structurally related, subtle differences in their saccharide sequence make it more difficult to determine common features in heparin–protein interactions.

Heparin is a linear, polydisperse polysaccharide consisting of repeating units of 1→4-linked pyranosyluronic acid and 2-amino-2-deoxyglucopyranose (glucosamine) residues. The uronic acid residues typically consist of 90% L-idopyranosyluronic acid (L-iduronic acid) and 10% D-glucopyranosyluronic acid (D-glucuronic acid). Structural variations of this disaccharide exist, leading to heparin microheterogeneity. The amino group of the glucosamine residue may be substituted with an acetyl or sulfo group or unsubstituted. The 3 and 6 positions of the glucosamine residues can either be substituted with an O-sulfo group or unsubstituted. The 3 and 6 positions of the glucosamine residues can either be substituted with an O-sulfo group or unsubstituted. The uronic acid, which can either be L-iduronic or D-glucuronic acid, may also contain a 2-O-sulfo group (Figure 1A). HS is also a linear copolymer of uronic acid 1→4 linked to glucosamine but has a more varied structure. D-glucuronic acid predominates in HS.
although substantial amounts of L-iduronic acid can be present. Additionally, HS is much less substituted in sulfo groups than heparin (Figure 1A).

Heparin has an average of 2.7 negative charges per disaccharide (2 negative charges per disaccharide for HS) provided by sulfo and carboxyl groups and, consequently, the most prominent type of interaction between heparin (or HS) and proteins is ionic. In fact, heparin-binding sites in proteins are characterized by the presence of clusters of positively charged basic amino acids that form ion pairs with spatially defined negatively charged sulfo or carboxyl groups of the GAG chain. Nonelectrostatic interactions such as hydrogen bonding and hydrophobic interactions can also contribute to the stability of heparin–protein complexes. The highly anionic nature of heparin has led people to believe that its binding with proteins is nonspecific in nature. However, elucidation of the structure activity relationship of the binding between heparin and AT has demonstrated that heparin can have defined sequences within its binding domain that interact with high affinity with its target proteins in a specific manner (Figure 1B). Furthermore, other sequences have been discovered that interact with some level of specificity to growth factors, and viral envelope proteins. These studies show that heparin–protein interactions depend on the defined patterns and orientations of the sulfo and carboxyl groups along the polysaccharide sequence in the polymer. A correct pattern of basic amino acids in the heparin-binding domain of the proteins is also necessary to ensure the appropriate affinity and specificity of the complex. The question arises whether heparin-binding sites share structural features such as analogous amino acid sequences or similar spatial arrangement of the basic residues.

### Common Structural Features in Heparin-Binding Domains

Electrostatic interactions play a major role in the binding of heparin to proteins, and basic amino acids such as arginine and lysine are present in the heparin-binding sites of most proteins. A number of studies have been undertaken to determine whether there is a consensus sequence of basic amino acids arranged in a specific way in heparin-binding sites. Cardin and Weintraub first compared the heparin-binding site amino acid sequence of 4 proteins: apolipoprotein B, apolipoprotein E, vitronectin, and platelet factor 4. They found that these regions are characterized by 2 consensus sequences of amino acids: XBBXBX and XBBBXXBX,
important role in heparin–protein interactions. Among them, serine and glycine have been found to be the most frequent nonbasic residues in heparin-binding peptides.\textsuperscript{28} Both have small side chains, providing minimal steric constraints and good flexibility for peptide interaction with GAG.

Elucidation of Heparin-Binding Domains

The studies on consensus sequences provide information on the structural features within heparin-binding sites that are important for GAG interaction and have facilitated the design of peptides that bind heparin efficiently.\textsuperscript{29} These consensus sequences are also useful in the absence of a 3D protein structure. To identify the critical amino acids in heparin-binding proteins and the essential functional groups in heparin interaction, the 3D structure of the heparin-binding site needs to be established. X-ray crystallography provides valuable high-resolution 3D structural information about the complexes. So far, the structures of heparin complexes with very different types of proteins have been described: AT, annexin V, fibroblast growth factors (FGFs) and their receptors, biosynthetic enzymes, and virus capsid proteins.\textsuperscript{30} In the crystal structure, proximity of basic amino acids to negatively charged groups in heparin permits the definition of the residues involved in binding. Nuclear magnetic resonance (NMR) spectroscopy has also provided solution structures containing comparable information on growth factors.\textsuperscript{31} A polypeptide E,\textsuperscript{32} adhesion molecules,\textsuperscript{33} and chemokines.\textsuperscript{34,33} Unfortunately, x-ray crystallography and NMR spectroscopy do not provide information on the relative importance of each residue in the binding. Site-directed mutagenesis of the wild-type protein is often required to obtain this information. Binding studies and measurement of biological activity of wild-type and mutant forms can enable detailed characterization of heparin-binding sites on proteins. Information regarding the structure, affinity, kinetics, and thermodynamics can be obtained through the combination of different techniques.\textsuperscript{5} Among them, affinity chromatography is the most commonly used technique in studying the binding affinity of heparin–protein interactions. Surface plasmon resonance is a powerful technique that has made it possible to study the kinetics of GAG–protein interaction, whereas isothermal titration calorimetry provides thermodynamic data. Using the combination of these and other techniques, a large number of heparin-binding sites have been studied,\textsuperscript{11,35–52} including growth factors,\textsuperscript{36,44,46,47,51} chemokines,\textsuperscript{52} and proteins involved in the anticoagulant system.\textsuperscript{35,40,41,43,48–50} To date, the best-characterized protein–heparin complex is the AT–heparin complex.\textsuperscript{53}

Heparin-Binding Domain of AT

Heparin functions as an anticoagulant primarily through activation of AT-mediated inhibition of blood coagulation factors such as thrombin and factor Xa. Several steps are involved in the interaction of heparin with AT and serine proteases. First, a low-affinity interaction between GAG and AT takes place, mediated by a well-defined unique pentasaccharide sequence within heparin (Figure 1B). This binding generates a conformational change in the structure of AT, which enables additional interactions between AT and hepa-
rin, resulting in stronger binding. The conformational change also expels a protease reactive loop in AT.54,55 A ternary complex is formed, after which the AT interaction reverts to low-affinity binding, resulting in the release of heparin from the covalent AT–protease complex.

The structure of the heparin-binding site in AT was mapped initially through the chemical modification of specific residues, resulting in decreased heparin affinity and by studying natural recombinant mutations associated in AT.56–64 Subsequently, the crystal structure of AT in complex with the heparin pentasaccharide was used to identify residues in AT involved in its interaction with heparin.65 In the N-terminus of the protein, A helix, and D helix, along with a loop on the N-terminal side of the D helix, are folded together to form a highly basic pentasaccharide-binding site. The basic residues in this site that interact with the pentasaccharide are Lys 11 and Arg 13 in the N-terminal end; Arg 46 and Arg 47 in the A helix; and Lys 114, Lys 125, and Arg 129 in the region of the D-helix and its N-terminal loop (Figure 2).

Additionally, an extended site outside the pentasaccharide-binding region composed of Arg 132, Lys 133, and Lys 136 in the C-terminal elongated region on the D-helix interacts electrostatically, although it is not required for AT activation.60,65,66 Comparison of the structure of the AT–pentasaccharide complex with that of the free protein shows that the pentasaccharide promotes important conformational changes in the inhibitor.67 These changes provide the correct orientation of the basic residues, leading to a stronger interaction with heparin and the expulsion of the protease reactive loop.68–70 Mutagenesis of the amino acids has been performed to elucidate the role of each residue in the binding. Three amino acids, Lys 114, Lys 125, and Arg 129, were found to be of greatest importance for pentasaccharide binding (Figure 2B).71–73 Mutation of any of these residues results in an important loss of binding energy, promoted by the loss of multiple ionic as well as nonionic interactions.71–73 This observation suggests that these residues act in a cooperative fashion, establishing a network of electrostatic interactions between the amino acids that mediate the interaction of AT and the pentasaccharide. In contrast, Arg 47, Arg 46, Lys 11, and Arg 13 contribute a smaller fraction of the binding energy (Figure 2C). Individual mutations of these residues result in a loss of binding energy, primarily accounted for by the loss of a single ionic interaction.74,75 An additional residue in the N-terminal region, Arg 24, is apparently located too far from heparin in the complex to establish any direct interaction (Figure 2D).65 However, mutation of this amino acid reduces AT affinity for the pentasaccharide.75 It has been proposed that Arg 24 stabilizes the AT–pentasaccharide complex by forming intramolecular ionic and nonionic interactions with Glu 113 and Asp 117, located in the pentasaccharide-induced P-helix. Analysis of the kinetics of pentasaccharide binding to wild-type and mutant forms of AT has enabled a dissection of the relative contributions of the different pentasaccharide binding residues to the 2 steps. Of all residues examined thus far, only Lys 125 makes a substantial contribution to the

Figure 2. A, Crystal structure of the AT–pentasaccharide complex (from Protein Data Bank65). The pentasaccharide is marked in red. The basic residues of AT involved in interaction with the heparin pentasaccharide are shown in B, C, and D. B, Lys 114, Lys 125, and Arg 129. C, Arg 47, Arg 46, Lys 11, and Arg 13. D, Arg 24, Phe 121, and Phe 122.
binding energy of the initial, low-affinity AT-pentasaccharide formed in the first step.\textsuperscript{76} Lys 11 has been also shown to be involved in the initial recognition of the pentasaccharide but to a lesser extent.\textsuperscript{75} Indeed, Lys 125 and Lys 11 are directed toward the same region in the x-ray crystal structure of the free AT and interact with the same carboxylate group in the pentasaccharide.\textsuperscript{65,67} The other basic residues do not participate in the initial binding of pentasaccharide to AT but contribute their binding energy primarily in the second step involving the conformational change.\textsuperscript{66,74,75}

Structural studies on heparin–AT complex have been focused primarily on basic residues in the heparin-binding site. These amino acids participate in 5 to 6 ionic interactions, contributing 40% of the binding energy.\textsuperscript{55,77,78} Nonionic interactions are responsible for the remaining 60% of the binding energy. Two nonbasic residues, Phe 121 and Phe 122 (Figure 2D), reside near these positively charged amino acids of the heparin-binding domain but make minimal direct contact with the pentasaccharide sequence.\textsuperscript{11} However, mutation of either Phe 121 or Phe 122 resulted in an important decrease of AT–heparin-binding affinity. Thus, nonionic effects associated with Phe 121 and Phe 122 appear to play a critical role in heparin binding and AT activation. More detailed molecular studies will be required to fully understand the interaction of heparin with AT.

### Heparin-Binding Domain of Other Vascular/Plasma Proteins

In addition to the crystal structure of AT in complex with heparin, high-resolution x-ray crystal structures have been obtained for a number of heparin-binding proteins during the past decade.\textsuperscript{13,15,16,79–82} Additionally, NMR titration studies have provided detailed information on heparin-binding sites.\textsuperscript{33,34,83–85} These binding sites lie on the surface of the proteins and are rich in basic amino acids that interact with heparin through ion pairing and hydrogen bonding. Spatial proximity occurs among the basic residues. Nevertheless, proximity in amino acid sequence is not required, as illustrated by the heparin-binding sites of AT.\textsuperscript{65} Lymphotactin,\textsuperscript{34} or foot and mouth disease virus proteins.\textsuperscript{79} Structural similarities have been found among the same family of proteins such as the growth factors,\textsuperscript{84} but this is not a general feature. Recently, chemokines such as lymphotactin have been pointed out as examples of the diversity of GAG-binding sites, where GAG-binding site specificity varies among family members.\textsuperscript{34} Differences in binding site structure represent a mechanism to confer specificity to heparin–protein interactions, essential for the proper function of GAGs in vascular biology.

The analysis of the conformation of free- and protein-bound saccharide residues can provide an important understanding of conformational changes within heparin that take place during heparin–protein interaction. Energy calculations suggest that the 2,6-disulfoglucosamine residue of the free GAG will occupy a $C_3$ conformation (indicating that carbon 4 is above the plane defined by carbons 2, 3, and 5 and the ring oxygen, and that carbon 1 is below the plane) and that there are 2 equienergetic conformations of the 2-sulfoiduronic acid residue: the $C_3$ chair and the $S_3$ skew boat.\textsuperscript{36,87} Solution NMR conformational studies suggested that the 2-sulfoiduronic acid residue prefers a $S_3$ conformation,\textsuperscript{88} whereas the 2,6-sulfoglucosamine residue resides in a $C_3$ conformation. In the crystal structure of FGF-2 bound to a heparin tetrasaccharide, the 2-sulfoiduronic acid residue makes contact with the protein and resides in a $C_3$ chair.\textsuperscript{13} In the annexin V–heparin tetrasaccharide complex, the 2-sulfoiduronic acid residue interacting with the protein resides in the $S_3$ conformer, whereas the noninteracting 2-sulfoiduronic acid residue resides in the $C_3$ chair.\textsuperscript{89} The foot and mouth disease virus oligosaccharide receptor\textsuperscript{79} and the 3-O-sulfotransferase-1-heparin oligosaccharide\textsuperscript{82} complexes show an interacting 2-sulfoiduronic acid residue in the $S_3$ conformation. In all the crystal structures reported to date, the bound 2,6-disulfoglucosamine residue resides in the $C_3$ conformation. These data suggest that when a protein binds to heparin, it induces a change in conformation of the 2-sulfoiduronic acid residue, resulting in a better fit and enhanced binding, but the conformation of the less flexible 2,6-disulfoglucosamine residue remains unaltered.

### Role of AT and Other Heparin-Binding Proteins in Vascular Biology

HS is biosynthesized as a proteoglycan through the same pathway as heparin;\textsuperscript{89} however, unlike heparin, the HS GAG chain remains connected to its core protein. Heparin is localized to the granules of mast cells and is only released locally in the allergic response. HSPG is extracellular, residing in the ECM and on the surface of all animal cells.\textsuperscript{80} The vessel wall is composed of a monolayer of endothelial cells (ECs) below which lie SMCs. Both ECs and SMCs contain membrane-associated HSPGs.\textsuperscript{91,92} HS on the endothelium is known to contain AT-binding sites.\textsuperscript{90} Interestingly, knockout mice, missing 3-OST-1 responsible for placement of the critical 3-O-sulfo group in the AT-binding pentasaccharide, do not show an abnormal coagulation phenotype.\textsuperscript{91} There are several possible explanations for the failure to observe this phenotype. First, AT interaction with 3-O-sulfo group containing HS might not be critical for homeostasis. This is unlikely because AT requires 3-O-sulfo group for tight binding, and AT deficiencies or mutations result in coagulation abnormalities. Second, overexpression of AT or alternative anticoagulant mechanisms might compensate for the loss of 3-O-sulfo group in HS. Third, another 3-OST isoform, such as 3-OST-5, which can synthesize the same sequence, might replace the lost function of 3-OST-1.\textsuperscript{94} Although the issue remains to be solved, it appears that the final explanation is the most plausible.

HSPGs play a complex role in the ECM, regulating a wide variety of biological processes, including coagulation, inflammation, angiogenesis, growth factors, cell adhesion, and others. In vascular biology, blood coagulation is triggered through enzymatically governed proteolytic processes. The reaction principally consists of a successive activation of proenzymes to proteolytic enzymes. Among them, thrombin is the key protease in coagulation. This clotting enzyme not only exhibits coagulant activities such as fibrinogen–fibrin conversion and activation of factors V, VIII, and XIII, but it also initiates the anticoagulant pathway via activation of...
protein C with the aid of thrombomodulin on the EC surface. Furthermore, it initiates several cellular and vascular reactions, platelet aggregation, the EC release reaction, and proliferation, as well as the contraction or dilatation of the vessel wall. Figure 3A schematically shows the regulation of thrombin coagulation function through the action of GAGs. HSPGs on the endothelium are known to contain AT-binding sites occupied by AT in an intact vascular wall (Figure 3A). When thrombin comes in contact with the AT-HSPG complex on the “healthy” intact vascular wall, thrombin is irreversibly inactivated as the covalent thrombin–AT (T-AT) complex (as described previously), thus preventing clot formation (Fbng →Fbn). However, SMCs are believed to contain an HSPG without an AT-binding site. When the vessel wall is wounded (Figure 3B), the endothelium containing HSPG with AT-binding sites is lost. Thus, T catalyzes fibrinogen (Fbng)→fibrin (Fbn) clot formation promoting coagulation. Inhibition of T by HC mediated by HSPGs also takes place in the wounded vessel, avoiding subluminal coagulation. C. Chemokines (Cx) are involved in the recruitment of lymphocytes that occur on inflammation. Cx form stable gradients through interaction with HSPG on the endothelium.

Conclusions

Heparin and HS bind a wide range of proteins of different structure and functionality, playing an important role in a variety of biological processes. The idea of common structural features among heparin-binding proteins has emerged, and efforts have been made to define consensus sequences or spatial arrangements of amino acids among heparin-binding domains. These studies have led to the understanding of GAG–protein interactions and have facilitated the search for potential GAG–protein-binding sites. Despite some common
characteristics, heparin-binding domains show no absolute dependency on specific sequences, making them difficult to predict from genomic data. The elucidation of heparin-binding domains still requires the in-depth study of each individual GAG–protein complex. Our improved understanding of the interactions on the molecular level should improve our understanding of normal and pathological vascular physiology as well as aid in design of new therapeutic agents.

References


