Heparan sulfate and heparin interactions with proteins

Maria C. Z. Meneghetti1, Ashley J. Hughes2,3, Timothy R. Rudd4,3, Helena B. Nader1, Andrew K. Powell5,3, Edwin A. Yates3,1 and Marcelo A. Lima1,3

1Departamento de Bioquímica, Universidade Federal de São Paulo (UNIFESP), Rua Três de Maio, São Paulo 40440-020, Brazil
2Department of Chemistry and Molecular Biology, University of Gothenburg, Gothenburg 40530, Sweden
3Department of Biochemistry, Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK
4The National Institute for Biological Standards and Control (NIBSC), South Mimms, Potters Bar, Hertfordshire EN6 3QG, UK
5School of Pharmacy and Biomolecular Science, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK

Heparan sulfate (HS) polysaccharides are ubiquitous components of the cell surface and extracellular matrix of all multicellular animals, whereas heparin is present within mast cells and can be viewed as a more sulfated, tissue-specific, HS variant. HS and heparin regulate biological processes through interactions with a large repertoire of proteins. Owing to these interactions and diverse effects observed during in vitro, ex vivo and in vivo experiments, manifold biological/pharmacological activities have been attributed to them. The properties that have been thought to bestow protein binding and biological activity upon HS and heparin vary from high levels of sequence specificity to a dependence on charge. In contrast to these opposing opinions, we will argue that the evidence supports both a level of redundancy and a degree of selectivity in the structure–activity relationship. The relationship between this apparent redundancy, the multi-dentate nature of heparin and HS polysaccharide chains, their involvement in protein networks and the multiple binding sites on proteins, each possessing different properties, will also be considered. Finally, the role of cations in modulating HS/heparin activity will be reviewed and some of the implications for structure–activity relationships and regulation will be discussed.

1. Introduction
1.1. Heparan sulfate and heparin similarities and differences
Heparan sulfate (HS) polysaccharides are a family of linear sulfated, heterogeneous polysaccharides found on the cell membrane and in the extracellular matrix as part of heparan sulfate proteoglycans (HSPGs). They are composed of repeating 1→4 linked disaccharide units, in which one monosaccharide is a α-D-glucosamine residue and the other an uronic acid (in a salt form—an uronate; figure 1). Heparin is a structurally similar polysaccharide found within mast cells as a component of serglycin proteoglycans and has been shown to differ in composition with several mammalian forms of HS [6]. Under the conventional definition, HS and heparin can be compared as follows: first, in heparin, the uronates are predominantly α-L-iduronate, whereas in HS, they are mainly its C-5 epimer, β-D-glucuronate. Second, in HS, the β-glucosamine residues are predominantly N-acetylated, whereas in heparin, they are N-sulfated. Finally, whereas at least 70–80% of heparin is composed of the disaccharide L-iduronate 2-O-sulfate α(1→4) β-D-glucosamine N,6-sulfate, in HS around 40–60% of the disaccharides consist of α-D-glucuronate β (1→4) D-glucosamine (1→, that can be either N-acetylated or N-sulfated). Together, these structural characteristics make heparin more sulfated and, hence, more charged than HS [6–8]. Furthermore, HS also has a much higher maximum average molecular weight (ca 50 kDa) than heparin (ca 20 kDa) [9]. The differences in underlying composition, as well
as between substitution pattern and domain structure, which could also lead to distinct glycosidic linkage geometry (torsion angles, phi and psi) and iduronate conformational equilibria (e.g., $\text{C}_2^\text{g}$, $\text{D}_2^\text{g}$, etc.) do, according to the conventional definitions, make HS fundamentally different to heparin. Furthermore, HS is considered typically to adopt longer, more flexible chains in solution than heparin [10,11]. It has become apparent, however, that the designations heparin or HS are less clear-cut than this description implies, and that polysaccharides isolated from some organisms appear to be hybrid [12,13]. Indeed, HS with features approaching heparin are produced in a variety of tissues [14–16] and cell types [17–19]. It is possible, therefore, to consider heparin a tissue-specific form of HS with the conventional definitions of heparin and HS residing towards two opposite extremes. More detailed studies across many species and tissues will provide a more complete picture of the extent of sulfation and structural diversity throughout the animal kingdom and a distinction from ‘heparin’, the finished pharmaceutical agent, must also be considered.

### 1.2. Heparin sulfate/heparin structural diversity from biosynthesis

The biosynthesis of both heparin and HS, as the carbohydrate components of proteoglycans, starts in the endoplasmic reticulum and proceeds through the Golgi apparatus where several polymerases and modifying enzymes and their respective isoforms may act on different substrates, or sets of substrates. The biosynthetic process and its machinery have been studied extensively [20] and a sequential hierarchical order of events, in which selective enzymatic modifications are required to allow further editing of the polymer, are widely accepted. While this model does provide for some of the possible substitution patterns observed in both heparin and HS, several disaccharides and their respective substitution patterns cannot be explained within the scheme, for example IdoA-GlcNAc6S, if the proposed order of events and their known specificities are respected rigidly [21]. Furthermore, unlike nucleic acids and proteins, HS and heparin biosynthesis is not template driven and the conventional experimental approaches of gene knockout and single enzyme inhibition do not provide conclusive answers, because compensatory mechanisms can be activated [22,23]. Further understanding of how the biosynthetic machinery functions, particularly at the sequence level, transcends the approaches cited above and the way in which such process are controlled is therefore still sought [20]. The process of biosynthesis results in the introduction of structural properties that can be viewed at a number of levels; the disaccharide composition, the order of disaccharides (sequence), the clustering of similar disaccharides within the chains (domains) and stereochemical variation arising from distinct monosaccharide composition, such as $\text{D}$-glucuronic acid and $\text{L}$-iduronic acid, all of which combine to bestow particular conformational, charge distribution and flexibility on a given sequence and, ultimately, define their biological function. It is worth emphasizing that structure, defined simply as saccharide sequence, has not been found to correspond to function in a simple manner.

### 1.3. Structural variation in heparan sulfate and heparin

Dynamic temporal and spatial variations in HS structure underpin the concept of the ‘heparanome’ [24]. Owing to the fact that HS occurs as an element of proteoglycans, the primary source of structural variation could arise from changes in proteoglycan expression. Nonetheless, changes in HS structural characteristics are thought to be more tissue specific rather than depending on the nature of its constituent proteoglycans [25]. One could assume that the overall HS assembly in similar tissues and/or cells are comparable but, such an assumption ignores this temporal variation within biological microenvironments. Substantial changes in both HS quantity and overall assembly in response to variation in environment have been shown in different cell cycle stages [26–29] and, in the case of heparin, shown to stimulate antithrombotic HS production by endothelial cells [17]. Heparin and HS can interact with a large number of proteins, which are being identified [30–32]. The poly- and oligosaccharide chains can interact with a very large number of proteins, which are being identified [33,34], and are defined as heparin binding proteins (HBPs) usually employing heparin as an experimental proxy for HS, and are known as HBPs. Ultimately, changes in HS structure may be particular to protein binding capabilities that are required for specific biological functions at different moments.

### 2. The aims and scope of the review

#### 2.1. Studies of heparan sulfate and heparin interacting with proteins encompass two broad fields of research: the biochemical and the physico-chemical

There are two broad areas of research that relate to the interactions of HS/heparin with proteins. One area employs...
biochemical approaches, including protein assays, in vitro cell assays as well as in vivo and ex vivo perturbations which, only relatively rarely, incorporate studies of the effects of changes to the complex chemical environment in which the polysaccharides can be found, such as pH, ionic strength and types of associated cation. The other area covers physicochemical techniques, including studies into the polyelectrolyte nature of the polysaccharides and the involvement of cation binding to HS/heparin chains, but only rarely investigates biological activity. As has been observed aptly by Kayitmazer et al. [30], there is often little communication between researchers in these two fields and those involved may employ very different experimental techniques. Debate regarding selectivity reflects not only this, but also the perspective, strategy and technical approaches employed. Thus, what may appear peripheral for one group may be considered ‘mainstream’ for the other.

Rather than attempt to review exhaustively the extensive biological literature in the heparin/HS field, which has been undertaken successfully elsewhere [3,31,35–42], the aim of this review is to focus on the findings from both biochemical approaches and physico-chemical studies, regarding HS/heparin polysaccharides and the relationship between sequence, structure and activity. It is not intended to restrict the review to recent articles, because the work in this area has been accumulating for decades and debate continues. Hence, relevant articles from across the breadth of the literature will be considered. A summary of the selected works, illustrated with examples from the most extensively studied systems, particularly relating to antithrombin (AT) and the fibroblast growth factor (FGF) protein family, will be made and a brief synthesis (§5), intended to be accessible to a wide audience and providing findings from both perspectives, will be attempted. It is hoped that the review will help to promote discussion regarding HS/heparin–protein interactions between researchers in both of these fields, and more widely, without particular favour towards either audience. Finally, aspects to which, it would seem, future efforts might be usefully directed will be suggested (§5.3).

2.2. The contrasting structure of oligo- and polysaccharides

Heparin and HS are expressed primarily as the polysaccharide component of HSPGs such as syndecans and glypicans in the case of HS, and serglycin in the case of heparin. They can also be subject to enzymatic cleavage to release free glycosaminoglycan (GAG) chains of varying lengths. The challenge has been to delineate the HS/heparin sequences that carry activity for a given protein and, eventually, a particular biological function. It has become apparent that activity in this system does not reside in unique HS/heparin sequences but, in a limited choice of potential binding sequences. Subtle relationships prevail between sulfation position and uronic acid identity on the one hand, and activity on the other hand. Because HS is relatively difficult to isolate from natural systems in substantial quantities, it is usually necessary to employ oligosaccharide fragments (§3.1) in isolated systems, so that binding properties to proteins can be interpreted yet, as intimated by the section above, there is evidence that oligosaccharides and polysaccharides can exhibit distinct properties [43–45]. Furthermore, the equivalence of the conformations that short sequences excised from the polysaccharide can adopt, compared with the intact parental polysaccharide [10,11] and the likely lack of domain structure and thermodynamic properties of the former, can be questioned. The manifold binding sites, which are characteristic of the polysaccharides, have usually been reduced substantially in oligosaccharides. These caveats are of less concern, however, if the aim is to obtain active structures for potential application as pharmaceutical agents, and it may be useful to distinguish this activity from attempts to understand the biological system per se, in which the polysaccharide component of HSPGs is usually involved. These two distinct aims can sometimes be conflated, one example being the frequently stated experimental aim of identifying the ‘minimal active sequence or fragment’. Having said this, the concentration of circulating oligosaccharide fragments derived from HS chains after heparanase degradation may be increased in certain circumstances, such as in the disease condition, mucopolysaccharidosis [46], may play a role in others [47] and disease outcome has been linked to heparanase in cancer [48,49].

3. The nature of specificity in heparan sulfate/protein interactions

An important contribution to the binding of heparin/HS polysaccharides to proteins involves binding of the negatively charged GAG to the amino acid residues lysine and arginine, and can also include protonated histidine residues at low pH values [50], as well as interactions with a range of other amino acids. A thorough analysis of the distribution and nature of these binding regions on the surfaces of the several hundred HS binding proteins so far identified will be interesting and warrants further investigation. Studies, which include the identification of interacting arginine side-chains considered to be stronger binders [51], will also be interesting to compare. Some of the data regarding binding specificity between heparin/HS and proteins were obtained using X-ray crystallography, in which small sugar fragments, usually heparin oligosaccharide fractions (obtained by size-exclusion chromatography and probably containing several different sequences) served as substitutes for the physiologically relevant polysaccharides. It must also be recalled that protein structural analysis by X-ray crystallography is subject inevitably to a number of limitations, despite its very high precision in measuring atomic position. These limitations include structural packing in which conformational artefacts can be created, incomplete solvation, which contrasts the normal physiological conditions, and the fact that crystallography cannot usually be employed to study protein–polysaccharide complexes nor, in many cases, analyse glycosylated and structurally disordered proteins. Moreover, heparin-binding processes are dynamic and protein structures obtained from crystallography can be viewed as ‘frozen’ in a particular conformation, potentially reporting conformations that may not represent the physiological situation.

3.1. Experimental tools: polysaccharides and oligosaccharides

The nature and selectivity of heparin and HS–protein interactions is the subject of widespread debate. Unfortunately, the experiments that could determine unequivocally the origins of specificity cannot be conducted at present for several reasons, which include the need to test vast numbers of structural variants in their pure forms [52]. It is the ability
to separate these species in sufficient quantities when using natural HS as a source, despite some progress [53–56], rather than the problem of sequencing them per se, which hinders progress. Heparin is often employed experimentally as a proxy for HS on account of its structural similarity (§1) and is considerably more highly charged than HS. Its purification process, intended to achieve high anticoagulant potency also favours chains with certain characteristics, leading to a more homogeneous mixture than most HS, tending to activate in some systems more extensively than tissue-derived HS where the active entity may be a lower proportion of the more heterogeneous mixture. For this reason, and also because they offer a route by which the relationship between substitution pattern, (in terms of the prevailing statistical content of O-, N-sulfate and N-acetyl groups), with activity in polysaccharides, rather than oligosaccharides (§2.2) can be explored, chemically modified heparin derivatives have been used [57–59]. Using these in their polysaccharide forms, or as a means of generating oligosaccharides following enzyme digestion, there have been attempts to examine the structure–activity relationship of heparin/HS [59–62]. Oligosaccharides can be selected, by affinity chromatography or filter trapping with salt elution, thereby tending to select high charge interactions [63–66]. Searches are often made among sugar sequences representative of only a very limited subset of the total possible sequences because, usually, the available sugars derive either from HS of a particular tissue, or from heparin, both of which have limited structural diversity. Early work attempted to identify unique sequences responsible for protein binding, again employing affinity chromatography, and eluting with a salt gradient (e.g. in the case of FGF-2) [67] although it was realized that other sequences could exert an effect. Interpreting these results as the case of FGF-2) [67] although it was realized that other sequences could exert an effect. Furthermore, non-carbohydrate compounds can bind both the pentasaccharide binding site and the extended heparin binding sites of AT, leading to enhanced anticoagulant activity [79], suggesting that alternative classes of compounds may find therapeutic roles.

The structure of AT in the heparin-bound, intermediate state illustrates such issues. It was reported that it was difficult to conclude that the form of AT obtained represented the kinetic intermediate identified from solution based experiments [80]. A recent X-ray crystallographic study of the solvent accessible areas of AT and thrombin, suggested that, contrary to what had long been held to be the case, the nature of heparin binding to the two proteins is fundamentally different; the polysaccharide ligand being much less strongly held in the case of thrombin. This analysis was based on the differences in mobility of the two main amino acid side chains responsible for heparin binding (Arg and Lys) [81]. More recently, it has been shown that the reactive centre loop, exposed upon heparin-binding, and considered to act as ‘bait’ for proteases, can populate conformational states in which the Arg side chain is exposed to the solvent [82].

Furthermore, such a sequence is yet to be found within HS—perhaps one of the natural ligands—and it is worth noting that, under normal physiological conditions, there may be very little, or no, circulating heparin, hence it is also important to distinguish biological anticoagulation from the pharmacological actions of anticoagulants. While the pentasaccharide sequence undoubtedly binds AT with high affinity and activates the protein, subsequent evidence has emerged suggesting that net charge plays a significant role in the affinity of heparin for AT (§4.1) and that the 3-O-sulfated group in the central glucosamine unit of the heparin binding pentasaccharide is not essential to activate AT [72,73]. In fact, other types of carbohydrate structures, distinct from GAGs, including those with low or no activity, have also been identified that can fulfil the structural requirements of AT binding [72] and a proposal has been made that the stabilization of AT is the key determinant of its activity [73]. It has also been shown that N-acetylation or N-sulfation is permitted in the substitution pattern of residues adjacent to and within, the first glucosamine residue, the pentasaccharide and that specific residues outside of the pentasaccharide region active for AT can influence AT activity [74–78]. This reinforces the idea that binding, even with low affinity, does not necessarily equate to activity. Furthermore, non-carbohydrate compounds can bind both the pentasaccharide binding site and the extended heparin binding sites of AT, leading to enhanced anticoagulant activity [79], suggesting that alternative classes of compounds may find therapeutic roles.

The structure of AT in the heparin-bound, intermediate state illustrates such issues. It was reported that it was difficult to conclude that the form of AT obtained represented the kinetic intermediate identified from solution based experiments [80]. A recent X-ray crystallographic study of the solvent accessible areas of AT and thrombin, suggested that, contrary to what had long been held to be the case, the nature of heparin binding to the two proteins is fundamentally different; the polysaccharide ligand being much less strongly held in the case of thrombin. This analysis was based on the differences in mobility of the two main amino acid side chains responsible for heparin binding (Arg and Lys) [81]. More recently, it has been shown that the reactive centre loop, exposed upon heparin-binding, and considered to act as ‘bait’ for proteases, can populate conformational states in which the Arg side chain is exposed to the solvent [82].

3.2. Antithrombin-heparin pentasaccharide: an exception that does not prove the rule?

The early work on elucidating the relationship between HS/heparin structure with protein binding and activity concentrated on analysis of the heparin sequence capable of binding and activating the serine protease inhibitor AT. This approach employed the fractionation of heparin followed by affinity chromatography, revealing a pentasaccharide that bound a particular sequence in heparin with a high degree of selectivity over other sequences, while retaining its anticoagulant activity. The pentasaccharide was first suggested as that possessing highest affinity under the experimental conditions employed (elution in high salt), which seem likely to have been selective for highly charged species [69,70]. Despite these reservations, which were expressed in the original reports, over time, the interactions with AT and the pentasaccharide sequence within the heparin polysaccharide have tended to be viewed by some [71] as the unique binding structure redolent of the ‘lock and key’ hypothesis and, furthermore, one that was prototypical of all heparin and HS interactions with proteins in general. A more nuanced view is now prevailing (vide infra).

3.3. Heparan sulfate/heparin interactions with fibroblast growth factors: representative of typical protein–heparan sulfate/heparin interactions?

The other well-studied class of HBP’s that interact primarily with HS, on the cell surface and in the extracellular space, are the FGFs and their cognate receptor tyrosine kinases (FGFRs). This family of signalling proteins forms complexes in which an FGF and an FGFR, or FGFRs (the nature of the signalling complex in vivo remains unclear) are brought together to form functioning signalling complexes in the presence of HS or heparin polysaccharides or oligosaccharides, acting as obligatory co-receptors. From the perspective of FGFR specificity, it was found that FGFR1 and FGFR2 had distinct kinetics and affinities for heparin, despite the apparent involvement of the same sulfate groups in heparin [83]. It is generally held that the naturally occurring co-receptor is HS but, heparin and its derivatives or fragments, and even...
analogous non-GAG structures are able to elicit both signalling and inhibitory activity [84].

In the literature relating to the HS/heparin sequences required for binding to FGFs, there are several claims to have uncovered ‘minimal binding sequences’ or similar, and it is important to assess these carefully. In most cases, the oligosaccharides identified originate from parent polysaccharides of rather restricted sequence diversity and this should temper the claims made. Furthermore, scrutiny of the methods employed to select bound saccharides reveals that they may rely on some form of salt elution, which could be biased towards those interactions involving charges and may be more likely to return highly charged species preferentially.

The employment of such a series of modified heparin derivatives, in which sequence diversity has been reduced by systematic desulfation, has also shown that no simple relationship exists between sulfation and charge distribution and their ability to support signalling with FGF-1 and -2 [84,43]. Furthermore, distinct carbohydrate structures can sometimes fulfill the role of HS/heparin, as has been shown for CS/DS extracted from brittlestars, which activated FGF-2 [85]. However, as was the case for AT, FGF-1 signalling correlated strongly with protein stabilization induced by the polysaccharides, whereas FGF-2 signalling did not [79]. For the FGFs, this complex situation is underpinned by the detection of widely ranging heparin binding properties among FGF subfamily members, which possess varied numbers of HBPs, identified by Lys residues (the chemistry of Arg residues proving thus far more problematic), that ranged from 1 to 3 (FGF-9) and their binding properties (kd values)

due to proving thus far more problematic), that ranged from 1 to 3 (FGF-9) and their binding properties (kd values from 38 to 620 nM and kass varied over 20-fold) [86]. Different sources of HSPGs are known to exert distinct activities [87], and HS is both spatially and temporally regulated, controlling FGF signalling [88]. Furthermore, it has been reported that the structural requirements for FGF activation may be more relaxed than those required for FGF binding [89] and, that the position of an interacting sequence within the HS chain can also be important [90].

An alternative route to structurally varied fragments was provided by oligosaccharides obtained from chemically modified heparin. In one case, random distributions of sulfate groups were generated in heparin by chemical means, and the polysaccharides partially depolymerized using enzymes to generate deliberately as many sequences as possible and chains of varying lengths. The sized pools were then screened for a biological activity (not merely binding) in increasing order of their sizes, the first pool capable of activating was then selected and further fractionated by strong anion exchange chromatography and screened again, this time starting from the first eluted fractions, corresponding to the least charged oligosaccharides. In this way, the smallest and least charged oligosaccharides capable of activating were identified and characterized, revealing that in principle, oligosaccharides with very low levels of sulfation were capable of activating FGF-1 and FGFR2c [91]. The implications of this in biological systems have yet to be explored in detail.

Compared with the FGFs, the interactions involving their receptors have been relatively little studied, but the structural basis of interactions between either FGF and HS, or FGFR and HS is distinct [92,93] and glycosylation of the FGFR is known to alter its activity [94]. The interactions between FGF-2 and heparin-derived oligosaccharides (in the case of a hexasaccharide) have been studied by crystallography, suggesting interactions between asparagusine and lysine (Asp 28, Asp 102, Lys 27, Lys 126, Lys 136) and glutamine (Gln 135) residues with the oligosaccharide [95]. Several complexes of FGF and FGFR receptor (FGFR) have also been reported, for example, incorporating heparin-derived oligosaccharides in which the oligosaccharide contacts both FGF-1 units of the dimeric FGF–FGFR complex but, only one of the two FGFRs [96]. In another model [97], the complex has a stoichiometry, the heparin contacting both FGF and FGFR and the 6-O-sulfate reportedly playing a key role.

It is established that binding of HS or heparin to FGF or FGFR alone does not correspond with the activity of the signalling complex [84]. The formation of the ternary complex bestows a set of structural constraints that are likely to be more selective than those imposed by the binding of the polysaccharide to either the FGF or FGFR alone. Indeed, selecting structures in terms of their activity, which involves interactions with both an FGF and its receptor [98], has allowed some of the structural requirements of activation of FGF and its receptors by HS to be addressed [43,99]. More recently, it has been shown that, for FGF-1, even the GAG polarity may be essential for HS–FGF binding and signalling [100].

Tissue-derived HS chromatographic fractions and more recently, several synthetic and defined oligosaccharides [101–103] have also been generated, with the purpose of exploring structure–function relationships [4,12,43,64,66,104–109]. However, it is conceivable that the use of such fractions may misrepresent activities performed by stretches of the full-length polysaccharide.

3.4. Interactions with protein networks

Heparin and HS are also endowed with a variety of functions regulated by their distribution in tissues [110] and facilitate a complex HS ‘interactome’ [33,34], the interactome of a molecule is the whole set of interactions a molecule is involved in. The HS interactome may have a wide influence on biological processes because, while HS interacts with a large variety of proteins [34,111], these HS-binding proteins themselves interact with a plethora of others, forming complex protein–protein interaction networks, in a way that their respective activities within a given network will change (figure 2). AT, for instance, when bound to heparin will have its interaction with factor II and X enhanced (figure 2a). Furthermore, once bound to HS, FGF-1 will bind to its cell-surface receptors, FGFR1 and others, eliciting a given cellular response (figure 2b). On the other hand, the function of biologically available heparin, in contrast to that supplied pharmaceutically, is not clear, and seems most likely to be linked to inflammatory responses when released by mast cells during inflammatory events [113]. Presumably, heparin will be available for protein binding to an interactome of comparable complexity, at least in the circulatory and lymphatic systems, as well as in tissues into which it can diffuse. Because HS is structurally diverse, its expression is regulated and linked to functional outcomes and the relationship between its structure and function is widely debated. Generally, there is little evidence of simple correlations between the sequence of particular disaccharide units and a given activity, and several perspectives have been presented in previous reviews [35–41,114]. The evidence suggests overwhelmingly that several saccharides with different disaccharide compositions, hence distinct sequences, can generate geometry and charge distributions appropriate to elicit similar biological effects. This
implies that the requirement to bind proteins can be fairly relaxed but, it should be emphasized, by no means do all sequences behave similarly. Taking this into account and considering their potential sequence diversity, whose interactions are determined ultimately by the three-dimensional architecture arising from that sequence, the huge information content seemingly available in heparin and HS chains may be lower than previously thought [115]. Furthermore, individual HS/heparin structures are capable of interacting with more than one protein [116,117], which has led to the proposal that the relationship between heparin/HS structure and activity should be viewed at the level of their interactions with multiple proteins [116,118,119]. It remains an open question as to whether all heparin/HS sequences interact to some degree with proteins and whether synergy effects are relevant. If this is the case, the analysis of interactions throughout networks will be essential.

The nature of HBP-interaction networks in disease states has recently been addressed. Two studies concentrated on attempting to establish the network of proteins interacting with HS/heparin in angiogenesis. The first was based on an analysis of the affinity and kinetics [120], and the second on an analysis of the supposed HS sulfation requirements of proteins involved in angiogenesis and construction of a connectivity map [121]. In another approach, the mRNA expression of HBPs in various stages of pancreatic disease were used to identify relevant HBP networks, which exhibited distinct characteristics for the various disease states [32]. In other work, microarray analysis of proteins involved in breast cancer showed that the expression of 105 proteins from a total of 1357 genes were influenced by heparin and these related to tumourigenicity [122].

4. Physico-chemical approaches

4.1. Fundamentals of polysaccharide–protein interactions: the polysaccharide as polyanion

One approach to understanding interactions between polysaccharides and proteins is to explore the fundamental properties of the species involved. This is the basis of the approach adopted by Dubin, Seyrek and co-workers, who have applied concepts from polymer and polyelectrolyte chemistry to their analyses [123] as a means of examining the interaction between the highly negatively charged polysaccharides and their protein binding partners, considered as charges in a given space. For example, interaction between AT and heparin bound on an affinity column exhibited a continuous, broad maximum in the apparent binding, rather than showing any abrupt changes (indicative of ‘high’ and ‘low’ affinity fractions), as theionic strength and pH required for elution from an affinity column was varied, suggesting relatively low specificity [124,125]. A degree of complementarity between the pattern of charges on the protein and those of the heparin was suggested as explaining selectivity [120]. In a study of the interaction between a polyanionic sulfated model compound and ß-lactoglobulin, employing capillary electrophoresis and small angle neutron scattering, increasing the chain flexibility of the saccharide was found to increase binding when the overall negative charge of the polymer dominated [124]. A similar property was uncovered relating to interactions between heparin and BSA [126] and is analogous to the effect of ionic strength on proteins binding to DNA [127]. Treating the interacting species as polyanions demonstrated that the overall attractive and repulsive forces participating in an interaction were important to consider. These depended on polymer chain length and charge density but the ionic conditions in the buffer were also critical [126]. In the case of AT, a low affinity binding fraction from heparin was found to contain low charge species [125] and Kd values, which peaked at particular salt concentrations, were very similar for heparin derivatives of similar charge [128]. Furthermore, a correlation between the charge density of both low molecular weight heparin and dermatan sulfate with apparent affinity was noted, leading to the suggestion that there are multiple binding sites on the polysaccharide chain for AT [125] and arguing against exquisite specificity. In a development of this theme, longer range interactions, which had hitherto been thought to be of relatively minor importance, were then considered in conjunction with modelling studies, supporting the idea that the strongest binding occurs when the heparin/HS polyanion sequences are able
to bind the protein in such a way as to optimize both attractive and repulsive electrostatic forces, hence, minimize global energy. Furthermore, simulations showed that there remained some conformational freedom in the bound carbohydrate chain, even when bound with high affinity, and there was evidence of multiple binding sites with different affinities [123]. It is important to attempt to emulate the ionic conditions that pertain physiologically in such investigations because it was observed that, employing ITC [129], only 30% of binding free energy was provided by charge–charge interactions.

4.2. Cation binding to heparan sulfate/heparin and its effects

The polyanionic nature of heparin and HS, arising from carboxylate groups, O- and N-sulfates, dictates that these polysaccharides must be associated with a large number of cations in solution. In evolutionary terms, it has been suggested that one of the first functions for such molecules may have been in metal ion sequestration and transfer [130]. Measuring the concentrations of cations in cells and tissues is not only extremely difficult, particularly given the likely variations over small distances, their temporal variation and regulation, as well as cell compartmentalization, but is compounded by the dearth of techniques able to make such measurements. Despite these difficulties, some information has emerged. It is known that physiological cation concentrations can affect heparin/HS activity significantly [131] and can vary, for example, for K+ from 3.0 mM extracellularly to 150 mM intracellularly which may be relevant if HS can gain entry into the cell, as has been suggested [132]. However, it is likely that, for example in the vicinity of ion channels, considerable variation in local concentration could prevail.

There are several examples of the effects of bound cations on the activity of GAGs. For example, the ratio of AT III/thrombin activation varied according to the identity of the cation, when present at 100 mM [133] and rarer cations, such as Cu2+ have also been found to increase during tumourigenesis and angiogenesis [134,135]. Studies of the structural effects of cation binding to heparin/HS saccharides using NMR include their influence on heparin [136], which indicated delocalized, long-range interactions with carboxylate groups for Na+, Mg2+ and Ca2+ at low pH values that, for Na+ and Mg2+, were also maintained at higher pH values. However, it has been noted [137] that not all heparin–cation interactions may be described in terms of relatively simple charge-density considerations, as described by Manning [138], because diffusion rates of cations were shown to be sensitive to the concentration of Na+ ions [139] and heparin exhibited an apparently lower charge density than predicted [140]. Obvious questions concerning the observation of altered properties and activities of heparin/HS with cations include whether such effects arise from the consequences of changing the charge distribution and/or the resulting altered conformation and flexibility. Detailed studies of small oligosaccharide model compounds have been conducted, in which a specific binding site was proposed between glucosamine and idurionate residues for Ca2+ ion binding, stiffening the molecule. However, for Na+ and Mg2+, no preferred binding site was apparent [141]. These findings agree with observations made of a heparin derivative (2-de-O-sulfated), in which the binding of divalent cations stiffened the heparin chain considerably, deduced from altered NMR T2 relaxation measurements. Binding of Ca2+ also changed substantially the conformation across glycosidic linkages, as well as the equilibrium of chair and skew-boat forms of the idurionate residues [142]. The effects of cation binding on conformation and dynamics have also been analysed with heparin-derived oligosaccharides possessing different substitution patterns, which exhibited considerable anisotropy and distinct internal motions. Furthermore, the recognition of HS by phage display antibodies can be altered by exposure to cations [143]. Altering the predominant cation from Na+ to Ca2+ induced conformational changes in the 2-O-sulfated iduronic acid residue, which was transferred to the conformation of the adjacent glucosamine moiety [144]. It is worth noting, however, that the validity of extrapolating results from small molecules to large polyanions has been questioned [125].

The consequences of altering the cation form of heparin/HS on biological activity can also be substantial, which has been known for some time in the FGF/FGFR signalling field [142,145]. Changes in signalling ability can be induced by cations bound to heparin/HS analogues. For example, for one chemically modified heparin analogue of HS, conversion of the predominant cation to Cu2+ employing cation exchange resin resulted in a complex active in FGF-2/R1c signalling [142], while changing the cation from Na+ to Ca2+ of another analogue, converted it from signalling to inhibitory [146], and involved a structural change at the level of FGF–HS interaction as measured in solution by circular dichroism spectroscopy. It has also been shown that Ca2+ is an essential activator of a heparin degrading enzyme, heparginase I, in which the heparin–Ca2+ complex is the true enzyme substrate, whereas Ca2+-free heparin is a competitive inhibitor [147].

Annexin V, an abundant anticoagulant and phospholipid binding protein binds HS in a calcium ion-dependent manner, binding via two binding sites on opposite faces of the protein [148,149]. The phospholipid binding and calcium-dependent annexin I has been reported as requiring N- and 2-O-sulfate groups for binding [150]. The influence of cations can also vary. For example, heparin stimulates the activation of human plasminogen (Pγ) by tissue-type Pγ activator and in this, HS differs from heparin, while the effect of the polysaccharide depends on the ionic strength, and can range from stimulatory to inhibitory [151]. A final example of the involvement of cations is the prion–protein interaction with low molecular weight heparin and HS, which occurs through the formation of oligomeric complexes stabilized by Cu(II) bridges. At low pH, this interaction involves protonated histidine residues but, at higher pH values the GAG binds the histidine-bound Cu(II) ion [50].

Understandably, owing to lack of sensitive and accessible instrumentation for the detection of cations in biological environments, the influence of cations is frequently ignored. Nevertheless, the examples highlighted indicate that cation–heparin/HS interactions need to be considered, if possible, when dissecting biological activities. A final consideration is the possible effect of cation binding on the protein. One protein, which has several pairs of acidic residues exposed on its surface, with distinct spacings, suitable for cation sequestration evident from the crystallographic model, is lysozyme [152]. Again, whether these examples can be extrapolated to physiological conditions is not known. Nonetheless, accumulating in vitro data suggest that it may be so, but conclusive answers may only become available after further technological development.
4.3. Other physico-chemical approaches
A combination of other physical chemical techniques including polarized light microscopy, reflection anisotropy and terahertz (10^{12} Hz) absorbance spectroscopies have been used to examine the properties of physiologically relevant cationic forms of heparin (Na^{+}, Ca^{2+}, Mg^{2+}, Cu^{2+}, Zn^{2+}), showing that heparin chains in several cation forms (100 mg ml^{-1}, in water) do not appear to exhibit substantial molecular ordering. Furthermore, heparin samples showed considerable absorption of terahertz radiation in contrast to neutral dextran water (100 mg ml^{-1}, in water) do not appear to exhibit substantial molecular ordering. The interactions between heparin and Cu^{2+} ions in particular stand out for their high degree of regioselectivity within the heparin chain. A combination of NMR, FTIR and EPR spectroscopies showed that the ion binds preferentially between the iduronate residue and the adjacent 6-O-sulfated glucosamine, adopting tetragonal coordination involving the carboxylic acid, the 6-O-sulfate, the ring oxygen of the iduronate and the glycosidic oxygen [154].

5. Concluding remarks: a degree of redundancy and selectivity
5.1. Summary of findings
The findings regarding the relationship between structure and activity can be summarized and a synthesis of the findings from the biochemical evidence and the physico-chemical approaches may be attempted. There are observations showing that several diverse carbohydrate scaffolds can support activity with particular proteins and that protein binding does not equate to activity directly. Indeed, inactive [155] and inhibitory HS structures have been described for some activities [107,146] but, neither binding nor activity depends on sequences and charge alone [107,146], although charge undoubtedly plays an important role bringing the interacting species into proximity [156,157]. However, activity does not depend on the presence of particular ‘key sulfates’ [158] yet, paradoxically, there is a sizeable and presumably energetically costly biosynthesis apparatus for HS and heparin [159], the regulation of which is currently unclear.

5.2. What happens as polysaccharide and protein approach each other?
At long range, heparin/HS represents a highly charged target for protein binding, that attracts the positively charged binding sites of proteins via lysine and arginine (as well as potentially, under appropriate pH conditions, histidine). As such, different HS binding proteins can bind with relatively low affinity to common sequences and translocate across HS chains [116]. As the distance between the two interacting species decreases, however, a degree of charge and shape complementarity may come into play, strong binding resulting from optimization of attractive and repulsive forces through compatible charge distribution and shape complementarity, leading to more protein-specific, sustained binding [116,160]. Some conformational change is usually evident in the protein and, although not often studied, also seems likely in the polysaccharide. The role of solvent water and the thermodynamic consequences of releasing solvated cations may also be important and would warrant further study. An additional consideration that may be expected to affect protein binding is the existence of multiple polysaccharide chains on, or near, the cell surface, whether they originate from the same HSPG, or separate HSPGs. The importance of multivalency in cooperative glycan–protein interactions has also been demonstrated [161].

Understandably, the emphasis has usually been on studying interactions between heparin/HS and individual proteins in isolated systems. However, it has been shown that heparin/HS interacts with hundreds of proteins, many of them associated with the extracellular matrix and involved in intercellular signalling [34]. According to most of the observations reported to date, there is little evidence for a high degree of sequence specificity, rather, support for selectivity with some redundancy, yet biosynthetic machinery comprising around 20 enzymes exists for HS/heparin biosynthesis. Ultimately, structure drives function, yet, ‘structure’ cannot be simplified to primary composition, sequence and substitution patterns. Perhaps, as for proteins, several of the binding modes may be explained by three-dimensional architecture in which—for HS/heparin—several sequences can lead to similar topological chemical entities; hence, the considerable redundancy observed.

5.3. A link to cation concentration and transport?
A key question in HS/heparin–protein interactions that remains largely unanswered concerns the role of cations, which can alter the activity of the molecule. If cations are not linked to the regulation of HS biosynthesis in some way, but simply encounter HS/heparin and these encounters modify activity essentially at random, then the nature of heparin/HS interactions with proteins seems even more loosely regulated. To begin to resolve this, it may be necessary to consider the interplay between the network of proteins interacting with HS/heparin and their relation to both the biosynthetic machinery and mechanisms for the regulation of cations. An appreciation of the effects of structural modification to HS/heparin on this network may reveal that particular regions of the protein network are affected. What limited work has taken place in this area has established the significant differences in the effect of calcium, lowering the affinity of heparin for thrombin [162,133] and among the signalling network of the FGF–FGFR system with heparin as the polysaccharide cofactor [118,119].

These ideas do go some way to explaining how, when an agent such as heparin is added to the body, despite (or perhaps because of) the hundreds of potential interactions that are possible in the circulatory and lymph systems and elsewhere following diffusion into tissues, very few side-effects are observed. Notwithstanding, the relatively low number of cases of such problems as heparin-induced thrombocytopenia, that can occur with prolonged heparin administration, the potentially adverse effects of adding pharmaceutical heparin may be absorbed largely by the network of proteins with which it interacts. It will be interesting to see whether particular ‘zonal regulation’ of the signalling network takes place, when it is perturbed by addition of HS, or its modification,
or whether a simple altering of the signalling intensity throughout the network is observed.

Authors’ contributions. M.A.L., A.K.F., T.R.R. and E.A.Y. conceived the review, whereas all the authors contributed to research, writing and editing.

Competing interests. We declare have no competing interests.

Funding. We received no funding for this study.

Acknowledgements. Support from the CAPES (E.A.Y., M.A.L. and H.B.N.) and FAPESP (M.A.L. and H.B.N.) is gratefully acknowledged. M.C.Z. is recipient of a PhD Fellowship from CNPq.


64. Guerini M, Guglielmi S, Beccati D, Torri G, Viskov C, Mourier P. 2006 Conformational transitions induced
in heparin octasaccharides by binding with anti-thrombin III. J. Biol. Chem. 399, 191–198. (doi:10.1042/BJ20060056)


85. Rudd TR et al. 2010 Comparable stabilisation, structural changes and activities can be induced in FGF by a variety of HS and non-GAG analogues: implications for sequence-activity relationships. Org. Biomol. Chem. 8, 5390–5397. (doi:10.1039/c0ob00246a)


