Glycoconj J (2014) 31:161–169
DOI 10.1007/s10719-013-9512-8

Molecular docking of heparin oligosaccharides with Hep-II heparin-binding domain of fibronectin reveals an interplay between the different positions of sulfate groups

Mathieu Carpentier · Agnès Denys · Fabrice Allain · Gérard Vergoten

Received: 3 September 2013 / Revised: 4 November 2013 / Accepted: 5 November 2013 / Published online: 19 November 2013

Abstract Fibronectin is a major component of the extracellular matrix and serves as support for cell adhesion and migration. Heparin and heparan sulfates (HS) have been reported to be high-affinity ligands for fibronectin. The strongest heparin/HS-binding site, named Hep-II, is located in the C-terminal repeat units FN12-14 of fibronectin. Mutational studies of recombinant fibronectin fragments and elucidation of the X-ray crystallographic structure of Hep-II in complex with heparin allowed localizing the main heparin/HS-binding site in FN13 to two parallel amino acid clusters: R1697, R1698, R1700 and R1714, R1716, R1745. Heparin, which is more sulfated than HS, is a better ligand for fibronectin, indicating that the sulfate density is important for the interactions. However, other studies demonstrated that the position of sulfate groups is also critical for high-affinity binding of the polysaccharides to fibronectin. In the current work, we used molecular docking of Hep-II domain of fibronectin with a series of differently sulfated dodecasaccharides of heparin to determine the implication of each sulfate position in the interaction. By using this approach, we confirmed the implication of R1697, R1698, R1700 and R1714 and we identified other amino acids possibly involved in the interaction. We also confirmed a hierarchic involvement of sulfate position as follows: 2S >> 6S > NS. Interestingly, the formation of stable complexes required a mutual adaptation between Hep-II domain and oligosaccharides, which was different according to the pattern of sulfation. Finally, we demonstrated that 3-*O*-sulfation of heparin stabilized even more the complex with Hep-II by creating new molecular interactions. Collectively, our models point out the complexity of the molecular interactions between heparin/HS and fibronectin.

Keywords Molecular docking · Fibronectin · Heparin · Heparan sulfate · Interaction

M. Carpentier (✉) · A. Denys · F. Allain · G. Vergoten
The Unité de Glycobiologie Structurale et Fonctionnelle, Unité Mixte de Recherche No. 8576 du CNRS, Institut de Recherche Fédératif No.147, Université Lille Nord de France - Université des Sciences & Technologies de Lille, 59655 Villeneuve d'Ascq Cedex, France
e-mail: mathieu.carpentier@univ-lille1.fr

Introduction

Fibronectin is a major component of the extracellular matrix (ECM), in which it is implicated in various physiological processes, including morphogenesis, tissue hemostasis or thrombosis, by providing a support for cell adhesion and migration (for review:[1]). These properties are dependent on the ability to create high-affinity interactions with a number of molecular partners. Thus, fibronectin interacts with transmembrane receptor proteins of the integrin family, which leads to linkage to cytoskeleton and assembly of focal adhesion. Additionally, cell surface heparan sulfates (HS) have been identified as necessary co-receptors for fibronectin [2].

Fibronectin exists either in a soluble dimeric form or an insoluble, fibrillar and multimeric form [3]. One monomer is composed by the repeat of modular segments of 40–90 amino acids, termed type I, II and III repeat units [4, 5]. From the N- to the C-terminal extremity, the polypeptide chain contains six type I units (FN_I1–FN_I6), two type II units, three type I units (FN_I7–FN_I9), 14 type III units (FN1–FN14), one type III unit containing a variable region (V), another type III unit (FN15) and finally, three type I units and a cysteine-rich region. These modules are arranged into functional binding domains along the length of a fibronectin monomer. There are four fibronectin-binding domains, allowing fibronectin to polymerize in the ECM. The type III unit FN10, which contains an RGD sequence, is critically involved in the binding to $\alpha_5\beta_1$

integrin [5], while the type III variable region contains two binding sequences for $\alpha_4\beta_1$ and $\alpha_4\beta_7$ integrins [6, 7]. A third site recognized by $\alpha_4\beta_1$ integrin is present in FN14 as part of the "heparin"-binding site Hep-II [7, 8].

Hep-II domain, which is located to FN12-FN13-FN14 units, is the strongest heparin/HS-binding site within fibronectin. It is strategically sandwiched between FN10 and type III variable region [9], which contains the integrin-binding sequences. A second heparin/HS-binding domain (Hep-I) is located at the N-terminal part and contains the first five repeats FN_I1–FN_I5 [10]. The affinity of isolated Hep-I domain for heparin/HS is 10–25 fold lower than that of Hep-II domain [11]. However, study using atomic force microscopy suggested that Hep-I and Hep-II domains may have comparable binding affinity for heparin in the whole fibronectin molecule [12]. Two secondary heparin/HS-binding sites have been identified within the type III variable region and in the type III repeat FN1 [13, 14].

The structural basis of HS/heparin backbone is a disaccharide unit, comprising a glucosamine residue (GlcNAc) and a glucuronic acid (GlcA). The fine structure of HS/heparin is however rendered more complex by the positioning and density of sulfate groups within the polymers. Glucosamines can be subject to *N*–deacetylation/*N*–sulfation, GlcA can be C5–epimerised into iduronic acid (IdUA) and *O*–sulfation may occur at C2 of the uronic acid (mostly IdUA), and at C6 of the glucosamine residue. Most rarely, *O*-sulfation may also occur at position 3 of GlcNS and $GlcNH_2$ units. Importantly, disaccharide unit sulfation does not occur uniformly along the HS chain, but is mostly restricted to specialised regions of the polysaccharide. Consequently, HS features a unique molecular organisation, with homogeneous, non or low sulfated regions (NA-domains) alternating with hypervariable, highly sulfated domains (NS-domains), bordered by short transition zones with intermediate sulfation [15]. Heparin, which is an analog of HS predominantly containing IdUA2S-GlcNS6S tri-sulfated disaccharide units, has been widely used as a structural surrogate of NS domains of HS to characterize interactions with protein ligands [16].

HS are involved in the interactions with a large panel of proteins, for which they modulate the biological functions by affecting their activation level and bioavailability, inducing formation of signalling complexes or providing a scaffold for the assembly of active molecular complexes [16, 17]. Although binding generally occurs between negatively charged groups in HS and positively charged amino acids residues in the protein ligands, accumulating data demonstrate that the HS large interactive properties are directly linked to its molecular diversity and complex structural organisation. This structural diversity is biosynthetically imprinted in a non template-driven manner and may be dynamically remodeled as cellular function changes [18]. Thus, modifications within HS result in highly heterogeneous molecules exhibiting variable length, sulfation pattern and domain organization, making HS as one of the most challenging biopolymer with regard to structural analysis and interactive properties [19].

A study with large recombinant fragments containing the Hep-II domain allowed localizing the main heparin-binding site to FN13 unit [20] and, in particular, to two parallel basic clusters: R1697, R1698, R1700 and R1714, R1716, R1745 [21]. X-ray crystallographic analysis of Hep-II is in agreement with the conformational organization of FN13 unit in two basic clusters and pointed out another possible heparin-binding site in FN14 unit [8]. However, these findings were only partially confirmed by NMR analysis [22]. This last study demonstrated that the binding reaction between a fragment corresponding to FN13-FN14 units and heparin-derived octasaccharides involves R1697, R1698, R1700 and R1714, while neither the two other arginine residues of the second basic cluster nor the FN14 unit were required for the interaction. This discrepancy may be due to the sizes of the binding partners, which are probably not large enough to mimic the interaction between heparin and Hep-II domain of fibronectin. Indeed, peptides derived from FN14 unit in Hep-II domain were reported to weakly bind to heparin at physiological ionic strength [23]. Moreover, Lyon *et al.* [24] demonstrated that octasaccharide (dp8) is the minimal length unit capable of interacting with Hep-II domain, but high-affinity binding was only observed with heparin-derived oligosaccharides from dp14. Even though all sulfate positions of heparin may make a contribution to Hep-II binding, there may be a hierarchy of involvement with 2S ≫ 6S > NS [24]. Surprisingly, the presence of *N*-sulfated glucosamines was reported to be essential for interaction between cell surface HS and Hep-II, whereas 2-*O*-sulfo groups were not significantly required [25]. The reasons for the difference in sulfate position requirements may be a preference of Hep-II domain for a particular pattern of sulfation, which may lead to different involvement of sulfo groups between heparin and HS.

In the current work, we have used molecular docking of Hep-II domain of fibronectin with a series of differently sulfated dodecasaccharides of heparin to determine the involvement of sulfate position in the interaction. Besides confirming previously published information on the structure of Hep-II domain in complex with heparin, we demonstrated that 3-*O*-sulfation was effective to potently stabilize the interactions. Moreover, our models suggest that the formation of stable complexes required a mutual adaptation between Hep II domain and oligosaccharides, which was different according to the pattern of sulfation.

Material and methods

In silico molecular docking

The structure files corresponding to the standard dodecasaccharide ($[IdUA2S\text{-}GlcNS6S]_6$) of heparin (1HPN)

and to the Hep-II domain of fibronectin (1FNH) were used for docking experiments. Heparin and derived oligosaccharides adopt a right-handed helical shape and a typical heparin disaccharide is 8–8.7 Å length. Contrary to what is mainly observed in pharmaceutical chemistry, where docking of a relatively small molecule occurs within a cleft inside a protein, sulfated interactions between sulfated glycosaminoglycan and protein usually take place at the protein surface. Due to possible intrinsic flexibility of the glycosaminoglycan ligand, the location of the interacting region within the intermolecular complex has to be carefully determined. Consequently, we decided to use a two-step procedure involving successive rigid and flexible docking techniques. To this end, rigid docking experiments were performed in a first step by using the different softwares GRAMM (Global Range Molecular Matching) [26], ESCHER [27], BiGGER [28] and HEX [29, 30]. Among the different solutions proposed by these programs, a zone containing R98, R99, R101 and R115 of Hep-II domain was retained because of a consensus in the results. The selected solution for the docking of Hep-II domain with the standard oligosaccharide of heparin is presented in Fig. 1. In a second step, the R zone was subsequently used for flexible docking experiments using the GOLD program [31]. A spherical binding site of 15 Å radius centered on the atom CA of R99 was used. Side chains of amino acids (R98, R99, R101, S113, W114, R115, T116, K117 and R146) within the binding site were

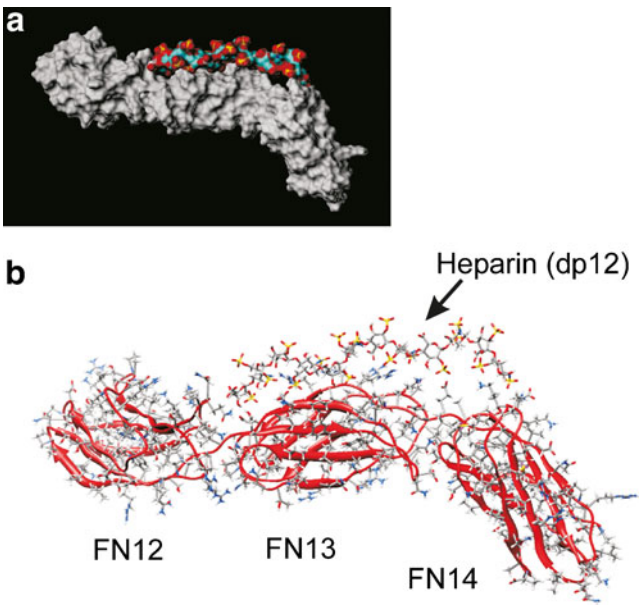

Fig. 1 Docking of the dp12 oligosaccharide $[IdUA2S\text{-}GlcNS6S]_6$ of heparin with the Hep-II domain of fibronectin. **a** Linear interactions of the Connoly's surfaces of the dp12 oligosaccharide of heparin (colored in cyan, red and orange) with the repeating units of the Hep-II domain. The oligosaccharide essentially interacts with FN13 and additionally with FN12 and FN14 (FN domains colored in grey). **b** Secondary structures of Hep-II domain are indicated in red arrows and chains in stick with heteroatomic colors

defined as flexible, while the oligosaccharide itself was kept partially rigid, with the exception of iduronate rings. Then, the three-dimensional conformation of the obtained complex was refined using energy minimization with side chains of amino acids in the binding site as well as oligosaccharide defined as mostly flexible. This final calculation allows predicting the best interacting conformation for oligosaccharide, taking into account the possible flexibility of iduronate rings. Indeed, the ligand is a dodecasaccharide of heparin, which is longer than hexasaccharides generally used. Thus, we studied the relative rigidity of the ligand by using a NVT molecular dynamics simulations of 1 ns length in a TIP4P water box. Such an integration time is long enough for a molecule of 300 atoms including hydrogen atoms (see for example Gandhi *et al.* 2012 [32]). The force field used in molecular dynamics simulation is the spectroscopic "SPASIBA" force field developed in our group for more than three decades and introduced into the CHARMM program [33, 34]. We decided to use the SPASIBA force field, as it was previously parameterized for glycosaminoglycans [33, 35]. In the derivation of SPASIBA parameters, our main concern was the development of a reliable force field able to reproduce accurately the vibrational modes. Effectively, as pointed out many years ago by S.W. Homans [36] a good fit between theoretical and experimental vibrational modes lends some confidence that the force field will give meaningful results in molecular dynamics simulations, which may require accuracy of the force field over the whole potential surface rather than in the region of the global minimum (conformational analysis using minimization techniques). By using these approaches, we obtained a refined model of dp12 interacting with Hep-II domain, which was in agreement with experimental data. On the contrary, the models in which dp12 was kept fully flexible were found aberrant and not reproducible (data not shown).

Molecular graphics and analysis of complexes

Molecular graphics and analysis were performed with the UCSF Chimera package [37]. Chimera (http://www.cgl.ucsf.edu/chimera/) is developed by the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco, with support from the National Institutes of Health (National Center for Research Resources grant 2P41RR001081, National Institute of General Medical Sciences grant 9P41GM103311).

Potential energy of interaction

In order to rank the complexes with energy values, we choose to calculate the potential energy of interaction, which may be determined using the simple expression: $E_{Interaction}=E_{Complex}$

$-(E_{Protein}+E_{Ligand})$. Energy values were calculated by using the MMFF94 Molecular Mechanics Force Field (see for example [38]) which is suitable for energy minimization of such compounds, not for molecular dynamics simulation.

Results

Molecular interaction between heparin and Hep-II domain of fibronectin

The binding energy of the complex between the Hep-II domain of fibronectin and the standard dp12 oligosaccharide of heparin [IdUA2S-GlcNS6S]$_6$ is equal to −3,017 kJ/mol. Twenty-seven H-bonds with sulfate or carboxyl groups were observed. Fine analysis of the complex revealed that H-bonds with amino acids residues are mainly located in the FN13 unit of the Hep-II domain (Fig. 2 and Table 1). Additionally, some interactions were also observed between amino acids in FN12 and FN14 units and the extremities of the oligosaccharide. However, these last interactions probably do not exist in the case of longer fragments of heparin.

In Table 1 are presented the amino acid residues of the Hep-II domain involved in the interactions with the dp12 oligosaccharide and their corresponding position in full-length fibronectin. As expected, we confirmed that R98, R99, R101 and R115 in FN13, which correspond to R1697, R1698, R1700 and R1714 in fibronectin, are involved in the interaction with heparin. Thus, these results validate our model, since it was established without taking into account previously published data concerning the interactions between heparin and fibronectin. In addition, we pointed out some interactions with other amino acids, *e.g.* V102, A105, T118, which are located on both sides of the arginine clusters in Hep-II domain.

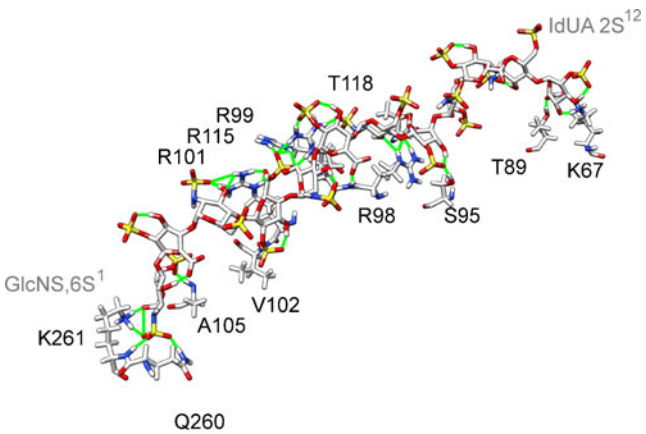

Fig. 2 Mapping of interacting amino acids of Hep-II with the standard oligosaccharide [IdUA2S-GlcNS6S]$_6$ of heparin. *Green lines* indicate H-Bonds

Table 1 H-bonds between the dp12 oligosaccharide [IdUA2S-GlcNS6S]$_6$ of heparin and Hep-II domain of fibronectin. The presence of multiple bonds between an interacting function and the same amino acid is indicated in brackets. The positions of amino acid residues involved in the interactions are numbered in Hep-II domain (FN12-14) and replaced in full-length fibronectin and FN13 unit. When peptide bonds rather than the side chains are involved in the formation of H-bonds, the indication of amino acid residues is followed by (NH) or (CO)

Monosaccharide	Interacting function	Amino acids in FN12-14	Amino acids in Fibronectin	Amino acids in FN13
GlcN^1	1-*O*H	K^{261}	K^{1860}	*(FN14)*
	N-SO$_3^-$	K^{261}	K^{1860}	*(FN14)*
		K^{261} (NH)	K^{1860} (NH)	*(FN14)*
		Q^{260}	Q^{1859}	*(FN14)*
	6-*O*-SO$_3^-$	A^{105}	A^{1704}	A^{13}
GlcN^3	*N*-SO$_3^-$	R^{101}	R^{1700}	R^9
	3-*O*H	R^{101}	R^{1700}	R^9
IdUA^4	2-*O*-SO$_3^-$	R^{101} (NH)	R^{1700} (NH)	R^9 (NH)
		V^{102} (NH)	V^{1791} (NH)	V^{10} (NH)
	-COO$^-$	R^{101}	R^{1700}	R^9
GlcN^5	6-*O*-SO$_3^-$ (x4)	R^{115}	R^{1714}	R^{23}
		R^{99}	R^{1698}	R^7
IdUA^6	2-*O*-SO$_3^-$	R^{99}	R^{1698}	R^7
	-COO$^-$	R^{99} (NH)	R^{1698} (NH)	R^7 (NH)
GlcN^7	*N*-SO$_3^-$	R^{99}	R^{1698}	R^7
	3-*O*H	R^{98}	R^{1697}	R^6
IdUA^8	COO$^-$ (*x*2)	R^{98}	R^{1697}	R^6
		T^{118}	T^{1717}	T^{26}
		T^{118} (NH)	T^{1717} (NH)	T^{26} (NH)
	2-*O*-SO$_3^-$	S^{95}	S^{1694}	S^3
IdUA^{12}	COO$^-$	T^{89}	T^{1688}	*(FN12)*
		K^{67}	K^{1666}	*(FN12)*
	4-*O*H	K^{67}	K^{1666}	*(FN12)*

Importance of the sulfate position in the interactions with Hep-II domain

We then explored the importance of the sulfate positions in the interactions with the Hep-II domain. To this end, a series of dp12 oligosaccharides were designed by site-directed desulfation and used in our docking experiments with the Hep-II domain. Calculation of the binding energy of each complex allowed us to define a hierarchic involvement of each sulfate position as follows: 2S >> 6S > NS. Indeed, the binding energies of the complexes formed with 2-*O*-desulfated (D2S), 6-*O*-desulfated (D6S) and *N*-desulfated (DNS) dp12 oligosaccharides were equal to −1,820 kJ/mol, −2,152 kJ/mol and −2,166 kJ/mol, respectively.

As expected, desulfation of oligosaccharides also reduced the number of H-bonds. However, we found that D2S (Fig. 3 and Table 2) and DNS oligosaccharides (Fig. 4 and Table 3)

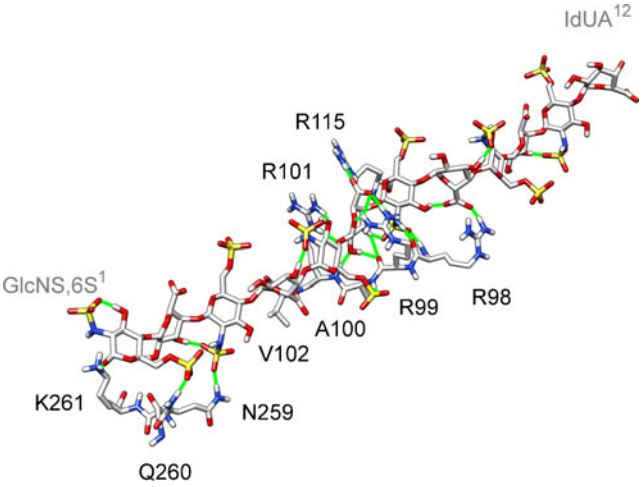

Fig. 3 Mapping of interacting amino acids of Hep-II with the D2S oligosaccharide [IdUA-GlcNS6S]$_6$ of heparin. *Green lines* indicate H-Bonds

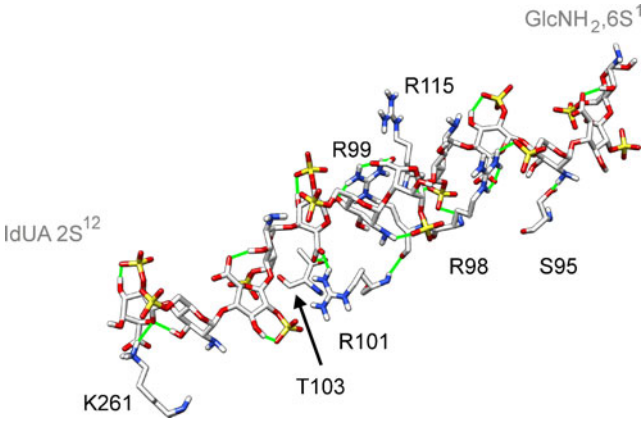

Fig. 4 Mapping of interacting amino acids of Hep-II with the DNS oligosaccharide [IdUA2S-GlcNH$_2$6S]$_6$ of heparin. *Green lines* indicate H-Bonds

established 19 H-bonds and 13 H-bonds, respectively, thus indicating that the binding energy is not directly related to the number of H-bonds. In the same way, we observed the formation of only 11 H-bonds with the D6S oligosaccharide (Fig. 5 and Table 4), even though the complex formed by the association of Hep-II with D6S oligosaccharide was more stable than the one formed with D2S oligosaccharides. These data clearly demonstrate that modifying the sulfate positions does not only determine the formation of H-bonds but also their binding energy. Interestingly, comparison of our models also highlight a mutual adaptation between Hep-II domain and desulfated oligosaccharides to form stable complexes, which involves different H-bonds with either the same amino acid residues or other ones.

We next analyzed the effect of 3-*O*-sulfation in the interactions between Hep-II domain and heparin, by adding 3-*O*-sulfate groups in the structure of the dp12 oligosaccharide (Fig. 6 and Table 5). The binding energy of the complex formed between 3-*O*-sulfated (R3S) oligosaccharide and Hep-II domain was equal to −3,251 kJ/mol. The minimum of energy of the complex formed with unmodified dp12 oligosaccharide was equal to −3,017 kJ/mol, thus indicating that additional 3-*O*-sulfation to the GlcNS6S residues has stabilized even more the interactions with Hep-II domain.

Table 2 H-bonds between the D2S oligosaccharide [IdUA-GlcNS6S]$_6$ of heparin and Hep-II domain of fibronectin. The positions of amino acid residues involved in the interactions are numbered in Hep-II domain. When peptide bonds rather than the side chains are involved in the formation of H-bonds, the indication of amino acid residues is followed by (NH) or (CO)

Monosaccharide	Interacting function	Amino acids in FN12-14
GlcN^1	1-*OH*	K^{261}
	6-*O*-SO$_3^-$	Q^{260}
GlcN^3	*N*-SO$_3^-$	N^{259}
IdUA^4	COO$^-$	V^{102} (NH)
GlcN^5	6-*O*-SO$_3^-$	A^{100} (NH)
	3-*OH*	R^{101}
	GlcN^5-*O*-IdUA^6	R^{101}
IdUA^6	2-*OH*	R^{99} (CO)
		R^{101} (NH)
	3-*OH*	R^{99} (CO)
	O(heterocycle)	R^{99}
	COO$^-$ (*x*2)	R^{99}
	(*x*2)	R^{115}
GlcN^7	*N*-SO$_3^-$	R^{98}(NH)
		R^{115}(NH)
	H(*N*-SO$_3^-$)	R^{98}(CO)
IdUA^8	COO$^-$	R^{98}

Table 3 H-bonds between the DNS oligosaccharide [IdUA2S-GlcNH$_2$6S]$_6$ and Hep-II domain of fibronectin. The presence of multiple bonds between an interacting function and the same amino acid is indicated in brackets. The positions of amino acid residues involved in the interactions are numbered in Hep-II domain. When peptide bonds rather than the side chains are involved in the formation of H-bonds, the indication of amino acid residues is followed by (NH) or (CO)

Monosaccharide	Interacting function	Amino acids in FN12-14
GlcN^3	NH$_2$	S^{95}
	6-*O*-SO$_3^-$	R^{98}
IdUA^4	COO$^-$ (*x*2)	R^{98}
	O(heterocycle)	R^{98}
GlcN^5	6-*O*-SO$_3^-$	R^{115}(NH)
IdUA^6	COO$^-$ (*x*2)	R^{99}
GlcN^7	GlcN^7-*O*-IdUA^8	R^{99}
IdUA^8	COO$^-$	R^{101}
	COO$^-$	T^{103}
IdUA^{12}	*O*(heterocycle)	K^{261}

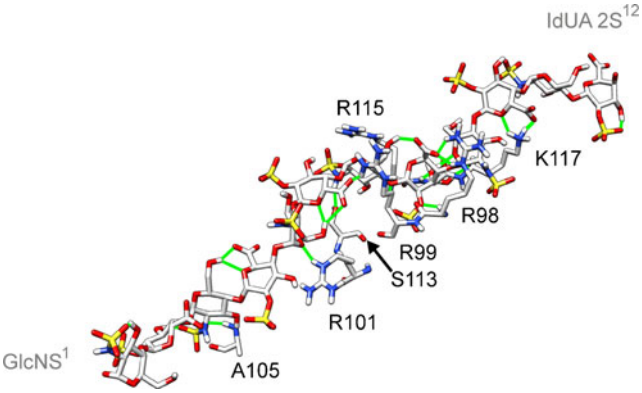

Fig. 5 Mapping of interacting amino acids of Hep-II with the D6S oligosaccharide [IdUA2S-GlcNS]$_6$ of heparin. *Green lines* indicate H-Bonds

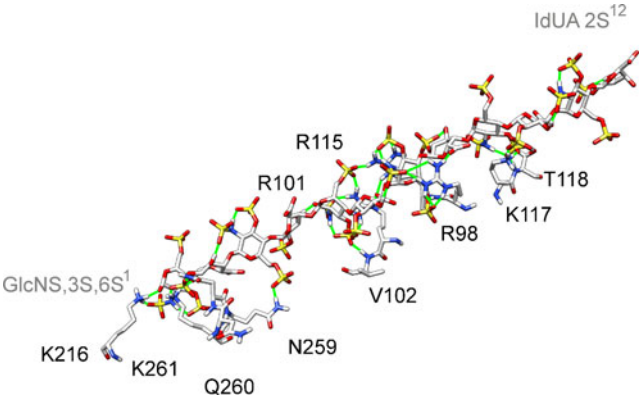

Fig. 6 Mapping of interacting amino acids of Hep-II with the R3S oligosaccharide [IdUA2S-GlcNS6S3S]$_6$ of heparin. *Green lines* indicate H-Bonds

Interestingly, we observed the formation of 23 H-bonds between the 3-*O*-sulfated oligosaccharide and Hep-II domain, instead of 24 with the non-modified oligosaccharide. This clearly indicates that the binding energy rather than the number of H-bonds is a critical factor for stabilizing the complex between heparin and Hep-II domain of fibronectin.

Comparison of the different models reveals that the molecular interactions between heparin oligosaccharides and Hep-II domain of fibronectin exhibit more complexity than what was predictable. Indeed, we might expect that the dp12 oligosaccharides take a position centered on the arginine cluster of FN13 unit and thereafter establish secondary interactions according to the presence or not of sulfate groups. In this model, the formation of H-bonds would be similar to a system of bar code, in which one sulfate would correspond to one interaction. Consistently, the number of established H-bonds would be directly related to the stability of the complexes. However, our docking experiments support another model. We found that modification in the sulfation pattern rather than in the number of sulfate groups leads to the formation of different complexes, in which one sulfate group can be involved in the interactions with the same amino acid residue or another one. Consequently, the formation of the more stable complexes implies a mutual adaptation between the Hep-II domain and

Table 4 H-bonds between the D6S oligosaccharide [IdUA2S-GlcNS]$_6$ of heparin and Hep-II domain of fibronectin. The presence of multiple bonds between an interacting function and the same amino acid is indicated in brackets. The positions of amino acid residues involved in the interactions are numbered in Hep-II domain. When peptide bonds rather than the side chains are involved in the formation of H-bonds, the indication of amino acid residues is followed by (NH) or (CO)

Monosaccharide	Interacting function	Amino acids in FN12-14
GlcN^3	*N*-SO_3^-	A^{105}(NH)
GlcN^5	*O*(heterocycle)	R^{101}
IdUA^6	COO^-	R^{99}
GlcN^7	*N*H-3-*OH*	S^{113}
		R^{115}(NH)
IdUA^8	2-*O*-SO_3^-	R^{98}(NH)
	COO^- (*x*2)	R^{98}
	IdUA^8-*O*-GlcN^9	R^{98}
IdUA^{10}	*O*(heterocycle)	K^{117}
	COO^-	K^{117}

Table 5 H-bonds between the R3S oligosaccharide [IdUA2S-GlcNS6S3S]$_6$ of heparin and Hep-II domain of fibronectin. The positions of amino acid residues involved in the interactions are numbered in Hep-II domain. When peptide bonds rather than the side chains are involved in the formation of H-bonds, the indication of amino acid residues is followed by (NH) or (CO)

Monosaccharide	Interacting function	Amino acids in FN12-14
GlcN^1	1-*O*H	K^{216}
	N-SO_3^-	K^{216}
	GlcNS^1-*O*-IdUA^2	K^{261}
IdUA^2	2-*O*-SO_3^-	K^{261}
		Q^{260}
GlcN^3	*6*-SO_3^-	N^{259}
GlcN^5	3-*O*-SO_3^-	V^{102} (NH)
		R^{101}
	N-SO_3^- (*x*2)	R^{101}
	6-*O*-SO_3^-	R^{101}
		R^{115}
IdUA^6	COO^-	R^{115}
GlcN^7	6-*O*-SO_3^- (*x*2)	R^{98}
	N-SO_3^- (*x*2)	R^{98}
	3-*O*-SO_3^-	R^{115}
IdUA^8	COO^-	R^{98}
GlcN^9	3-*O*-SO_3^-	K^{117}
		T^{118}(NH)
		T^{118}
IdUA^{10}	3-*O*-SO_3^-	K^{117}

the dp12 oligosaccharides, which is highly dependent on interplay between the different positions of sulfate groups.

Discussion

Experimental approaches to analyze the interactions between proteins and HS/heparin are limited because of the large size of the fragment needed for an efficient binding, the flexibility of the linkages and the high density of negative charges. Alternatively, molecular modeling has proven to be a powerful method to study protein-heparin/HS complexes [39].

By using molecular docking, we proposed models of the interactions between the Hep-II domain of fibronectin and a series of heparin oligosaccharides. First, we confirmed that the amino acids residues R98, R99, R101, and R115 located in FN13 unit are involved predominantly in the interaction with the standard heparin dodecasaccharide $[IdUA2S\text{-}GlcNS6S]_6$. We then defined a hierarchic involvement of sulfate positions in the force of interactions as follows: 2S >> 6S > NS. We also found that 3-*O*-sulfation, which is a rare modification in heparin and HS, stabilizes the intermolecular interactions of the complex formed with the Hep-II domain of fibronectin.

The overall specificity of these interactions is compatible with previous experimentally-determined results on the structure of Hep-II domain in complex with heparin, thus indicating that our model is relevant [24]. Unfortunately, our findings were only partially confirmed by NMR analysis [22]. This last study demonstrated that the binding reaction between a fragment corresponding to FN13-FN14 units and heparin-derived octasaccharides involves R98, R99, R101 and R115, while other basic residues were not required for the interaction. However, the authors have used a short fragment corresponding to FN13-FN14 units and heparin-derived octasaccharides as ligands. The discrepancy may be due to the sizes of the binding partners, which are probably not large enough to mimic the interaction between heparin and Hep-II domain of fibronectin. Indeed, peptides derived from FN14 unit in Hep-II domain were reported to weakly bind to heparin [23], and heparin-derived oligosaccharides from dp12 are required for high-affinity binding to fibronectin [24]. That may explain why neither our docking experiments nor previous experimental data are in total agreement with NMR data.

Molecular docking with GOLD program has been widely used to predict specific interactions between heparin/HS and proteins. In many studies, the authors have used the high-resolution crystal structure of proteins complexed with heparin fragments to further predict the specificity of the interaction with HS sequences. As example, Raghuraman *et al.* (2006) [40] have described a computational virtual screening approach with GOLD to predict the high specificity of the heparin/HS sequences interacting with antithrombin. To this end, they have used a high-resolution crystal structure of the protein complexed with heparin fragments, and they have predicted the specificity of the interactions of differently sulfated HS sequences. Thus, screening of a computational virtual library of heparin hexasaccharides using a dual filter strategy, in which the first filter was the predicted affinity of antithrombin and the second filter was the self-consistency of docking, allowed them to identify specific interacting sequences. In the same way, Raghuraman *et al.* (2010) [41] have used a virtual screening library of dermatan sulfate hexasaccharides to study the interaction with heparin cofactor II. If the crystal structure for a protein is not available, a model can be constructed by homology and then used for docking heparin analogs or GAGs mimetics into the binding site. As example, Gandhi *et al.* (2012) [32] have recently used this approach to propose models of the interactions of HS with heparanase.

Several modeling works have used GOLD program to point out specificity and selectivity of GAG binding to proteins. In this way, the three-dimensional structure of the extracellular immunoglobulin domains of PECAM-1 was constructed using homology modeling and threading methods. Potential heparin/HS binding sites were predicted on the basis of amino acid consensus sequences and a comparison with other known structures of sulfate-binding proteins [42]. More recently, another modeling work was performed to study the specificity of heparin/HS binding to antithrombin or to thrombin. Analysis of the complexes suggests the presence of a reasonably sized bifurcated cavity in antithrombin that facilitates a firm 'hand-shake' with H/HS and tight interaction with water molecules. Such a significant cavity was detected in the binding site of antithrombin, but not in thrombin, which may explain the heparin/HS specificity of binding [43]. In their works, Gandhi *et al.* (2008) [42] and Mosier *et al.* (2012) [43] used short oligosaccharides, *i.e.* pentasaccharides, as ligands for the interactions. In these models, short oligosaccharides interact with a cavity in the binding site of the protein.

However, previous works have demonstrated that the minimal length for interaction between heparin fragments and Hep-II is an octasaccharide and that a dp14 was required for high-affinity binding [24]. Long oligosaccharides bind to the full amplitude of a domain of fibronectin thus adopting a deployed configuration. Thus, we decided to study the flexibility of the ligands by using the "SPASIBA" force field prior docking experiments with GOLD program. Even though this procedure is probably not required for small oligosaccharides, it is adapted to oligosaccharides of dp longer than hexasaccharides.

In other cases, computational approach to study GAG-protein interactions was complicated because of the presence of solvent water for the interactions. Indeed, docking results could be better by inclusion of water molecules at the protein binding site. For example, in the case of the complexes formed between IL-8 and heparin disaccharides, half of the

interactions in GAG-protein interfaces are water-mediated [44]. In our docking experiments, inclusion of water molecules did not however modify the interactions between fibronectin and heparin/HS (data not shown). This is probably due to the length of the dodecasaccharide chain and to the linearity of the interaction between the ligand and binding site in Hep-II domain of fibronectin.

In most of the studies on the structure of intermolecular complexes between HS/heparin and proteins, the sulfated oligosaccharide is usually considered as a simple extended binding site. However, our observations point out a more complex system of mutual adaptation of both partners for the interactions. According to this model, the presence of sulfate groups at one position does not only account for the number of H-bonds, but also determines the interactions between other sulfate groups and amino acid residues in the intermolecular complex. This may explain the differences between the structural disposition of heparin and HS in complexes formed with fibronectin [22, 24, 25] and possibly other heparin-binding proteins.

References

1. Pankov, R., Yamada, K.M.: Fibronectin at a glance. J. Cell Sci. **115**, 3861–3863 (2002)
2. Woods, A., Couchman, J.R.: Syndecan 4 heparan sulfate proteoglycan is a selectively enriched and widespread focal adhesion component. Mol. Biol. Cell **5**, 183–192 (1994)
3. Petersen, T.E., Thøgersen, H.C., Skorstengaard, K., Vibe-Pedersen, K., Sahl, P., Sottrup-Jensen, L., Magnusson, S.: Isolation and characterization of cDNA clones for human and bovine fibronectins. Proc. Natl. Acad. Sci. U. S. A. **80**, 137–141 (1983)
4. Pierschbacher, M.D., Ruoslahti, E.: Variants of the cell recognition site of fibronectin that retain attachment-promoting activity. Proc. Natl. Acad. Sci. U. S. A. **81**, 5985–5988 (1984)
5. Ruoslahti, E.: Fibronectin and its receptors. Annu. Rev. Biochem. **57**, 375–413 (1988)
6. Humphries, M.J., Komoriya, A., Akiyama, S.K., Olden, K., Yamada, K.M.: Identification of two distinct regions of the type III connecting segment of human plasma fibronectin that promote cell type-specific adhesion. J. Biol. Chem. **262**, 6886–6892 (1987)
7. Sharma, A., Askari, J.A., Humphries, M.J., Jones, E.Y., Stuart, D.I.: Elucidation of the structural features of heparan sulfate important for interaction with the Hep-2 domain of fibronectin. EMBO J. **18**, 1468–1479 (1999)
8. Garcia-Pardo, A., Rostagno, A., Frangione, B.: Primary structure of human plasma fibronectin. Characterization of a 38 kDa domain containing the C-terminal heparin-binding site (Hep III site) and a region of molecular heterogeneity. Biochem. J. **241**, 923–928 (1987)
9. Kishore, R., Samuel, M., Khan, M.Y., Hand, J., Frenz, D.A., Newman, S.A.: Interaction of the NH2-terminal domain of fibronectin with heparin. Role of the omega-loops of the type I modules. J. Biol. Chem. **272**, 17078–17085 (1997)
10. Benecky, M.J., Kolvenbach, C.G., Amrani, D.L., Mosesson, M.W.: Evidence that binding to the carboxyl-terminal heparin-binding domain (Hep II) dominates the interaction between plasma fibronectin and heparin. Biochemistry **27**, 7565–7571 (1988)
11. Ingham, K.C., Brew, S.A., Atha, D.A.: Interaction of heparin with fibronectin and isolated fibronectin domains. Biochem. J. **272**, 605–611 (1990)
12. Lin, H., Lal, R., Clegg, D.O.: Imaging and mapping heparin-binding sites on single fibronectin molecules with atomic force microscopy. Biochemistry **39**, 3192–3196 (2000)
13. Mostafavi-Pour, Z., Askari, J.A., Whittard, J.D., Humphries, M.J.: Identification of a novel heparin-binding site in the alternatively spliced IIICS region of fibronectin: roles of integrins and proteoglycans in cell adhesion to fibronectin splice variants. Matrix Biol. **20**, 63–73 (2001)
14. Gui, L., Wojciechowski, K., Gildner, C.D., Nedelkovska, H., Hocking, D.C.: Identification of the heparin-binding determinants within fibronectin repeat III1: role in cell spreading and growth. J. Biol. Chem. **46**, 34816–34825 (2006)
15. Esko, J.D., Selleck, S.B.: Order out of chaos: assembly of ligand binding sites in heparan sulfate. Annu. Rev. Biochem. **71**, 435–471 (2002)
16. Capila, I., Linhardt, R.J.: Heparin-protein interactions. Angew. Chem., Int. Ed. Engl. **41**, 391–412 (2002)
17. Ori, A., Wilkinson, M.C., Fernig, D.G.: The heparanome and regulation of cell function: structures, functions and challenges. Front. Biosci. **13**, 4309–4338 (2008)
18. Sugahara, K., Kitagawa, H.: Heparin and heparan sulfate biosynthesis. IUBMB Life **54**, 163–175 (2002)
19. Shriver, Z., Capila, I., Venkataraman, G., Sasisekharan, R.: Heparin and heparan sulfate: analyzing structure and microheterogeneity. Handb. Exp. Pharmacol. **207**, 159–176 (2012)
20. Barkalow, F.J., Schwarzbauer, J.E.: Localization of the major heparin-binding site in fibronectin. J. Biol. Chem. **266**, 7812–7818 (1991)
21. Busby, T.F., Argraves, W.S., Brew, S.A., Pechik, I., Gilliland, G.L., Ingham, K.C.: Heparin binding by fibronectin module III-13 involves six discontinuous basic residues brought together to form a cationic cradle. J. Biol. Chem. **270**, 18558–18562 (1995)
22. Sachchidanand, Lequin, O., Staunton, D., Mulloy, B., Forster, M.J., Yoshida, K., Campbell, I.D.: Mapping the heparin-binding site on the 13-14F3 fragment of fibronectin. J. Biol. Chem. **277**, 50629–50635 (2002)
23. Ingham, K.C., Brew, S.A., Migliorini, M.M., Busby, T.F.: Heparin binding by fibronectin module III-13 involves six discontinuous basic residues brought together to form a cationic cradle. Biochemistry **32**, 12548–12553 (1993)
24. Lyon, M., Rushton, G., Askari, J.A., Humphries, M.J., Gallagher, J.T.: Elucidation of the structural features of heparan sulfate important for interaction with the Hep-2 domain of fibronectin. J. Biol. Chem. **275**, 4599–4606 (2000)
25. Mahalingam, Y., Gallagher, J.T., Couchman, J.R.: Cellular adhesion responses to the heparin-binding (HepII) domain of fibronectin require heparan sulfate with specific properties. J. Bio. Chem. **282**, 3221–3230 (2007)
26. Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A.A., Aflalo, C., Vakser, I.A.: Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. Proc. Natl. Acad. Sci. U. S. A. **89**, 2195–2199 (1992)
27. Ausiello, G., Cesareni, G., Helmer-Citterich, M.: ESCHER: a new docking procedure applied to the reconstruction of protein tertiary structure. Proteins **28**, 556–567 (1997)
28. Palma, P.N., Krippahl, L., Wampler, J.E., Moura, J.: BiGGER: a new (soft) docking algorithm for predicting protein interactions. J. Proteins **39**, 372–384 (2000)
29. Ritchie, D.W., Kozakov, D., Vajda, S.: Accelerating and focusing protein-protein docking correlations using multi-dimensional rotational FFT generating functions. Bioinformatics **24**, 1865–1873 (2008)
30. Hex's home page: http://www.loria.fr (~ ritchied/hex) 2013. Accessed 20 April 2013

31. Jones, G., Willet, P., Glen, R.C., Leach, A.R., Taylor, R.: Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol. **267**, 727–748 (1997)
32. Gandhi, N.S., Freeman, C., Parish, C.R., Mancera, R.L.: Computational analyses of the catalytic and heparin-binding sites and their interactions with glycosaminoglycans in glycoside hydrolase family 79 endo-β-D-glucuronidase (heparanase). Glycobiology **22**, 35–55 (2012)
33. Vergoten, G., Mazur, I., Lagant, P., Michalski, J.C., Zanetta, J.P.: The SPASIBA force field as an essential tool for studying the structure and dynamics of saccharides. Biochimie **85**, 65–73 (2003)
34. Lagant, P., Nolde, D., Stote, R., Vergoten, G., Karplus, M.: Increasing normal modes analysis accuracy: the SPASIBA spectroscopic force field introduced into the CHARMM program. J. Phys. Chem. **108**, 4019–4029 (2004)
35. Meziane-Tani, M., Lagant, P., Semmound, A., Vergoten, G.: The SPASIBA force field for chondroitin sulfate: vibrational analysis of D-glucuronic and N-acetyl-D-galactosamine 4-sulfate sodium salts. J. Phys. Chem. **110**, 11359–11370 (2006)
36. Homans, S.W.: A molecular mechanical force field for the conformational analysis of oligosaccharides: comparison of theoretical and crystal structures of Man alpha 1-3Man beta 1-4GlcNAc. Biochemistry **29**, 9110–9118 (1990)
37. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E.: UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. **25**, 1605–1612 (2004)
38. Halgren, T.A.: MMFF VII. Characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. J. Comput. Chem. **20**, 730–748 (1999)
39. Sapay, N., Cabannes, E., Petitou, M., Imberty, A.: Molecular modeling of the interaction between heparan sulfate and cellular growth factors: bringing pieces together. Glycobiology **21**, 1181–1193 (2011)
40. Raghuraman, A., Mosier, P.D., Desai, U.R.: Finding a needle in a haystack: development of a combinatorial virtual screening approach for identifying high specificity heparin/heparan sulfate sequence(s). J. Med. Chem. **49**, 3553–3562 (2006)
41. Raghuraman, A., Mosier, P.D., Desai, U.R.: Understanding dermatan sulfate-heparin cofactor II interaction through virtual library screening. ACS Med. Chem. Lett. **1**, 281–285 (2010)
42. Gandhi, N.S., Coombe, D.R., Mancera, R.L.: Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 1. Molecular modeling studies. Biochemistry **47**, 4851–4862 (2008)
43. Mosier, P.D., Krishnasamy, C., Kellogg, G.E., Desai, U.R.: PLoS One (2012) doi: 7/e48632
44. Samsonov, S.A., Teyra, J., Pisabarro, M.T.: Docking glycosaminoglycans to proteins: analysis of solvent inclusion. J. Comput. Aided Mol. Des. **25**, 477–489 (2011)