Mapping of heparin/heparan sulfate binding sites on αvβ3 integrin by molecular docking

ARTICLE *in* JOURNAL OF MOLECULAR RECOGNITION · FEBRUARY 2013

Impact Factor: 2.15 · DOI: 10.1002/jmr.2250 · Source: PubMed

Research Article

Journal of Molecular Recognition

Received: 30 July 2012, Revised: 15 October 2012, Accepted: 15 October 2012, Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jmr.2250

Mapping of heparin/heparan sulfate binding sites on αvβ3 integrin by molecular docking

Lionel Ballut[a†], Nicolas Sapay[b†], Émilie Chautard[a], Anne Imberty[b] and Sylvie Ricard-Blum[a*]

Heparin/heparan sulfate interact with growth factors, chemokines, extracellular proteins, and receptors. Integrins are αβ heterodimers that serve as receptors for extracellular proteins, regulate cell behavior, and participate in extracellular matrix assembly. Heparin binds to RGD-dependent integrins (αIIbβ3, α5β1, αvβ3, and αvβ5) and to RGD-independent integrins (α4β1, αXβ2, and αMβ2), but their binding sites have not been located on integrins. We report the mapping of heparin binding sites on the ectodomain of αvβ3 integrin by molecular modeling. The surface of the ectodomain was scanned with small rigid probes mimicking the sulfated domains of heparan sulfate. Docking results were clustered into binding spots. The best results were selected for further docking simulations with heparin hexasaccharide. Six potential binding spots containing lysine and/or arginine residues were identified on the ectodomain of αvβ3 integrin. Heparin would mostly bind to the top of the genu domain, the Calf-I domain of the α subunit, and the top of the β subunit of RGD-dependent integrins. Three spots were close enough from each other on the integrin surface to form an extended binding site that could interact with heparin/heparan sulfate chains. Because heparin does not bind to the same integrin site as protein ligands, no steric hindrance prevents the formation of ternary complexes comprising the integrin, its protein ligand, and heparin/heparan sulfate. The basic amino acid residues predicted to interact with heparin are conserved in the sequences of RGD-dependent but not of RGD-independent integrins suggesting that heparin/heparan sulfate could bind to different sites on these two integrin subfamilies. Copyright © 2013 John Wiley & Sons, Ltd.
Supporting information may be found in the online version of this paper

Keywords: heparin; heparan sulfate; integrin; molecular docking; interaction

INTRODUCTION

Heparin (HP) and heparan sulfate (HS) are complex polysaccharides belonging to the glycosaminoglycan family. Both glycosaminoglycans are made of the repeat of disaccharide units comprising one hexuronic acid and one glucosamine that are modified by sulfation and acylation during biosynthesis (Bülow and Hobert, 2006) (Figure 1(A)). HS chains can be covalently attached to several proteins to form proteoglycans located in the extracellular matrix and at the surface of the cells. HP and HS bind to a variety of proteins such as enzymes, enzyme inhibitors, chemokines, growth factors, lipid-binding proteins, morphogens, nuclear proteins, extracellular matrix proteins, cell adhesion proteins and cell surface receptors such as tyrosine-kinase growth factor receptors (Imberty *et al*., 2007, Gandhi and Mancera, 2008, Esko and Linhardt, 2009).

HS also bind to integrins, which are αβ heterodimers that act as cell surface receptors for extracellular proteins (Barczyk *et al*., 2010). The α subunit comprises a β-propeller, a thigh domain and two calf domains, whereas the β subunit contains a plexin–semaphorin–integrin domain, a hybrid domain, a βA domain, and four cysteine-rich epidermal growth factor repeats (Figure 1(B)). Integrins control cell adhesion, migration, growth, differentiation, and apoptosis in development and in physiopathological situations. HP binds to the leukocyte-specific receptors α4β1, αXβ2, and αMβ2 integrins (Diamond *et al*., 1995, Vorup-Jensen *et al*., 2005, Vorup-Jensen *et al*., 2007) and to four RGD-dependent integrins, αIIbβ3 (Sobel and Adelman, 1988), α5β1, αvβ3, and αvβ5 (Faye *et al*., 2009). The characterization of the interface of the HS/integrins complex would be of interest to determine if ternary complexes comprised of integrin, protein ligand, and HS could form *in vivo* and play a biological role. It could be the case with α5β1, αvβ3, and αvβ5 integrins, which are involved in angiogenesis (Hynes, 2007) and bind both to endostatin, the C-terminal anti-angiogenic fragment of collagen XVIII (Rehn *et al*., 2001, Faye *et al*., 2009), and to HP/HS (Faye *et al*., 2009). HS-integrin interactions do not contribute significantly to cell adhesion but mediate cell spreading (Faye *et al*., 2009). We have shown by inhibition experiments and molecular modeling that

* *Correspondence to: S. Ricard-Blum, UMR 5086 CNRS-Université Lyon 1, Institut de Biologie et Chimie des Protéines, 7 passage du Vercors, 69367 Lyon Cedex 07, France.*
E-mail: s.ricard-blum@ibcp.fr

† *These authors contributed equally to this work.*

a *L. Ballut, É. Chautard, S. Ricard-Blum*
UMR 5086 CNRS-Université Lyon 1, Institut de Biologie et Chimie des Protéines, 7 passage du Vercors, 69367 Lyon Cedex 07, France

b *N. Sapay, A. Imberty*
Centre de Recherche sur les Macromolécules Végétales, UPR5301 CNRS (affiliated with Université de Grenoble), 601 rue de la Chimie, BP 53, 38041 Grenoble Cedex 9, France

Abbreviations: *HS, heparan sulfate; IdoAS, 2-sulfate-L-iduronic acid; GlcNS, 2,6-sulfate-D-glucosamine; $MeSO_4^-$, methylsulfate.*

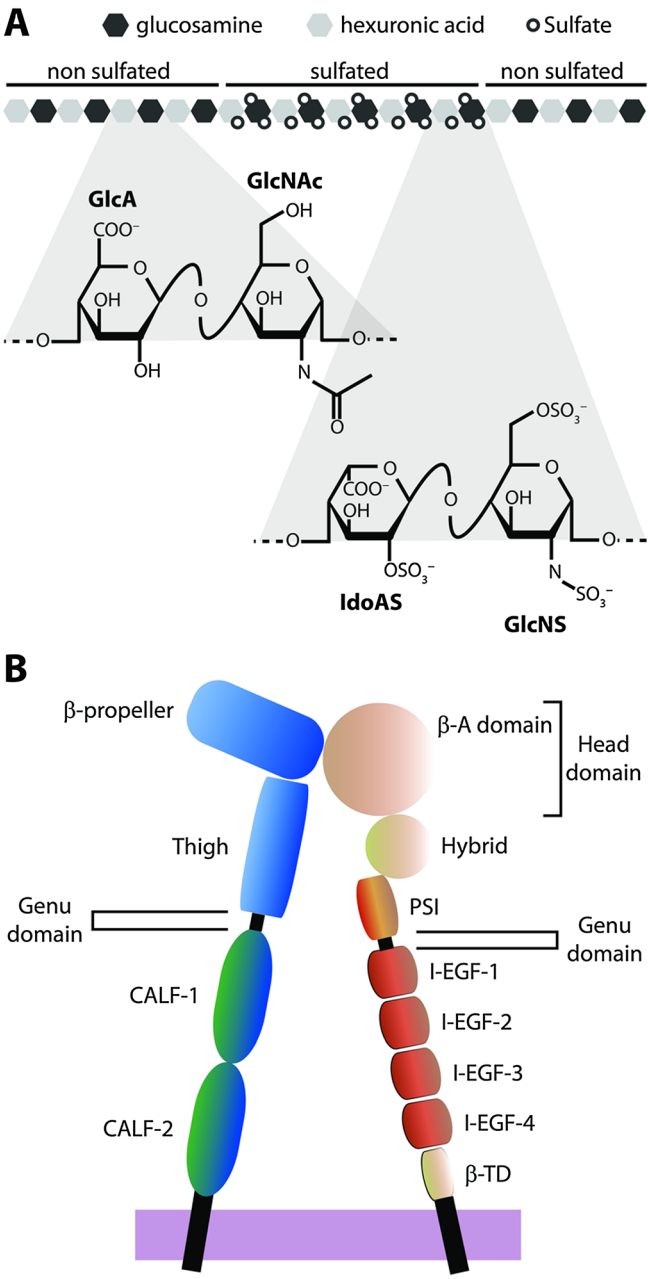

Figure 1. (A) Schematic representation of heparan sulfate (GlcA, glucuronic acid; IdoA, iduronic acid; GlcNAc, N-acetylated glucosamine; IdoAS, 2-sulfated iduronic acid; GlcNS, 6-sulfated-N-sulfated glucosamine. (B) Domain structure of the αvβ3 integrin (blue, αv subunit; red/orange, β3 subunit; EGF, epidermal growth factor; TD, tail domain).

endostatin binds on the head of the αvβ3 integrin, at the interface between the β-propeller domain of the αv subunit and the βA domain of the β3 subunit (Faye *et al.*, 2009), but we did not determine either the location of HP/HS binding site(s) on the αvβ3 integrin or the role of the glycosaminoglycans, which also interact with endostatin (Ricard-Blum *et al.*, 2004), in the complex.

We report here the molecular modeling of the HS-αvβ3 integrin complex, based on the X-ray structure of the ectodomain of the human αvβ3 integrin, which we have previously used to predict the localization of the endostatin binding site on this integrin. The extracellular domain is bent in the crystal structure, but although the bent conformation is most probably adopted by inactive integrins (reviewed by Askari *et al.*, 2009), this conformation of the ectodomain of αvβ3 integrin is able to stably bind a physiological ligand in solution (Adair *et al.*, 2005). Molecular modeling of HS–protein complexes is known as a difficult task because of the flexibility of the ligands, including ring shape fluctuation in the iduronic ring (chair $^1C_4$ and skew ^2S_O), the occurrence of many charged groups, and the general absence of well-defined pocket on the protein surface (Forster and Mulloy 2006, Imberty and Pérez, 2006). Few molecular docking methods have been specifically elaborated to the case of HS–protein complex (Bitomsky and Wade, 1999, Sadir *et al.*, 2001, Ricard-Blum *et al.*, 2004, Mahoney *et al.*, 2005, Raghuraman *et al.*, 2006, Sapay *et al.*, 2011), but because of the large size of the integrin ectodomain (956 and 690 amino acid residues for the αv and β3 subunits, respectively), we developed a new approach based on overlapping grids with fragments of increasing size. We predicted the existence of six binding sites that all contain at least one lysine or arginine residue interacting with sulfate moieties. Five potential binding sites are located on the α subunit, and one is located on the β subunit of RGD-binding integrins. Three of those sites are close enough to form a large binding surface for HP/HS. According to the best scores, HP would mostly bind to the top of the integrin genu domain, to the Calf-I domain of the α subunit, and on the top of the β subunit of RGD-dependent integrins. Because HP/HS does not bind to the same integrin site than protein ligands, ternary complexes comprised of integrin, one ligand and HP/HS could form *in vivo*. The basic amino acid residues predicted to interact with HP are conserved in the sequences of human RGD-dependent integrins, but not in RGD-independent integrins suggesting that HP/HS could bind to different sites on these two integrin subgroups.

MATERIALS AND METHODS

Workflow

The modeling protocol was divided into four steps: (i) molecular docking of a methyl sulfate probe by rigid molecular docking simulation on the entire integrin surface; (ii) molecular docking of a series of half-rigid four HP disaccharides (HS_2) on the entire integrin surface; (iii) selection of the best locations for the rigid docking of an HP hexasaccharide (HS_6); and (iv) rigid docking of a series of 40 HS_6 conformers on the selected locations.

Preparation of the integrin structure

The tridimensional coordinates of the ectodomain of the αvβ3 integrin were obtained from the 1JV2 entry (Xiong *et al.*, 2001) in the Protein Data Bank (PDB; Berman *et al.*, 2000). The calcium ions Ca^{2+} and the N-linked glycans were removed. The N-termini and C-termini of each integrin subunit were methylated, except the N-terminus of the αv subunit, which corresponds to chain extremity after removal of the signal peptide. Finally, the hydrogen atoms were added to the structure and their positions were optimized. The preparation of the protein structure was carried out with SybylX 1.2 (www.tripos.com). The structure was minimized using the Amber 99SB force field (Hornak *et al.*, 2006). The Amber partial charges were reused in all the following docking simulations. Amino acids were numbered as in the UniprotKB (The Uniprot Consortium, 2012) entries P06756 and P05106, corresponding to the human αv and β3 subunits, respectively.

Preparation of methylsulfate and HP oligosaccharide structures

The structure of methyl sulfate ($MeSO_4^-$) was built manually with SybylX 1.2 and the Tripos force field. Four HS_2 and two hexasaccharides were extracted from the NMR structure of an HP dodecamer (PDB entry 1HPN) (Mulloy *et al.*, 1993) as a function of the reducing sugar units and/or the ring puckering of the iduronic acid residues IdoAS (Table 1). All oligosaccharides were energy minimized in vacuum with their backbone heavy atoms restrained to their initial position (i.e. carbon and oxygen atoms from the glycosidic rings and linkages). Minimization was performed with NAMD 2.7b1 and the GLYCAM06 force field (Kirschner *et al.*, 2008). The partial charges for the methyl sulfate, IdoAS and GlcNS residues were the same than those previously used to build the model HS-cellular growth factors complexes (Sapay *et al.*, 2011). They have been derived using a protocol consistent with GLYCAM06.

In order to generate a library of hexasaccharides conformers, a series of 40 structures were generated by molecular dynamics simulation in water. The two HS_6 structures (Table 1) were placed in a water box of, viz. $50 \times 56 \times 65$ Å^3. The system was neutralized with 12 sodium ions and equilibrated for 1 ns to 298 K and one bar by Langevin dynamics with a 2 fs time step. Non-bonded interactions were cut off at 12 Å. A switch function was applied above 10 Å. Long range electrostatics was treated by the particle mesh Ewald method (Darden *et al.*, 1993, Essmann *et al.*, 1995) with a grid spacing of approximately 1 Å. The two equilibrated systems were then simulated for 20 ns with coordinates saved every 5 ps using the same parameters than for equilibration. Finally, the oligosaccharide coordinates were extracted every nanosecond from both trajectories for a total of 40 conformations of the HS_6 oligosaccharide. Hence, each conformation is a unique combination of torsion angles around the glycosidic linkage and iduronic acid ring conformations. Molecular dynamics simulations were performed with NAMD 2.7b1 (Phillips *et al.*, 2005). Analysis of the trajectories and the docking results were performed with VMD 1.8.7 (Humphrey *et al.*, 1996), R 2.10 (R-development-core-team, 2005), and the ggplot2 package (Wickham, 2009). A summary of the main structural features of the conformer library is available as supplementary data (Supplementary Table S1 and Supplementary Figure S1).

Molecular docking simulations

All the molecular docking simulations were performed with the Lamarckian genetic algorithm of Autodock 4.2 (Morris *et al.*, 1998, Morris *et al.*, 2009) using the same partial charges than for the preparation of protein and ligands. For each simulation, 150 runs of genetic algorithm were performed using a population of 200 individuals and a maximum of 5×10^6 score evaluations or 27×10^3 generations. The convergence of 150 runs was checked, and only the result with the best Autodock score was retained for further analysis. It should be noted that a similar protocol was used previously for the Fibroblast Growth Factor 2 in complex with an HP oligosaccharide. Autodock 4.2 was able to retrieve successfully the binding site of oligosaccharide co-crystallized with protein (Sapay *et al.*, 2011, Sapay *et al.*, 2013), although with a different ligand orientation. Because the integrin structure was too large to be treated at once, its surface was split into 46 overlapping areas corresponding to 46 Autodock grids with a volume of $47.25 \times 47.25 \times 47.25$ Å^3 and a node spacing of 0.375 Å. The distance between grid centers was half the grid length, that is, 22.5 Å. This overlapping coverage of the integrin surface was necessary to avoid boundary effects between two neighbor grids.

The methylsulfate and the HS_6 oligosaccharides from the library of conformers were treated as rigid during the calculation. For HS_2 compounds, only the glycosidic linkage and the ring conformation were considered as rigid.

Sequence alignments

In order to determine if the amino acid residues predicted to interact with HP/HS are conserved in integrins, the sequences of all human integrin subunits retrieved from the UniprotKB database were aligned using T-COFFEE (Notredame *et al.*, 2000). The results were visualized using the Multiseq plugin of VMD (Roberts *et al.*, 2006). The sequences of human integrins (The UniProt Consortium, 2012) were divided into two groups corresponding to α and β subunits.

RESULTS

Exploration of the integrin surface with a methyl sulfate probe

The determination of the preferred localization of sulfate anions on the surface of protein, either by crystal structure analysis (Lortat-Jacob *et al.*, 2002, Gandhi *et al.*, 2008) or by GRID analysis (Ricard-Blum *et al.*, 2004) has been shown to be an efficient way to determine possible HP sulfate binding sites. The few crystal structures of the ectodomain of the αvβ3 integrin available in the PDB database do not contain sulfate ions. Therefore, in a first attempt to find the potential sites of interaction between the ectodomain of the αvβ3 integrin and HP/HS, we performed a

Table 1. List of the heparin oligosaccharides extracted from the PDB entry 1HPN (Mulloy *et al.*, 1993). The reducing end is indicated by "–OH"

Name	Type	Oligosaccharide	Ring conformation of the GlcNS units	Ring conformation of the IdoAS units
IG : ^2S_O	HS_2	IdoAS-GlcNS–OH	$^4C_1$	^2S_O
IG : $^1C_4$	HS_2	IdoAS-GlcNS–OH	$^4C_1$	$^1C_4$
GI : ^2S_O	HS_2	GlcNS-IdoAS–OH	$^4C_1$	^2S_O
GI : $^1C_4$	HS_2	GlcNS-IdoAS–OH	$^4C_1$	$^1C_4$
HS_6 : ^2S_o	HS_6	IdoAS-GlcNS-IdoAS-GlcNS-IdoAS-GlcNS–OH	$^4C_1$	^2S_O
HS_6 : $^1C_4$	HS_6	IdoAS-GlcNS-IdoAS-GlcNS-IdoAS-GlcNS–OH	$^4C_1$	$^1C_4$

docking simulation of methyl sulfate ($MeSO_4^-$) to mimic the sulfate moieties of HS. This probe was docked to integrin αvβ3 using the 46 overlapping grids covering the ectodomain of the integrin and the best docking score was selected for each grid. The 46 results were distributed into approximately 20 spots on the protein surface according to the amino acids in contact (Figure 2(A)). The Autodock scores ranged between −2.11 and −0.15 kcal mol^{-1}, the best scores being shared by two spots only labeled A and B in Figure 2(A) (see also the Supplementary Tables S2 and S3, Figure S2). Spot A was centered on residues K^{645} and K^{646} of the αv subunit, and spot B on the residue R^{65} of the same subunit. They both corresponded to a highly electropositive patch on the protein surface (Figure 2(B)), which was close to another electropositive patch. This suggested that a larger electronegative molecule, like the sulfated domains of HS chains, could potentially interact with the ectodomain of the αvβ3 integrin at those locations. A longer probe, reflecting the size of an HS oligosaccharide, was thus needed.

Determination of hot spots compatible with HP/HS binding

In order to refine the results obtained with the methylsulfate probe, a series of four HS_2 were used as ligands to explore the integrin surface. They correspond to disaccharide repeats found in the sulfated domains of HS chains and took into account the two main ring shapes of iduronic acid. Each HS_2 probe was docked using the 46 grids covering the integrin surface. A total of 184 (4 × 46) docking results was obtained, and the best score was selected for each grid. The Autodock scores ranged from −4.26 to +1.29 kcal mol^{-1} and were distributed over approximately 20 clusters (Figure 3(A) and Supplementary Table S4). Most clusters were in the same location than those identified with the methyl sulfate probe, but they involved a larger number of amino acid residues. However, in contrast to the results obtained with the methyl sulfate probe, the best docking results were shared by six clusters, defining six spots on the integrin surface. Two of them (spots V and II) corresponded to spots A and B reported earlier, and four new ones were identified, including one on the β3 subunit. Each of the six spots (Table 2 and Figure 3) comprises one to four basic residues, most often lysine residue(s). The disaccharide–integrin interactions were mostly mediated by direct contacts between sulfate moieties and lysine and/or arginine residues (Figure 3(B)). The best scores (spots I, III, and VI) were obtained for the binding of HS_2 to the integrin via two sulfate groups. For each spot, the docking results of

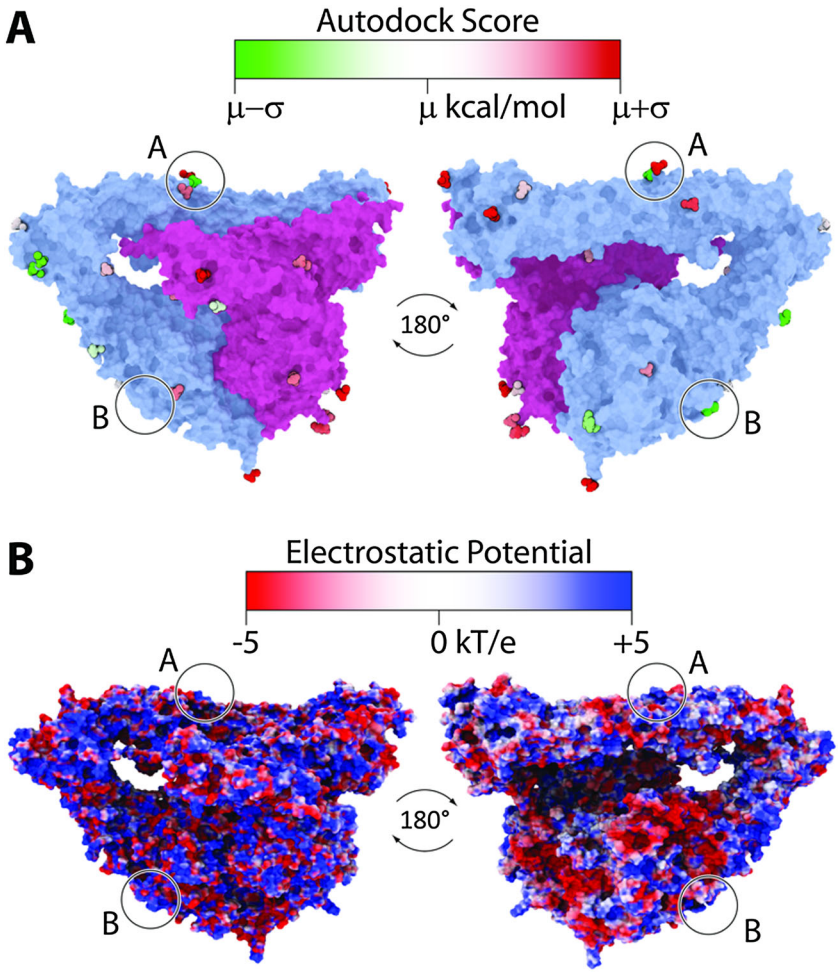

Figure 2. Summary of the docking results obtained with $MeSO4^-$. (A) The 46 $MeSO4^-$ compounds docked on the ectodomain of αvβ3 integrin (1JV2). The αv and β3 subunits are respectively blue and mauve. The $MeSO4^-$ compounds are colored as a function of their deviation from the mean Autodock score μ (= −1.28 kcal mol^{-1}) from green (≤ −1.72 kcal mol^{-1}) to red (≥ −0.85 kcal mol^{-1}). (B) Electrostatic potential of the αvβ3 ectodomain. The positions of the best docking results are circled in black (score ≤ −1.72 kcal mol^{-1}).

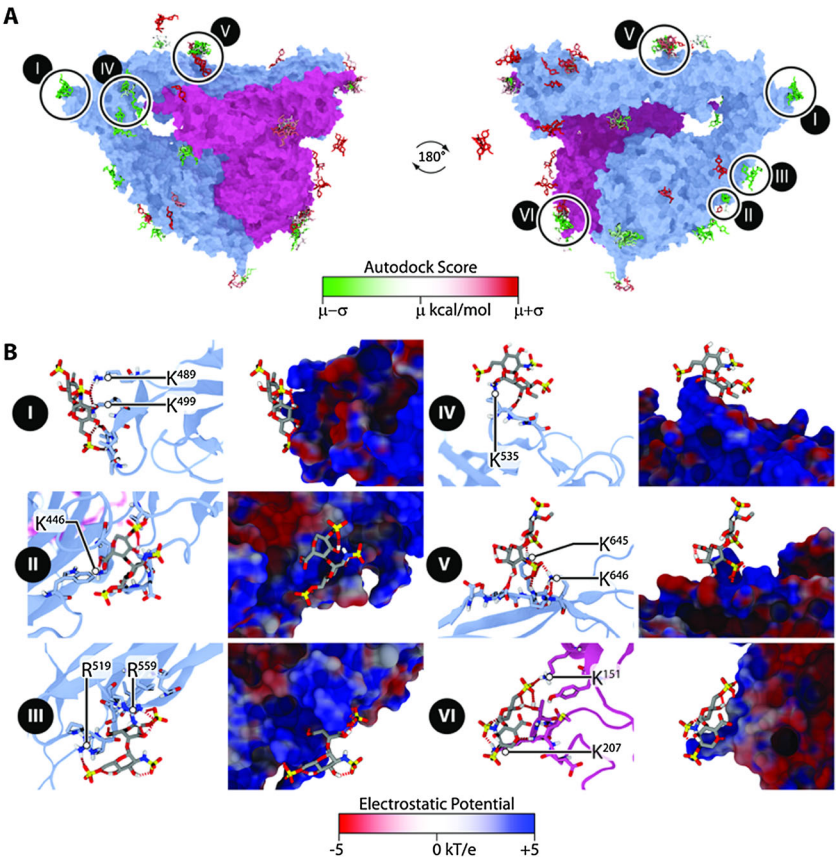

Figure 3. Summary of the docking results obtained with HS2. (A) The 184 heparin disaccharide structures docked on the ectodomain of αvβ3 integrin (1JV2). The αv and β3 subunits are respectively blue and mauve. The HS2 compounds are colored from green (≤ -2.59 kcal mol^{-1}) to red (≥ -0.48 kcal mol^{-1}) as a function of their deviation from the mean Autodock score μ ($= -1.54$ kcal mol^{-1}). The positions of the six best spots are circled in black. (B) Details of the best docking result obtained for each spot. On the left panel is shown the local structure of the protein. The HS2 compound and the amino acids within 3 Å are fully represented; the rest of the protein is shown in cartoon representation. The polar interactions between the protein and the ligand are depicted as a dashed red line. Charged basic residues are indicated for each spot. On the right panel is shown the local electrostatic potential on the integrin surface.

Table 2. List of the best docking results obtained with HS_2 over the 46 grids, that is, with an Autodock score lower than the mean minus one standard deviation (-2.59 kcal mol^{-1}). For each grid, the result is averaged over the four HS_2 oligosaccharides. The amino acids within 3 Å of the compounds are also indicated

Grid	Best score	Mean score $\pm\sigma$	Subunit	Amino acids in contact with HS_2	Spot
46	−4.00	-3.56 ± 0.31	αv	K^{489}, T^{490}, T^{496}, A^{497}, L^{498}, K^{499}	I
40	−3.99	-3.52 ± 0.35	αv	K^{489}, T^{490}, T^{496}, A^{497}, L^{498}, K^{499}	I
45	−3.64	-3.44 ± 0.19	αv	K^{489}, L^{498}, K^{499}	I
32	−3.37	-2.73 ± 0.38	αv	S^{61}, S^{62}, R^{63}, F^{65}, K^{88}, W^{91}, R^{96}, K^{446}, A^{599}, D^{600}, T^{601}	II
39	−3.62	-3.19 ± 0.42	αv	R^{519}, K^{520}, S^{558}, R^{559}, L^{562}	III
36	−3.24	-2.76 ± 0.33	αv	Q^{534}, K^{535}, A^{537}	IV
10	−2.98	-2.68 ± 0.27	αv	D^{643}, Q^{644}, K^{645}, K^{646}	V
16	−4.26	-3.06 ± 0.73	β3	Y^{148}, K^{151}, M^{206}, K^{207}, T^{208}	VI
$\mu \pm \sigma$		-1.54 ± 1.05			

the clusters correspond to rather different orientation of the disaccharide on the surface. This was not due to the heterogeneity of the sugar ring conformation, but rather to the orientation of the disaccharide on the integrin surface because of the absence of physical constraints limiting the orientation of the disaccharide.

Refinement of hot spots and identification of amino acids interacting with an HS_6

To refine the predictions made with the disaccharide probes and to identify the amino acids, which interact with HP on each spot, further docking experiments were performed with HS_6. We have

previously shown that a hexasaccharide inhibited the binding of one RGD-binding integrin to HP and HS by 30 and 41%, respectively (Faye *et al.*, 2009).

The number of grids was reduced to the six spots identified earlier. They were centered on six lysine residues, K^{489}, K^{446}, K^{520}, K^{535}, K^{646}, and K^{151} for spots I to VI, respectively (Figure 4(A)). A set of 40 conformations of HS_6 was sampled from two independent molecular dynamics simulation of HS_6 (with different ring shape for the iduronic acid residue) in explicit water to better take into account its intrinsic flexibility during the rigid–rigid docking calculation. The 40 conformations were docked using the six grids, giving a total of 240 docking results. The Autodock score ranged from −1.97 to +1.65 kcal mol^{-1}. Again, for a same spot, different orientations of the hexasaccharide ligand were reported (Figure 4(B)). However, the distribution of the score is narrow and approximately the same for the six spots (ca. 0.25 kcal mol^{-1}). The best scores were obtained for spots I and III (Table 3 and Figure 4(A)). Spot I was located at the top of the genu domain and spots II and III on the thigh domain on the α_V subunit (Figure 5(A)). The combination of spots II, III, and IV encompasses both sides of the αv subunit (Figure 5(A)), which could bind a single HP molecule, assuming a twist of the chain (Figure 5(A)). Spot VI was located on the head region of the β subunit. The worst scores were obtained for spot V, which is located on the Calf-I domain on the α subunit. None of these spots encompasses a known calcium binding site or an N-glycosylation site. Noteworthy, the interaction between HS_6 and spots I, II, III, V, or VI was established with three to four consecutive monosaccharide units for the best docking models (Figure 4(A) and Supplementary Table S2). The interaction was mediated by the two first and the two last monosaccharide units on spot IV.

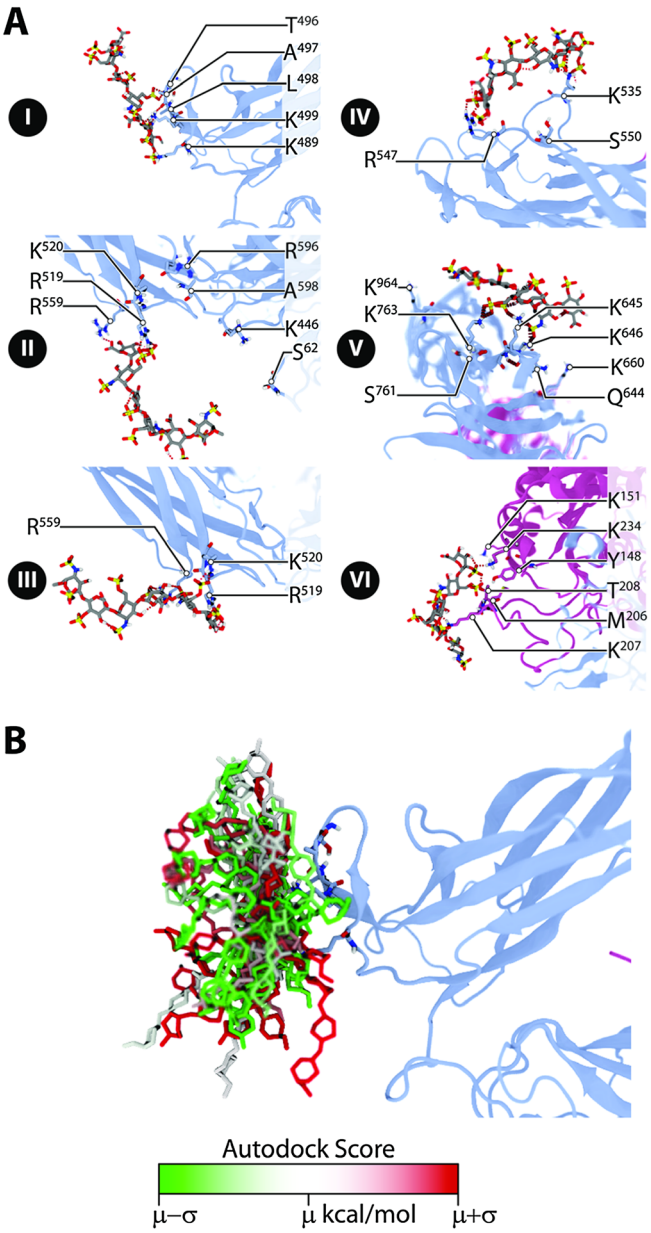

Figure 4. Summary of the docking results obtained with HS6. (A) Details of the best docking results obtained for each spot. The αv and β3 subunits are blue and mauve, respectively. The non-bonded polar interactions are represented as dashed red lines. (B) Superimposition of the 40 docking results obtained for spot I. HS6 compounds are colored from green (≤ −1.67 kcal mol^{-1}) to red (≥ −1.20 kcal mol^{-1}) as a function of their deviation from the mean Autodock score 1 (= −1.44 kcal mol^{-1}).

Analysis of basic amino acid conservation in integrin sequences

As observed for methylsulfate and HS_2, the binding of the integrin subunits to HS_6 was mostly mediated by direct contacts between basic amino acid residues and the negatively charged moieties of the oligosaccharide (Table 3). Furthermore, the amino acid residues interacting with the HS_6 were the same than those interacting with the HS_2. To determine if these amino acid residues were conserved in all the HP-binding integrins, we aligned the sequences of human integrins (Supplementary Tables S5 and S6). Among the residues of the αv subunit predicted to bind with HS_2 and HS_6, only the basic residues (K/R) in position 489 on spot I and in position 535 on spot IV were strictly conserved in the four RGD-dependent integrins binding to HP (αvβ3, αvβ5, α5β1, and $\alpha IIb\beta3$) and in the α8β1 RGD-dependent integrin that has not been reported to bind HP so far. The residue R^{547} was conserved only in the αv and α5 subunits. The single strictly conserved residue in the β subunit of RGD-dependent integrins binding to HP was the lysine residue 151 on spot VI. Globally, the most conserved residues in the integrin family were K^{489}, R^{547}, R^{559}, and K^{763} located on the αv subunit and K^{151} and K^{234} located on the β3 subunit (Figure 5(B)). The amino acid residues predicted to contact HP/HS and conserved in the α and β subunits of the HP-binding RGD-dependent integrins (K/R^{489} and K^{151}, respectively) are replaced by a hydrophobic residue (F^{489}) in αIIb, αM, and αX subunits and in β2 subunit (L^{151}, Supplementary Tables S5 and S6). Basic residues found in RGD-binding integrins were thus not conserved in non-RGD HP-binding integrins, the leukocyte-specific receptors α4β1, αXβ2, and αMβ2. Only one basic (K^{547}) residue, potentially corresponding to spot IV, was present in the αX and αM chains (Supplementary Tables S5 and S6). These results suggest that HP/HS could bind to different sites on the RGD-dependent integrins and on the RGD-independent leukocyte-specific integrins.

HP and protein ligand binding sites on integrins

Integrins serve as receptors for extracellular proteins. It is thus interesting to determine if the potential HP binding sites on integrins are close or overlap the binding sites of the protein

Table 3. Summary of the docking results obtained with for the six selected spots. For each spot, the result is averaged over the 40 HS_6 conformations. The amino acids within 3 Å of the oligosaccharide are also indicated. The percentage in parenthesis represents the number of times a residue is in contact with HS_6, compared with all the other residues of the spot. For example, K^{499} is in contact with HS_6 in 36% of the 184 docking simulations achieved on spot I

Spot	Best score	Mean score $\pm \sigma$	Subunit	Amino acids in contact with HS_6
I	−1.97	-1.44 ± 0.23	αv	K^{499} (36%), K^{489} (33%), A^{497} (12%), L^{498} (10%), T^{496} (6%), others (3%)
II	−0.55	-0.22 ± 0.28	αv	R^{519} (44%), K^{520} (19%), R^{559} (16%), S^{62} (3%), K^{446} (3%), R^{596} (3%), A^{598} (3%), others (9%)
III	−1.71	-0.96 ± 0.31	αv	R^{519} (33%), R^{559} (32%), K^{520} (30%), others (5%)
IV	−1.05	-0.43 ± 0.32	αv	K^{535} (53%), R^{547} (29%), S^{550} (3%), others (15%)
V	0.66	1.20 ± 0.23	αv	K^{645} (43%), K^{646} (18%), K^{763} (14%), S^{761} (4%), Q^{644} (3%), K^{660} (3%), K^{964} (3%), others (12%)
VI	−0.05	0.62 ± 0.32	β3	K^{207} (58%), T^{208} (13%), K^{151} (8%), K^{234} (8%), M^{206} (6%), Y^{148} (3%), others (4%)
$\mu \pm \sigma$		-0.20 ± 0.98		

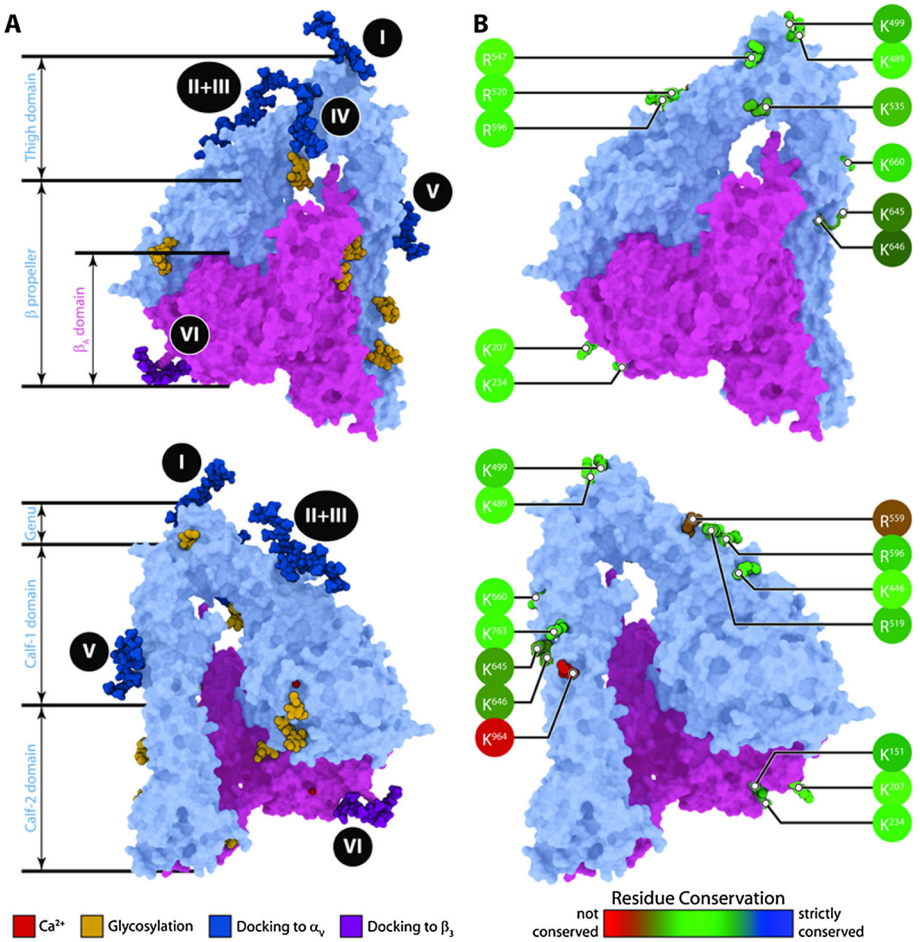

Figure 5. Overview of the best docking results obtained with HS6. (A) Overview of six spots on the 3D structure of the ectodomain of the αvβ3 integrin. The αv and β3 subunits are blue and mauve, respectively. (B) Residue conservation within the human integrin sequences of the basic amino acids of each spot. Amino acids are colored as a function of the sequence similarity within the multiple sequence alignments of the α and β subunits based on the BLOSUM60 matrix.

ligands of integrins. When we superimposed the docking results obtained with HS_6 on the 3D structure of the ectodomain of αvβ3 associated with a cyclic peptide containing the RGD motif (PDB entry 1L5G; Xiong *et al*., 2002), the spot VI and the RGD ligand appear to be close, but no steric clash was observed with the HS_6 (Figure 6). It should also be true for the other

Journal of Molecular Recognition

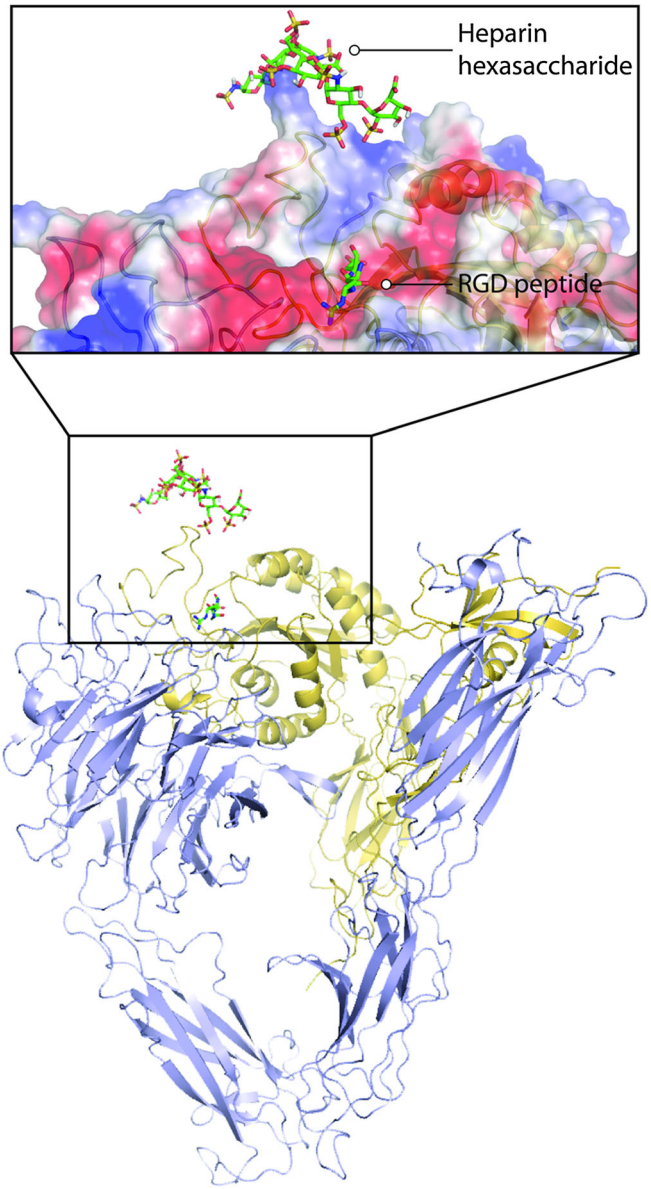

Figure 6. Predicted structure of a ternary complex formed by the αvβ3 integrin (αv, blue; β3, gold), an heparin hexasaccharide and a ligand of integrin represented by an RGD peptide. The structure of the integrin is the crystal structure (1JV2) used in this study for molecular modeling. The position of the heparin hexasaccharide corresponds to the best Autodock score for the spot VI. RGD coordinates have been obtained from the structure of integrin αvβ3 co-crystallized with a cyclic RGD peptide (1L5G, Xiong *et al*., 2002).

RGD-dependent integrins because the RGD peptide interacts in a similar manner with both the α and β subunits in the three available structures of the ectodomain of integrin in complex with an RGD peptide (Nagae *et al*., 2012). The arginine residue of the peptide forms hydrogen bonds with an aspartate or a glutamine residue of the α subunits (Supplementary Table S7). The aspartate residue of the peptide coordinates the magnesium ion of metal ion dependent adhesion site found in the β subunits. In metal ion dependent adhesion site, the two serine and the glutamate residues coordinating magnesium are very well conserved between all the β subunits of the RGD binding integrins.

This suggests the possible formation of ternary complexes between HS, the αvβ3 integrin and a protein ligand (e.g., endostatin) on the head of the integrin. HS could play several roles via its interaction with integrins. It could bring the extracellular ligand in close proximity to the ligand binding site on the integrin headpiece without playing any role in the activation of the receptor leading to its high affinity conformation. It could also contribute to integrin activation in specific conditions. This is supported by the fact that the syndecan-1-dependent activation of the αvβ3 integrin, which is dependent on syndecan-1 binding to a ligand, requires or not the presence of HS chains of the syndecan-1 depending on the ligand (Beauvais *et al*., 2004).

DISCUSSION

We have previously shown that αvβ3 integrin binds to HP and HS, and that this interaction mediates cell spreading (Faye *et al*., 2009). We report here the characterization of this interaction at the molecular level using three series of molecular docking calculation. The entire surface of the integrin ectodomain was first explored with methyl sulfate and four HS_2 corresponding to fragments of HS sulfated domain. The docking results were ranked and six spots were predicted to be the most likely HS binding sites on the integrin surface. Five of them are located on the α subunit and one on the β subunit of the integrin. They contain at least one lysine or arginine residue interacting with the sulfate moiety of the disaccharides. This last result is in agreement with the inhibition by HP (100%), but not by the carboxymethyl dextran lacking sulfate groups, of the binding of the α5β1 ectodomain to HP (Faye *et al*., 2009). All spots are compatible with available biochemical information about integrin surface because they do not overlap with calcium binding sites or an N-glycosylation sites (Schlessinger *et al*., 2000, Baglin *et al*., 2002, Carter *et al*., 2005, Murphy *et al*., 2007, Yang *et al*., 2010).

Each predicted binding spot was submitted to supplemental docking calculation with an HS_6 in order to obtain a binding model. Hexasaccharides are usually sufficient to promote the binding to growth factors such as fibroblast growth factor-2 (Schlessinger *et al*., 2000), CXCL12α (stromal cell derived factor 1α; Murphy *et al*., 2007), and hepatocyte growth factor (Lietha *et al*., 2001). The docking results did not converge toward a unique binding mode for each of the six spots. They gave a set of results reflecting different binding possibilities but lysine or arginine residues were clearly involved in the binding of HP oligosaccharides. This absence of convergence is likely due to the lack of physical constraints able to orient the oligosaccharides in a particular direction on the integrin surface. Indeed, the six spots have all a good accessibility to the solvent. This is in agreement with available structural data on HP binding sites on soluble proteins (Imberty *et al*., 2007, Gandhi and Mancera, 2008, Esko and Linhardt, 2009). Such sites do not consist of deep grooves, but of flat surfaces or very shallow grooves lined by at least four (Schlessinger *et al*., 2000) and up to 16 (Murphy *et al*., 2007) lysine or arginine residues. A careful inspection of the αvβ3 integrin electrostatic surface did not reveal the presence of a basic groove, although there were many electropositive areas, and a single spot includes two to five basic residues at most for integrins (Table 3). This suggests that HP or HS binding site on integrins would not be continuous, but distributed on several discrete locations. The identification of spots II and III and possibly of spots IV and V is in agreement with this hypothesis.

In vivo RGD-binding integrins would not bind to a single sulfated domain of HS, but too few of them separated by non-sulfated regions, which have an enhanced flexibility (Khan *et al*., 2010, Khan *et al*., 2011) and would allow a bending of the glycosaminoglycan chain.

Seven integrins (αvβ3, αvβ5, αIIbβ3, α5β1, α4β1, αXβ2, and αMβ2) have been shown to interact with HP and/or HS (Sobel and Adelman, 1988, Diamond *et al*., 1995, Vorup-Jensen *et al*., 2005, Vorup-Jensen *et al*., 2007, Faye *et al*., 2009). According to our results, HP would mostly bind to the top of the integrin genu domain, to the Calf-I domain of the α subunit, and on the top of the β subunit of RGD-dependent integrins. Recently, the HP binding site on αIIbβ3 was mapped by surface plasmon resonance in two regions of the αIIb subunit, the N-terminal half of the head domain and in the segment containing the thigh and calf domains (Yagi *et al*., 2012). No binding has been reported for the C-terminal half of the head region, the N-terminal two thirds of the thigh domain, or for the β3 subunit. Our docking simulations did not predict any binding site in the β-propeller forming the head domain of αvβ3, spot I being located in the genu part, spots II to IV in the thigh domain and spot V in the Calf-I domain. However, some clusters of basic amino acids are present in the genu part and in the Calf-I domain of the αIIbβ3 integrin (Yagi *et al*., 2012), which are close to the location of spots I (the best docking score for HS_6) and V on the αvβ3 structure. The RGD-binding integrins could have a common HS binding site in this part of the protein. It is worth noting that among the amino acid residues in contact with HP/HS in spot I on αvβ3 integrin, one basic amino acid (K/R^{489}) is conserved in the sequences of the α subunit of all RGD- and HP-binding integrins. Another amino acid residue (K^{151}) is strictly conserved in the β subunit of the RGD-binding and HP-binding integrins. These residues are not conserved in α4β1, αXβ2, and αMβ2, which bind HP but do not recognize the RGD sequence in their ligands. Basic amino acids residues are replaced by a hydrophobic residue both in αIIb, αM, and αX integrin subunits (F^{489}) and in β2 subunit (L^{151}), suggesting that HP/HS binds to different sites on the two subgroups of integrins. This is in agreement with the binding of HP to the αI domain of αXβ2 and αMβ2 integrins, which is absent in the αv and αIIb chains (Vorup-Jensen *et al*., 2007). Because mechanisms of activation are not generic for all integrins (Margadant *et al*., 2011) HS could play different roles in their crosstalk with different integrins and/or integrin subgroups. The characterization of the ternary complex between integrin, endostatin, and HS at the surface of activated endothelial cells would provide a better understanding of the molecular mechanisms underlying the anti-angiogenic activities of endostatin. It would also be of interest to determine if other ternary complexes regulating angiogenesis (e.g., transglutaminase 2 - HS - integrin and/or MMP 2 - HS - integrin) form at the surface of activated endothelial cells. Several other protein fragments, cleaved from extracellular proteins and exhibiting anti-angiogenic properties, bind to integrins (Ricard-Blum and Ballut, 2011) and could also be involved in the formation of ternary complexes.

CONCLUSION

In conclusion, HP would mostly bind to the top of the genu domain of αvβ3 integrin, to the Calf-I domain of the αv subunit, and on the top of the β3 subunit of RGD-dependent integrins. Because HP/HS does not bind to the same integrin site than protein ligands, ternary complexes comprised of integrin, one ligand and HP/HS could form *in vivo*. The basic amino acid residues predicted to interact with HP are conserved in the sequences of human RGD-dependent integrins, but not in RGD-independent integrins suggesting that HP/HS could bind to different sites on these two integrin subfamilies. The role played by HS in the activation of integrins warrants further investigation.

Acknowledgements

This work was funded by IXXI 2011 (Institut Rhône-Alpin des Systèmes Complexes) to SRB, by the grant ANR-07-NANO-050-03 to LB and SRB and by the "Pôle de Compétitivité MEDICEN" as part of the DOSCA FUI project to NS and AI. The authors thank Dr. Xavier Robert (UMR 5086 CNRS - Université Lyon 1, France) for providing free computational time.

REFERENCES

Adair BD, Xiong JP, Maddock C, Goodman SL, Arnaout MA, Yeager M. 2005. Three-dimensional EM structure of the ectodomain of integrin αVβ3 in a complex with fibronectin. *J. Cell Biol.* **168**: 1109–1118.

Askari JA, Buckley PA, Moul AP, Humphries MJ. 2009. Linking integrin conformation to function. *J. Cell Sci.* **122**: 165–170.

Baglin TP, Carrell RW, Church FC, Esmon CT, Huntington JA. 2002. Crystal structures of native and thrombin-complexed heparin cofactor II reveal a multistep allosteric mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **99**: 11079–11084.

Barczyk M, Carracedo S, Gullberg D. 2010. Integrins. *Cell Tissue Res.* **339**: 269–280.

Beauvais DM, Burbach BJ, Rapraeger AC. 2004. The syndecan-1 ectodomain regulates alphavbeta3 integrin activity in human mammary carcinoma cells. *J. Cell Biol.* **167**: 171–181.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. *Nucleic Acids Res.* **28**: 235–242.

Bitomsky W, Wade RC. 1999. Docking of glycosaminoglycans to heparin binding proteins: validation for aFGF, bFGF, and antithrombin and application to IL-8. *J. Am. Chem. Soc.* **121**: 3004–3013.

Bülow HE, Hobert O. 2006. The molecular diversity of glycosaminoglycans shapes animal development. *Annu. Rev. Cell Dev. Biol.* **22**: 375–407.

Carter WJ, Cama E, Huntington JA. 2005. Crystal structure of thrombin bound to heparin. *J. Biol. Chem.* **280**: 2745–2749.

Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**: 10089–10092.

Diamond MS, Alon R, Parkos CA, Quinn MT, Springer TA. 1995. Heparin is an adhesive ligand for the leukocyte integrin Mac-1 (CD11b/CD1). *J. Cell Biol.* **130**: 1473–1482.

Esko JD, Linhardt RJ. 2009. Proteins that bind sulfated glycosaminoglycans. In *Essentials of glycobiology*. Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (eds), 2nd ed, Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY).

Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**: 8577–8593.

Faye C, Moreau C, Chautard E, Jetne R, Fukai N, Ruggiero F, Humphries MJ, Olsen BR, Ricard-Blum S. 2009. Molecular interplay between endostatin, integrins, and heparan sulfate. *J. Biol. Chem.* **284**: 22029–22040.

Forster M, Mulloy B. 2006. Computational approaches to the identification of heparin-binding sites on the surfaces of proteins. *Biochem. Soc. Trans.* **34**: 431–434.

Gandhi NS, Mancera RL. 2008. The structure of glycosaminoglycans and their interactions with proteins. *Chem. Biol. Drug Des.* **72**: 455–482.

Gandhi NS, Coombe DR, Mancera RL. 2008. Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 1. Molecular modeling studies. *Biochemistry* **47**: 4851–4862.
Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**: 712–725.
Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. *J. Mol. Graphics* **14**: 33–38, 27–28.
Hynes RO. 2007. Cell-matrix adhesion in vascular development. *J. Thromb. Haemostasis Suppl* **1**: 32–40.
Imberty A, Pérez S. 2006. Molecular modeling of glycosaminoglycans and interactions with protein receptors – methods and progress. In *New developments in therapeutic glycomics*. Delehedde M, Lortat-Jacob H (eds), Research Signpost, Trivandrum India; 185–201.
Imberty A, Lortat-Jacob H, Pérez S. 2007. Structural view of glycosaminoglycan-protein interactions. *Carbohydr. Res.* **342**: 430–439.
Khan S, Gor J, Mulloy B, Perkins SJ. 2010. Semi-rigid solution structures of heparin by constrained X-ray scattering modelling: new insight into heparin-protein complexes. *J. Mol. Biol.* **395**: 504–521.
Khan S, Rodriguez E, Patel R, Gor J, Mulloy B, Perkins SJ. 2011. The solution structure of heparan sulfate differs from that of heparin: implications for function. *J. Biol. Chem.* **286**: 24842–24854.
Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, Woods RJ. 2008. GLYCAM06: a generalizable biomolecular force field. *Carbohydrates. J. Comput. Chem.* **29**: 622–655.
Lietha D, Chirgadze DY, Mulloy B, Blundell TL, Gherardi E. 2001. Crystal structures of NK1-heparin complexes reveal the basis for NK1 activity and enable engineering of potent agonists of the MET receptor. *EMBO J.* **20**: 5543–5555.
Lortat-Jacob H, Grosdidier A, Imberty A. 2002. Structural diversity of heparan sulfate binding domains in chemokines. *Proc. Nat. Acad. Sci. U. S. A.* **99**: 1229–1234.
Mahoney DJ, Mulloy B, Forster J, Blundell CD, Fries E, Milner CM, Day AJ. 2005. Characterization of the interaction between tumor necrosis factor-stimulated gene-6 and heparin: implications for the inhibition of plasmin in extracellular matrix microenvironments. *J. Biol. Chem.* **280**: 27044–27055.
Margadant C, Monsuur HN, Norman JC, Sonnenberg A. 2011. Mechanisms of integrin activation and trafficking. *Curr. Opin. Cell Biol.* **23**: 607–614.
Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. 1998. Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. *J. Comput. Chem.* **19**: 1639–1662.
Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. 2009. AutoDock4 And AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**: 2785–2791.
Mulloy B, Forster MJ, Jones C, Davies DB. 1993. N.m.r. and molecular-modelling studies of the solution conformation of heparin. *Biochem. J.* **293**: 849–858.
Murphy JW, Cho Y, Sachpatzidis A, Fan C, Hodsdon ME, Lolis E. 2007. Structural and functional basis of CXCL12 (stromal cell-derived factor-1 alpha) binding to heparin. *J. Biol. Chem.* **282**: 10018–10027.
Nagae M, Re S, Mihara E, Nogi T, Sugita Y, Takagi J. 2012. Crystal structure of α5β1 integrin ectodomain: atomic detail of the fibronectin receptor. *J. Cell Biol.* **197**: 131–140.
Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**: 205–217.
Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**: 1781–1802.
Raghuraman A, Mosier PD, Desai UR. 2006. 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.
R-development-core-team. 2005. R: a language and environment for statistical computing. Vienna.
Rehn M, Veikkola T, Kukk-Valdre E, Nakamura H, Ilmonen M, Lombardo C, Pihlajaniemi T, Alitalo K, Vuori K. 2001. Interaction of endostatin with integrins implicated in angiogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **98**: 1024–1029.
Ricard-Blum S, Ballut L. 2011. Matricryptins derived from collagens and proteoglycans. *Front. Biosci., Landmark Ed.* **16**: 674–697.
Ricard-Blum S, Féraud O, Lortat-Jacob H, Rencurosi A, Fukai N, Dkhissi F, Vittet D, Imberty A, Olsen BR, van der Rest M. 2004. Characterization of endostatin binding to heparin and heparan sulfate by surface plasmon resonance and molecular modeling: role of divalent cations. *J. Biol. Chem.* **279**: 2927–2936.
Roberts E, Eargle J, Wright D, Luthey-Schulten Z. 2006. Multiseq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinf.* **7**: 382.
Sadir R, Baleux F, Grosdidier A, Imberty A, Lortat-Jacob H. 2001. Characterization of the stromal cell-derived factor-1alpha-heparin complex. *J. Biol. Chem.* **276**: 8288–8296.
Sapay N, Cabannes É, Petitou M, Imberty A. 2011. Molecular modeling of the interaction between heparan sulfate and cellular growth factors: bringing pieces together. *Glycobiology* **21**: 1181–1193.
Sapay N, Nurisso A, Imberty A. 2013. Simulation of carbohydrates: from molecular docking to molecular dynamics simulation. *Methods Mol. Biol.* **924**: 469–483.
Schlessinger J, Plotnikov AN, Ibrahimi OA, Eliseenkova AV, Yeh BK, Yayon A, Linhardt RJ, Mohammadi M. 2000. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol. Cell* **6**: 743–750.
Sobel M, Adelman B. 1988. Characterization of platelet binding of heparins and other glycosaminoglycans. *Thromb. Res.* **50**: 815–826.
The Uniprot Consortium. 2012. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucl Acids Res.* **40**: D71–D75.
Vorup-Jensen T, Carman CV, Shimaoka M, Schuck P, Svitel J, Springer TA. 2005. Exposure of acidic residues as a danger signal for recognition of fibrinogen and other macromolecules by integrin alphaXbeta2. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 1614–1619.
Vorup-Jensen T, Chi L, Gjelstrup LC, Jensen UB, Jewett CA, Xie C, Shimaoka M, Linhardt RJ, Springer TA. 2007. Binding between the integrin alphaXbeta2 (CD11c/CD18) and heparin. *J. Biol. Chem.* **282**: 30869–30877.
Wickham H. 2009. *ggplot2: elegant graphics for data analysis*; Springer: New York.
Xiong J-P, Stehle T, Diefenbach B, Zhang R, Dunker R, Scott DL, Joachimiak A, Goodman SL, Arnaout MA. 2001. Crystal structure of the extracellular segment of integrin alphaVbeta3. *Science* **294**: 339–345.
Xiong J-P, Stehle T, Zhang R, Joachimiak A, Frech M, Goodman SL, Arnaout MA. 2002. Crystal structure of the extracellular segment of integrin alphaVbeta3 in complex with an Arg-Gly-Asp ligand. *Science* **296**: 151–155.
Yagi M, Murray J, Strand K, Blystone S, Interland G, Suda Y, Sobel M. 2012. Heparin modulates the conformation and signaling of platelet integrin αIIbβ3. *Thromb. Res.* **129**: 743–749.
Yang IS, Kim TG, Park BS, Cho KJ, Lee J-H, Park Y, Kim KH. 2010. Crystal structures of aprotinin and its complex with sucrose octasulfate reveal multiple modes of interactions with implications for heparin binding. *Biochem. Biophys. Res. Commun.* **397**: 429–435.

