BIOCHEMISTRY

 Volume 47, Number 17 April 29, 2008

Articles

Platelet Endothelial Cell Adhesion Molecule 1 (PECAM-1) and Its Interactions with Glycosaminoglycans: 1. Molecular Modeling Studies[†]

Neha S. Gandhi,[‡,§] Deirdre R. Coombe,[‡,§] and Ricardo L. Mancera*[,‡,§,||]

Western Australian Biomedical Research Institute, and Schools of Pharmacy and Biomedical Sciences, Curtin University of Technology, GPO Box U1987, Perth, Western Australia 6845, Australia

Received December 17, 2007; Revised Manuscript Received January 24, 2008

ABSTRACT: Platelet endothelial cell adhesion molecule 1 (PECAM-1) has many functions, including its roles in leukocyte extravasation as part of the inflammatory response and in the maintenance of vascular integrity through its contribution to endothelial cell−cell adhesion. PECAM-1 has been shown to mediate cell−cell adhesion through homophilic binding events that involve interactions between domain 1 of PECAM-1 molecules on adjacent cells. However, various heterophilic ligands of PECAM-1 have also been proposed. The possible interaction of PECAM-1 with glycosaminoglycans (GAGs) is the focus of this study. The three-dimensional structure of the extracellular immunoglobulin (Ig) domains of PECAM-1 were constructed using homology modeling and threading methods. Potential heparin/heparan sulfate-binding sites were predicted on the basis of their amino acid consensus sequences and a comparison with known structures of sulfate-binding proteins. Heparin and other GAG fragments have been docked to investigate the structural determinants of their protein-binding specificity and selectivity. The modeling has predicted two regions in PECAM-1 that appear to bind heparin oligosaccharides. A high-affinity binding site was located in Ig domains 2 and 3, and evidence for a low-affinity site in Ig domains 5 and 6 was obtained. These GAG-binding regions were distinct from regions involved in PECAM-1 homophilic interactions.

Platelet endothelial cell adhesion molecule 1 (PECAM-1)[1]/cluster determinant 31 (CD31) is a member of the cell adhesion molecule (CAM) subgroup of the immunoglobulin (Ig) superfamily and is expressed on the surface of circulating platelets, monocytes, neutrophils, and certain T-cell subsets (*1*). PECAM-1 is implicated in numerous functions, including the transendothelial migration of leukocytes (diapedesis), angiogenesis, integrin activation, suppression of apoptosis, and platelet aggregation (*1*). PECAM-1 has been mapped to region 17q23 on chromosome 17. Its mRNA is highly

[†] Part of this work is funded by the award of a Curtin Internal Grant.

* To whom correspondence should be addressed. Telephone: +61-8-9266-1017. Fax: +61-8-9266-3245. E-mail: r.mancera@curtin.edu.au.

[‡] Western Australian Biomedical Research Institute.

[§] School of Biomedical Sciences.

[||] School of Pharmacy.

[1] Abbreviations: PECAM-1, platelet endothelial cell adhesion molecule 1; CAM, cell adhesion molecule; Ig, immunoglobulin; CD31, cluster determinant 31; GAG, glycosaminoglycan; ITIM, immunoreceptor tyrosine-based inhibitory motif; IdoA, α-L-iduronic acid; GlcA, *β*-D-glucuronic acid; GlcA2S, 2-*O*-sulfated glucuronic acid; IdoA2S, 2-*O*-sulfated iduronic acid; GlcN, D-glucosamine; GlcNS6S, *N*-sulfo-D-glucosamine-6-*O*-sulfate; Gal*N*Ac4S *N*-acetyl-D-galactosamine-4-*O*-sulfate; DS, dermatan sulfate; CS, chondroitin-4-sulfate; HS, heparan sulfate; CDD, Conserved Domain Database; BLOSUM, blocks of amino acid substitution matrix; ADT, AutoDock tools; VCAM-1, vascular cell adhesion molecule 1; NCAM, neural cell adhesion molecule; ICAM-1, intercellular adhesion molecule 1; LFA, lymphocyte function-associated antigen 1; CEA, carcino-embryogenic antigen; rmsd, root-mean-square deviation.

10.1021/bi702455e CCC: $40.75
Published on Web 04/05/2008

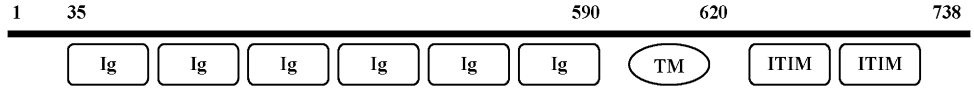

FIGURE 1: Structural organization of the domains of PECAM-1. The line indicates the relative positions of the amino acids with various structural features of PECAM-1 located in Table S1 in the Supporting Information. Ig, Ig-like domains; TM, transmembrane domains; ITIM, immunoreceptor tyrosine-based inhibitory motifs.

expressed in the kidney, lung, and trachea and, at lower levels, in the brain, heart, and liver. It is not expressed by fibroblasts, epithelial cells, or red blood cells (*2*). PECAM-1 is an important glycoprotein target in drug-induced immune thrombocytopenia (*3*) and has clinical importance in pathological disorders, such as cancer, thrombosis, and inflammatory and infectious diseases (*4–6*).

The 130 kD translated sequence of PECAM-1 contains six extracellular C2-type Ig-like domains, one transmembrane domain, and a 118 amino acid cytoplasmic domain (Figure 1) (*1, 4, 5*). PECAM-1 is differentially glycosylated, having both N- and O-linked glycosylation sites (*7*). Other post-translational modifications of PECAM-1 include palmitoylation of Cys 595 (*8*) and phosphorylation of the cytoplasmic tail (*9*). Its cytoplasmic tail becomes phosphorylated on serine and tyrosine residues following cellular activation, creating binding sites for SHP-2 (Src homology 2-containing tyrosine phosphatases) and perhaps other cytosolic signaling molecules (*9*). There are several PECAM-1 variants, which arise from alternative splicing of exons encoding either the transmembrane or cytoplasmic domain (*2*). These PECAM-1 variants are expressed in a cell-type and species-specific pattern in the human, rat, and mouse (*2*).

Five of the six C2-type Ig domains in PECAM-1 comprise a β-sandwich made of seven antiparallel β strands joined with a Greek key topology. Domain 5 is incomplete, having β strands that form only one side of the sandwich (*10*). The key feature of a C2 Ig fold is a core structure formed by two β sheets packed face to face. This structure is the key feature of cell adhesion molecules of the Ig superfamily, and it has been shown to be involved in a variety of cell−cell interactions (*11*). Six disulfide bridges are present in the Ig domains of human PECAM-1, along with nine consensus N-glycosylation sites; glycosylation accounts for about 20% of the apparent molecular size of PECAM-1. The major structural features of the Ig domains of human PECAM-1, as described in Swiss-Prot (P16284), are given in Table S1 in the Supporting Information.

The interactions of PECAM-1 with its ligands are complex. PECAM-1 engages in both homophilic (antiparallel interdigitation of opposing PECAM-1 molecules on adjacent cells) (*12*) and heterophilic (binding to other ligands) binding events, depending upon the conditions of the interaction and the ligands available (*1*). PECAM-1−PECAM-1 interactions can be both cis (interactions between adjacent PECAM-1 molecules in the same membrane) and trans (interactions between two PECAM-1 molecules in distinct membranes) associations (*13, 14*). PECAM-1 trans homophilic interactions require Ig domain 1 (*15*) and residues Asp 38, Asp 60, Lys 77, Asp 78, and Lys 116 (amino acid numbering from SWISS-PROT) have been implicated in this binding (*13, 16, 17*). Human−mouse chimeric studies have suggested that these homophilic interactions of PECAM-1 are species-specific (*17*). Moreover, antibodies that recognize the epitope "CAVNEG" in Ig domain 6 have been shown to enhance homophilic adhesion (*18*).

A number of heterophilic ligands have been proposed for PECAM-1. These include the integrin $\alpha_v\beta_3$, CD38, a 120 kDa ligand on T cells, the neutrophil antigen CD177, and heparan sulfate (HS) chains (*19–22*). These heterophilic interactions may be the result of the direct binding of PECAM-1 to another molecule or the result of secondary interactions mediated by non-PECAM-1 molecules, whose activation is PECAM-1-dependent. These molecules are believed to bind regions encompassing Ig domains 1–3 of PECAM-1 (*19–23*). PECAM-1 has been also reported to have high-affinity binding sites for Mn^{2+} cations, involving acidic residues in region 463–475 of Ig domain 5 and a cluster of acidic residues in regions 512–522 and 561–576 (amino acids numbered according to SWISS-PROT) in Ig domain 6 (*24*).

A number of studies have suggested that cell surface glycosaminoglycans (GAGs) act as ligands for PECAM-1 (*22*) by binding to a GAG consensus binding sequence (L-K-R-E-K-N) in Ig domain 2 (*22*). Interestingly, a similar GAG recognition sequence (residues 131–148) is located in loop 2 of Ig domain 2 of the neural cell adhesion molecule (NCAM), a heparin, and HS-binding protein (*22, 25*). The crystal structures of NCAM Ig domains forming zipper adhesion complexes (a dimerization pattern where at least two monomers are intertwined) reveal that heparin and HS-binding sites are solvent-exposed, suggesting that the association of NCAM with hetero- and homophilic ligands may occur simultaneously (*26, 27*). Experimental data have suggested that PECAM-1 could mediate L-cell aggregation by binding in a heterophilic fashion to HS on cell surfaces. Cell aggregation was blocked by iduronic acid-containing GAGs, including heparin, HS, and dermatan sulfate (weakly effective), but not by hyaluronic acid or the chondroitin sulfates; it was concluded that PECAM-1 bound iduronic acid containing GAGs (*22, 28*). However, others have argued that cell-surface GAGs are not ligands for PECAM-1 (*29*).

In the past decade, extensive X-ray crystallography, site-directed mutagenesis, and molecular modeling studies have provided an accurate structural characterization of several GAG fragments with biological activity, in either solution or their bioactive conformations bound to their target proteins (*30–34*). These studies have advanced the understanding of the molecular basis of GAG−protein interactions in the regulation of various physiological processes. Given the multiple activities of PECAM-1, an understanding of the nature of its interactions with all of its possible ligands is important because of the potential for PECAM-1 to become a drug target. In contrast to other heparin-binding proteins, the crystal structure of PECAM-1 has not yet been determined, either on its own or in a complex with heparin fragments. Here, we report the construction of a structural model of the extracellular domains of PECAM-1, the location of predicted sulfate-binding regions in PECAM-1 domains, and the modeling of various GAG fragments interacting with PECAM-1 domains.

MATERIALS AND METHODS

Homology Modeling of PECAM-1. Sequences of the human, mouse, rat, and porcine PECAM-1 and alternatively spliced PECAM-1 isoforms were retrieved from the SWISS-PROT protein sequence database. Multiple sequence alignments were performed with ClustalW (*35*) using BLOSUM matrices to quantify the sequence similarity between individual subunits of PECAM-1 in different species and isoforms. A PSI-BLAST (*36*) search against the Protein Data Bank (PDB) was performed to find sequences that were homologous with human PECAM-1 to identify proteins of known structure that can be used as a global template for homology modeling. Different sequence and secondary-structure prediction algorithms [PredictProtein (*37*) and PSIPRED (*38*)] were also used to predict the structure of residues comprising the subunits. The Ig-domain sequences, as classified by SWISS-PROT (see Table S1 in the Supporting Information), were submitted to the fold-recognition servers Phyre (*39*) (a successor algorithm of 3D-PSSM) and CBS Meta Server (*40*).

Alignments of the Ig domains of PECAM-1 with crystal structures of known Ig folds using LALIGN/PLALIGN (*41*) and PAM 120 matrices were performed. This expands earlier work using sequence-similarity searches only (*10, 42, 43*) and allowed us to include new structural information available from time to time in the subsequent stages of our modeling studies. The statistical significance of an alignment was computed by aligning the two sequences and then shuffling the second sequence between 200 and 1000 times using the PRSS module (*41*).

A preliminary sequence analysis with the sequences of PECAM-1 and its relatives was performed to investigate evolutionary relationships. The alignment between the PECAM-1 sequence and the template obtained from Phyre (*39*) was used to build the global alignment. The signal, transmembrane, and cytoplasmic regions were not modeled because of a lack of proper protein folds. The amino acid numbering from SWISS-PROT was used.

Structure construction, assignment of six disulfide bridges, optimization, and visualization were performed using the molecular-modeling package DS Modeling 1.7 (Accelrys, Inc.). Loops were built using the loop-modeling protocol in MODELLER (*44, 45*). Essential hydrogen and charges were added to the structure. The metal coordination site in Ig domain 6 was modeled using a manganese cation (Mn^{2+}), with a metal–oxygen distance of 2.15−2.25 Å, which is characteristic of metals bound to proteins (*46, 47*). Energy minimization of the modeled structure was carried out to remove unfavorable interactions. The CHARMm forcefield (*48*) was used with the smart minimizer method, which begins with the steepest descent method and is followed by the conjugate gradients method until the gradient reached a value below 0.001 kcal/mol. This was followed by molecular dynamics simulations (with a nonbonded cutoff of 10 Å and a dielectric constant of 4, at a temperature of 300 K for 20 ps using a time step of 1 fs), with the backbone atoms of the Ig domains kept fixed. The quality of the resultant protein structure was tested using PROCHECK (*49*), Eval23D (*40*), and Verify3D (*40*). Electrostatic potential calculations were performed using the DELPHI program in DS Modeling 1.7 (Accelrys, Inc.) using the atomic partial charges assigned by CHARMm, with a protein interior dielectric constant of 4, a solvent dielectric constant of 80, and an ionic strength of 0.145 M.

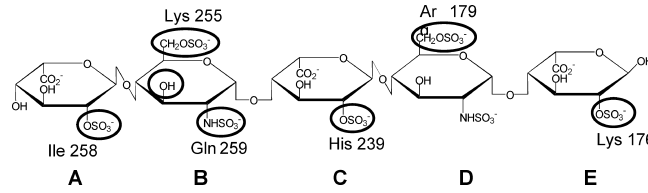

FIGURE 2: Structure of the pentasaccharide (4dIdoA2S−Glc*N*S6S−IdoA2S−Glc*N*S6S−IdoA2S) extracted from the hexasaccharide, for which there is a crystal structure (*53*). The unsaturated UA2S residue at the nonreducing terminus was modified to create a 4-deoxy-IdoA2S (4dIdoA2S) for the docking simulations. A−E refer to the labels of each of the residues of the pentamer. Circled anionic groups are critical for high-affinity interactions with Ig domains 2 and 3 of PECAM-1.

The PDB was surveyed for sulfate-binding motifs using BLAST searches. Briefly, the PECAM-1 Ig domains were split into a set of overlapping fragments 15 amino acids long; each overlap was 5 amino acids. For each fragment, a sequence-similarity BLAST search was performed from the PDB. Each hit was checked for a sulfate-binding motif in PDBsum.

Docking of GAG Fragments. PatchDock (*50, 51*) was used to dock heparin and other GAG fragments to the PECAM-1 model. PatchDock is a fast geometry-based molecular-docking algorithm that works by optimizing shape complementarity (hence, it is not an energy grid-based method). No constraints were used to define the binding site. This allowed the program to explore the entire surface of PECAM-1 and find appropriate interaction regions using root-mean-square deviation (rmsd) clustering to reduce the number of potential binding modes (*50*).

Most three-dimensional X-ray structures of GAG−protein complexes determined thus far involve relatively small oligosaccharides (from di- to hexasaccharides) of varying affinity for their protein targets. Hence, to determine the *minimal* length of the heparin fragments required for binding to the Ig domains of PECAM-1, docking simulations with di- and pentasaccharides were performed. The structure of the heparin pentasaccharide was obtained from the crystal structure of annexin A2 complexed with an unsaturated hexasaccharide in PDB code 2HYV (*52*). Because no electron density was observed for the sixth saccharide (*53*), the pentasaccharide was extracted directly from the structure. The residue at the nonreducing end was modified from the unsaturated UA2S by the addition of hydrogen to the double bond between C-4 and C-5 to create a 4-deoxy IdoA2S residue (4dIdoA2S) (Figure 2). Hence, the sequence of this modified pentasaccharide is 4dIdoA2S(1 → 4)Glc*N*S6S(1 → 4)IdoA2S(1 → 4)Glc*N*S6S(1 → 4)IdoA2S. The pyranose rings of the glucosamine residues adopt a $^4C_1$ chair conformation, whereas iduronic acids can adopt either a $^1C_4$ chair or a ^2S_o skew-boat conformation. The pentasaccharide has been modeled with the iduronic acid at the nonreducing end in a $^1H_2$ conformation and the central and terminal iduronic acids in the $^1C_4$ conformation. The structure of the disaccharide [IdoA2S(1 → 4)Glc*N*S6S] was extracted from the reported NMR structures of a heparin dodecasaccharide fragment (PDB code 1HPN) (*33*). We chose the iduronic acid in the $^1C_4$ chair conformation and the glucosamine in

Table 1: Templates Used for Building the Initial Models of the Ig Domains of PECAM-1 Identified by LALIGN/PLALIGN[a]

		amino acid range in PECAM-1											
		domain 1 (34–128)		domain 2 (139–234)		domain 3 (232–342)		domain 4 (328–414)		domain 5 (415–510)		domain 6 (499–604)	
template crystal structures	PDB code	percent identity	amino acid	percent identity	amino acid	percent identity	amino acid	percent identity	amino acid	percent identity	amino acid	percent identity	amino acid
VCAM (I) - C2 set	1VCA	19.8	81	18.7	75	41.2	17	25.9	81	19	76	21.6	74
VCAM(II) - I set	1VCA	43.8	16	22.2	72	30.4	23	21	81	25	45	33.3	18
CD8 (I) - V set	1HNF	25	40	27.8	18	21.7	83	27	37			27.3	22
CD8 (II) - C2 set	1HNF			50	6	31.6	19	27	63	26.9	52	23.1	78
CEA (I) - V set	1L6Z (mouse)	22.5	40	36.6	11	15.2	33	25.6	39			24.6	61
CEA (II) - I set	1L6Z	20.9	67	24.1	83	25	40	23.5	85	24.6	61		
CD16/FcGR (I) - I set	1FCG	23.7	76	50	6	33.3	36	27.7	65	20.7	58	50	12
CD16/FcGR (II) - I set	1FCG	26.5	68	22.7	44	30.2	63	26.2	84	24.6	65	24.7	81
NCAM (I) - I set	1QZ1 (rat)	28.1	32	33.3	12	20.4	49	20.4	49	20	40	20	30
NCAM (II) - I set	1QZ1	20.7	82	40	10	35	40	21.2	66	38.1	21	31.2	64
NCAM (III) - I set	1QZ1	38.1	21	27.3	55	22.7	22	23.8	63	22.4	67	21.2	99
ICAM (I) - C2 set	1P53	37.5	24	31.2	32	25.3	79	25	88	31.8	22	21.4	28
ICAM (II) - I set	1P53	38.5	13	35.3	34	29.2	24	19.3	83	42.9	14	29.4	17
ICAM (III) - I set	1P53	29.2	24	36.4	11	25.8	31	23.8	42	33.3	33	38.5	13

[a] The data in the table report the percentage identity and the length of the amino acid aligned range that shares identity with the query sequence (extracellular domains of PECAM-1) obtained by local sequence alignments.

the $^4C_1$ chair conformation to compare the docking results with those obtained for a similar conformation of the pentasaccharide.

Docking of the dermatan sulfate (DS) tetrasaccharide (PDB code 1HM2) (*52*) and a chondroitin-4-sulfate (CS) pentasaccharide (PDB code 1C4S) (*52*) was also performed. There is no crystal structure available for the DS pentasaccharide. The modeled DS tetrasaccharide consisted of IdoA(1 → 3)Gal*N*Ac4S(1 → 4)IdoA(1 → 3)Gal*N*Ac4S, and the modeled CS consisted of GlcA(1 → 3)Gal*N*Ac4S(1 → 4)GlcA(1 → 3)Gal*N*Ac4S(1 → 4)GlcA. Hydrogen atoms were added to these oligosaccharides, and the resultant structures were energy-minimized to optimize the orientation of rotatable groups. The surface area, atomic contact energy, and the binding score computed by PatchDock for the heparin pentasaccharide were extracted.

Further docking simulations were performed using AutoDock 3.0 (*54*). This program allows for flexibility in the ligand structure but uses a rigid-body approximation for the protein receptor. Auto Dock Tools were used to prepare the PECAM-1 molecule by adding appropriate hydrogens, partial atomic charges, and solvation parameters. Ligand rotatable bonds for all docked ligands were defined using the AutoTors module of AutoDock. The ligands were atom-typed manually to ensure that they complied with the carbohydrate forcefield in AMBER (*55*). The ligands were energy-minimized to optimize the orientation of their hydrogen atoms. A grid spacing of 0.37 Å and a distance-dependent dielectric constant of 4.0 [as defined by Mehler and Solmajer (*56*)] were used for the binding energy calculations, covering the putative binding site surface. Using AutoDock's Lamarckian genetic algorithm, heparin fragments were subjected to 200 search runs using a population of 200 individuals. The grid box was defined with a constant grid spacing of 0.37 Å around each heparin fragment using the binding poses obtained from PatchDock with respect to Ig domains 2 and 3, Ig domain 4, and Ig domains 5 and 6 of PECAM-1.

Because of the flexibility and size of the di- and pentasaccharides, the number of energy evaluations and the size of the genetic population were optimized to ensure convergence of the calculated energies, starting with a minimum of 5×10^6 and a maximum of 50×10^6 energy evaluations. Cluster analysis was performed on the resulting binding poses using a rmsd tolerance of 1.0 Å. Because AutoDock cannot handle more than 32 rotatable bonds, the docking of the heparin fragments was performed, keeping the hydroxyl groups fixed. The lowest docking energy binding scores of the disaccharides with full rotational freedom of their hydroxyl groups were similar to those obtained when the hydroxyl groups were fixed, confirming that the initial orientation of the hydroxyl groups was appropriate for interactions with PECAM-1 (these interactions were more important for nonheparin fragments, as discussed below).

RESULTS AND DISCUSSION

Comparative Modeling. A standard homology-building procedure was adopted to construct a three- dimensional model of the extracellular domains of PECAM-1, starting with similarity searches and followed by secondary-structure predictions. The secondary structure as predicted by PSIPRED can be found in Figure S1 in the Supporting Information, and the percentage of secondary-structure type and solvent accessibility can be found in Table S2 in the Supporting Information. BLAST searches detected conserved Ig regions in domains 1, 4, and 6. Proteins with the required sequence similarity (>35%) were not found following a global search using the entire extracellular region of PECAM-1 or when individual PECAM-1 Ig domains were considered. In local sequence-similarity searches, the identities of the aligned sequences varied from 18 to 23%. The percentage identity of various alignments was very low, as shown in Table 1. Consequently, threading and comparative modeling techniques were used to model the extracellular domains of PECAM-1. The alignment of the PECAM-1 Ig-like domains

with other members of the Ig-CAM family performed with LALIGN/PLALIGN suggests that PECAM-1 evolved through gene duplication (*10*). Threading servers predicted various I set topologies (*57*) for Ig folds with various templates, including VCAM-1, NCAM, ICAM-1, CEA, and different isoforms of FcγR. The crystal structures of VCAM-1, NCAM, and ICAM-1 confirmed the I-type topology in these molecules (*26*). We modeled extracellular domains 1 and 2 with multiple templates, including NCAM and VCAM, as reported by others (*13, 15*).

The evolutionary analysis using ClustalW suggested that NCAM, CEA, and FcγR are more closely related to human PECAM-1; hence, we chose to derive the models with these templates. The Ig domain 1 of PECAM-1 was modeled with NCAM (Ig domain 2) and VCAM (Ig domain 1) as templates. Ig domain 2 of PECAM-1 was modeled using the structural features of NCAM (Ig domain 3) and VCAM (Ig domains 1 and 2). The Ig I set topology in domain 3 of PECAM-1 was modeled with multiple templates, including CD8 and ICAM. Extracellular domains 4–6 of PECAM-1 showed a preference for various folds of FcγR, as predicted by fold-recognition programs, such as Phyre (*39*).

The quality of the final model of the extracellular domains of human PECAM-1 was evaluated using the PROCHECK program (*49*) at a 2.5 Å resolution, which gives Ramachandran plots and a quantitative distribution of the geometric parameters within the allowed conformational space. The percentage of residues in the most favored, allowed, and disallowed conformations were 82.7, 13.7, and 1.0, respectively. The distribution of the ψ/φ angles of the model is within the allowed region, and only five residues are in disallowed regions, as shown in the Ramachandran plot in Figure S2 in the Supporting Information. These residues were not further optimized because of the extensive presence of loops and glycine and proline residues. Eval23D (*40*) and Verify3D (*40*) also predicted a good model, with 3D profile scores of 0.035 and 0.117, respectively. The models of the various Ig domains of human PECAM-1 that were derived are shown in Figure 3. The electrostatic potential surface representation shows that Ig domains 2−6 have positive electrostatic surfaces (colored blue in Figure 3), which may constitute binding surfaces for anionic ligands, such as GAGs.

Various approaches have been used to identify heparin/GAG-binding sites on the surface of proteins. These have been based on amino acid composition (*58, 59*), secondary structure (*60*), spatial distribution of the basic amino acids (*61*), and the surface properties of proteins (*62*). While consensus sequences, such as XBBXBX and XBBBXXBX (where B is a basic residue and X can be any residue), have been suggested for heparin binding (*63*), they are neither necessary nor sufficient to define a GAG-binding site. GAG-binding sites generally consist of a cluster of basic residues on the protein surface but not necessarily in a continuous sequence. Consequently, a database search for sulfate-binding structural motifs was performed. It has been observed that arginine and lysine have the highest propensity to bind GAGs. Interestingly, basic residues, such as arginine and lysine, comprise approximately 13% of the total amino acids present in the various domains of PECAM-1. The percentage composition of amino acids was predicted using the PHD module of PredictProtein (*37*) (Table S3 in the Supporting Information).

A survey of sulfate-binding regions revealed several positively charged regions in Ig domains 2−4 and 6. Residues 177–182 in Ig domain 2 have a high sequence identity with the sulfate-binding site in the bacterial protein disulfide oxidoreductase (PDB code 2AYT). Other sulfate-binding motifs were also found: (1) residues 207–223 in Ig domain 2, which are homologous to sulfate-binding motifs in bacterial protein SecA translocation ATPase (PDB codes 1M6N and 1M74); (2) residues 254–258 and 278–286 in Ig domain 3, which are homologous to sulfate-binding motifs in snake phospholipase A2 (MIPLA3) (PDB code 1OZY); and (3) residues 330–342 in Ig domain 4, which are homologous to sulfate-binding motifs in the multiple sugar-binding transport ATP-binding protein (PDB code 2D62) and ribosomal protein S6 kinase α 5 (PDB code 1VZO). Interestingly, the region 563–571 in Ig domain 6 of PECAM-1 showed homology to the sulfate-binding motifs in the crystal structure of HIV-2 reverse transcriptase (PDB code 1MU2). In our model of human PECAM-1, the relative distance between the sulfate groups when modeled binding to their binding motifs in Ig domain 2 was approximately 8–9 Å; similarly, the sulfates modeled as binding to their motifs in Ig domain 3 are approximately 8–9 Å apart (Figure 3c). This spacing between sulfate-binding motifs is also observed in the crystal structure of the heparin-binding protein, artemin (PDB code 2ASK) (*64*). Two sulfate-binding motifs were identified in Ig domain 4 (Figure 3a), placed at a distance of 24 Å from each other. No sulfate-binding motif was identified in Ig domain 5. A single sulfate-binding motif was predicted by the survey in Ig domain 6. The clusters of basic amino acids in Ig domains 2 and 3 are located approximately 20 Å apart. A comparison of the spatial distribution of basic amino acids in other heparin-binding proteins that have a β-sheet topology (such as apolipoprotein E, antithrombin III, and NCAM) suggests that a 20 Å long region can accommodate a GAG pentasaccharide (*61*).

Ig domain 5 of PECAM-1 is connected to Ig domain 4 through a long flexible loop. Ig domain 5 forms only half of the Ig fold because it lacks the characteristic β strands a, b, d, and e, as reported earlier (*10*). We obtained the b strand from the structure of the KK50.4 T-cell receptor β chain (PBD code 2ESV: E). The region 449–507, including the d and e strands in Ig domain 5, showed 49% homology with monomeric isocitrate dehydrogenase complexed with isocitrate and Mn^{2+} (PDB code 1ITW) when a PSI-BLAST search was performed. However, after 3D fragments of domain 5 were surveyed in the PBD, no sulfate-binding motifs were detected.

Construction of Ig-like domains 1 and 2 of PECAM-1 was carried out by modeling the side chains of specific surface-exposed amino acids to facilitate interactions with known homo- and heterophilic ligands. The side chains of the key residues Asp 38, Asp 60, Lys 77, Asp 78, and Lys 116 (SWISS-PROT numbering) that mediate PECAM-1 homophilic binding were repositioned in Ig domain 1 (see Figure 3b) on the basis of PECAM-1 models described earlier based on the VCAM-1 structure (*13, 15*). A ClustalW (*35*) sequence analysis of PECAM-1 shows that residues Asp 60, Lys 77, Asp 78, and Lys 116 are highly conserved in the human, bovine, pig, rat, and mouse, whereas Asp 38 in human PECAM-1 is replaced by His 38 in other species.

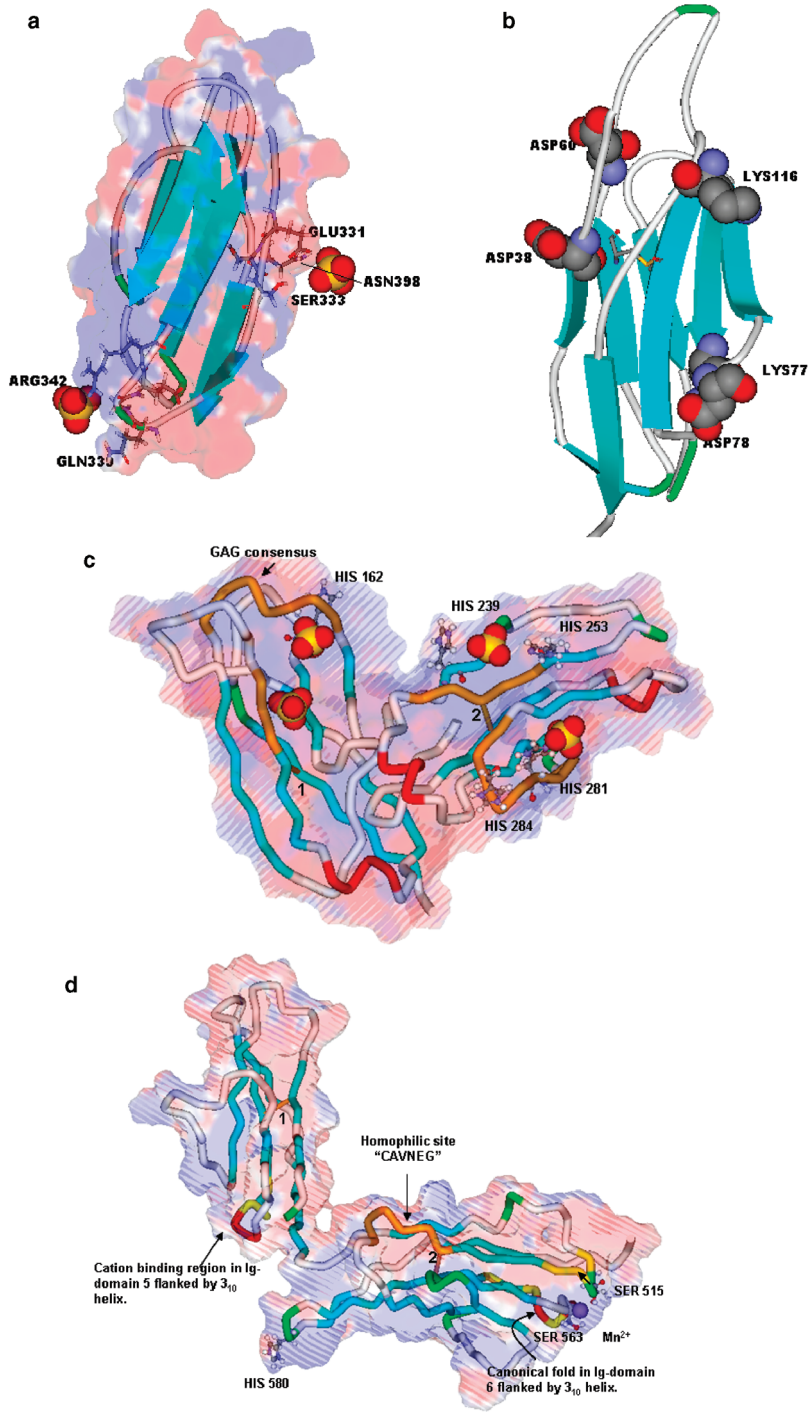

FIGURE 3: Modeling of the six extracellular domains of PECAM-1 and their predicted sulfate-binding regions. Three-dimensional models of the Ig domains of human PECAM-1 are shown in a schematic representation according to secondary structure. The surface is represented by electrostatic potential (negative potential in red and positive potential in blue) for extracellular domains 2−6. The sulfates are shown in CPK. (a) Ribbon diagram of Ig domain 4 with the sulfate-binding motifs indicated by highlighted residues (in a stick representation). (b) Homophilic binding sites reported by Newton et al. (*13*) in Ig domain 1 are shown in CPK representation. The disulfide bond between Cys 57 and Cys 109 is shown in a stick representation. (c) Ig domains 2 and 3 are shown in a toothpaste representation according to their secondary structure. Numbers 1 and 2 indicate the disulfide bridges, as mentioned in Table S1 in the Supporting Information. The protonated histidines are shown in a ball and stick representation, and the positive electrostatic potential (shown in blue) surface represents the regions consisting of basic residues found by a survey of sulfate-binding motifs (in orange), which may constitute high-affinity binding sites for GAGs. (d) The cation coordination site in Ig domains 5 and 6, as described by Jackson et al. (*24*), is represented by a yellow tube. Numbers 1 and 2 indicate the disulfide bridges, as mentioned in Table 1. The cation binding regions in Ig domains 5 and 6 are flanked by the 3_{10} helix. The homophilic binding site identified in Ig domain 6 (*18*) is shown in orange. The regions of positive electrostatic potential (shown in blue) may contribute to low-affinity binding sites for GAGs.

This is consistent with homophilic interactions of PECAM-1 mediated by Ig domain 1 being species-specific (*15*).

Ig domains 1–3 of PECAM-1 were said to participate in $\alpha_v\beta_3$-mediated heterotypic binding. This binding was inferred to be cation- and temperature-dependent (*65*); however, the validity of $\alpha_v\beta_3$ as the PECAM-1 ligand has been challenged (*16*). Our survey of structural motifs from the PBD indicated that PECAM-1 Ig domains 1–3 lacked the integrin-binding

motifs found in the members of the Ig superfamily (IgSF) that have been described as integrin binding; these include VCAM, ICAM-1, and ICAM-2. This suggests that the cell adhesion attributed to direct PECAM-1−$\alpha_v\beta_3$ interactions may occur by an indirect mechanism, which involves PECAM-1 and $\alpha_v\beta_3$, as reported previously (*16*).

PECAM-1 has a large presence of charged residues on the surface of Ig domains 1−3. Modeling of the protonated and unprotonated states of these residues provided useful information. Basic amino acids are clustered in our model of PECAM-1, with histidine side chains from the β sheets positioned within 3.0 to 5.0 Å of the sulfate-binding motifs. Ig domains 2 and 3 are brought into close contact by a loop of only three residues long. However, the flexibility of this loop results in either an open or a closed conformation of that portion of the PECAM-1 molecule. In our homology model, the best-fit loop modeled connects Ig domains 2 and 3 in a closed conformation. In a closed conformation, it is expected that a GAG fragment may be able to bind to PECAM-1 in a way that involves domains 2 and 3. Modeling of these domains (Figure 3c) showed the sulfates bound to two clusters of residues: Q-A-R (207–209) and L-K-R-E-K-N (177–182), with His 162 in close vicinity in Ig domain 2. The side chains of His 239 and His 253 are oriented toward the region I-K-C-T-I (254–258), and His 281 and His 298 are in close vicinity of the region 278–286 in Ig domain 3 (Figure 3c). Histidines were modeled in a positively charged form with both N_δ and N_ε atoms protonated. Modeling of surface-exposed histidines in Ig domains 2 and 3 in close vicinity to the sulfate-binding motifs suggests that these histidines may assist in the binding of PECAM-1 to heparin or HS in a pH-dependent manner, as was observed for the chemokine CXCL 12 (*66*), the cytokine VEGF (*47*), and the mouse mast cell protease 7 (*67*). Modeling of a fully open conformation was seen to result in the clusters of basic residues in Ig domains 2 and 3 being too far apart to constitute a GAG-binding site because of a twist in the relative orientation of these domains.

We also modeled cation coordination sites in Ig domains 5 and 6 of PECAM-1. A survey of the PDB for metal-binding sites in this region did not detect any metal-binding motifs. The acidic residues in the region 463–475 are placed at a distance of 8–9 Å, too far apart to constitute a metal coordination site (Figure 3d). However, the presence of a cluster of basic residues in the vicinity (second coordination shell) of a cation-binding site could act as an electrostatic anchor for a metal ion. In region 565–572 of Ig domain 6, we found a canonical fold consisting of a β sheet followed by a 3_{10} helix (Figure 2d). Such folds have also been described for the anti-HIV-1 V3 Fab 2219 structure (PDB codes 2B0S: L and 2B1A: L) and the CD8 α ectodomain fragment (PDB code 1BQH: G, H, I, and K).

A previously described cation-binding site in Ig domain 6 was shown to have higher affinity for Mn^{2+} than for Ca^{2+} or Zn^{2+} (*24*). Consequently, the conformation of the Ig-domain 6 cation-coordination metal-ion-dependent adhesion site (MIDAS) was modeled in the presence of Mn^{2+}, as previously described (*68, 69*). In those papers, the cell adhesion molecules Mac-1(CD11b/CD18/αMβ2), LFA-1(αLβ2), and CD2 had the metal ion coordinated to water molecules as well as serine, aspartate, and glutamate residues. In our model, Glu 514, Asp 518, and Glu 569 are located 6 Å from each other, and therefore, it is possible that these residues may similarly coordinate to Mn^{2+} through water molecules in the MIDAS site (*68*). Our model of Ig domain 6 differs in terms of the position of the metal coordination site flanked by the 3_{10} helix from that of the model described by Jackson and co-workers (*24*). To validate our modeling of the topology of the cation-binding site, searches were performed in the Conserved Domain Dataset for Ig domains 5 and 6. These searches suggested structures with similar folds to those that we derived in our model. The existence of divalent cation-binding sites proximal to the protein surface might allow Mn^{2+} to interact efficiently with histidine-containing ligands (*70*). Interestingly, the neutrophil-specific antigen CD177 was recently predicted to mediate a heterophilic interaction with PECAM-1 through the cation-coordination site present in Ig domain 6 (*71*).

Docking of GAGs. A heparin pentasaccharide (Figure 2) was docked to all Ig domains of PECAM-1 on the basis of shape complementarity using PatchDock in an effort to obtain initial binding modes of the pentasaccharide with each domain. The best binding mode of the pentasaccharide was to Ig domains 2 and 3 and was determined to have an approximate interaction surface area of 1200 Å^2, an atomic contact energy (ACE) (*72*) of 290, and a geometric shape complementarity score of 9830. The second best binding mode of the pentasaccharide was obtained with Ig domains 5 and 6, having an approximate interaction surface area of 865 Å^2, an ACE of 152, and a geometric shape complementarity score of 7598. While there is no evidence of the accuracy of these measures for carbohydrate−protein interactions, the scores suggest the presence of a high- and low-affinity GAG-binding site in Ig domains 2 and 3 and Ig domains 5 and 6, respectively.

To obtain better estimates of the free energies of binding, the AutoDock program was used. The top ranking binding mode of heparin that was obtained resulted in an improved fit between the negatively charged pentasaccharide and the positively charged regions in Ig domains 2 and 3 (Figure 4a). Docking of disaccharides to Ig domains 5 and 6 of PECAM-1 using AutoDock suggested a better fit and lower free energies of binding on the electropositive surface of these domains than that obtained with the pentasaccharide and PATCHDOCK.

The results of the docking studies indicating that a heparin pentasaccharide binds a region on Ig domains 2 and 3 are consistent with the predicted location of the sulfate-binding motifs. The key interactions of the heparin pentasaccharide and Ig domains 2 and 3 identified by the docking simulations involve Lys 176, Leu 177, Arg 179, His 239, Lys 255, Gln 259, and Ile 258 (Figure 4b). The protonated $N_{\varepsilon 2}$ in His 239 makes an electrostatic interaction with the 2-*O*-sulfate of the central iduronic acid of the pentasaccharide (residue C Figure 2). Residues Ile 258 (main chain) and Lys 255 (side chain) make hydrogen bonding and electrostatic contacts with the 2-*O*-sulfate in residue A and the 6-*O*-sulfate in residue–r-esidue B, respectively (Figure 2). The side chain of Gln 259 makes a hydrogen bond with the *N*-sulfate in residue B in the top-ranked binding mode obtained but makes a hydrogen bond with the hydroxyl group of residue B in the second-ranked binding mode. The *O*-sulfates in residues E and D establish ionic interactions with the charged side chains of Lys 176 and Arg 179, respectively. Furthermore, Arg 209

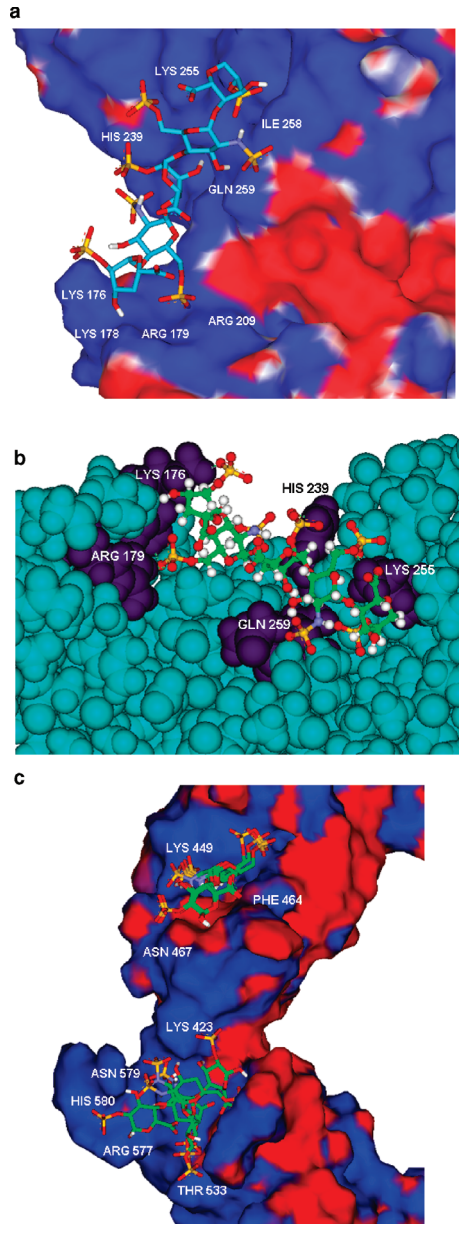

FIGURE 4: Predicted binding modes for sulfated pentasaccharide with various domains of human PECAM-1, which is represented with an electrostatic potential surface (negative potential in red and positive potential in blue). The electrostatic potential surfaces were calculated and displayed using the DELPHI module in Discovery Studio (Accelrys, Inc.). (a) Electrostatic potential surface representation of Ig domains 2 and 3 of PECAM-1. The pentasaccharide fragment is shown as sticks. (b) Predicted binding mode for a sulfated pentasaccharide in Ig domains 2 and 3 of PECAM-1, showing those amino acids (in purple) that interact with the fragment. (c) Predicted binding modes for sulfated disaccharides with Ig domains 5 and 6. The protein surface is colored according to the sign of the electrostatic potential (blue for positive areas and red for negative areas). Low binding energy clusters are depicted for the disaccharides, showing that amino acids form basic (positively charged) clusters on the surface of the protein.

in Ig domain 2 is in close proximity to the GAG consensus region ("L-K-R-E-K-N"), and hence, it is possible that its guanidine group may interact with the charged residues of the pentasaccharide. However, this was not observed in the docking simulations, because it would require a change in the conformation of the main chain of residues 207–209 to bring Arg 209 closer to this cluster of basic residues.

The docking calculations predicted the free energy of binding and the dissociation constant (K_i) of the ligands with the extracellular domains of PECAM-1 at slightly acidic pH. The predicted free energy of binding and dissociation constant of the best binding mode of the heparin pentasaccharide with Ig domains 2 and 3 were computed to be –17.22 kcal/mol and 4.93 nM, respectively. The second-ranked binding mode was predicted to have a free energy of binding and a dissociation constant of –16.91 kcal/mol and 9.04 nM, respectively. When these calculations were repeated at a neutral pH (leaving the histidine residues unprotonated), the free energy of binding increased to approximately –3 kcal/mol, because of the loss of ionic interactions with His 239.

These calculations were made assuming a closed configuration of Ig domains 2 and 3 and the $^1C_4$ chair conformation of the IdoA2S. It is likely that in nature Ig domains 2 and 3 will not always be in such close proximity, hence allowing longer fragments to interact with both domains. Moreover, although IdoA residues exist in two conformations of nearly equal energy ($^1C_4$ chair and ^2S_o skew-boat conformations), internal IdoA residues will favor the ^2S_o skew-boat conformation because in the $^1C_4$ form the bulky carboxylate group is equatorial and all other substituents are in axial positions (*73*). Docking of the pentasaccharide with these alternative IdoA conformations would likely result in a lower affinity of binding, and hence, experimental determinations would reflect an average lower binding affinity.

The docking simulations also predicted the binding of heparin fragments to Ig domains 5 and 6. In this case, docking of sulfated disaccharides resulted in two clusters with significant numbers of related binding modes (Figure 4c). The disaccharides with the lowest energies of binding in each cluster were seen to interact with the positively charged accessible surfaces of Ig domains 5 and 6. However, the interactions are predicted to be weak: the computed free energies of binding in clusters 1 (Ig domain 5) and 2 (Ig domain 6) were –6.54 and −6.23 kcal/mol, respectively, resulting in dissociation constants of 15.4 and 26.9 μM, respectively. The third lowest energy cluster was predicted to have a free energy of binding and dissociation constant of –6.13 kcal/mol and 32.2 μM, respectively. A number of amino acids (Lys 423, Lys 446, Lys 449, Asn 467, Arg 577, and His 580) in Ig domains 5 and 6 interact with the disaccharides, mostly through ionic interactions with their sulfates. In addition, the main chain of Phe 464 interacts with the ionized carboxylate, and Thr 533 (main chain) interacts with the 2-*O*-sulfate of the IdoA residue of the disaccharide. The side chain of Glu 470 makes a hydrogen bond with the amine group of glucosamine, whereas the main chains of Gly 528 and Ser 529 make hydrogen bonds with the hydroxyl groups of the disaccharides.

The binding of the disaccharides to Ig domains 5 and 6 did not involve the cation-binding region. The predicted sulfate-binding motif (region 563–571) in Ig domain 6 partially overlaps with the cation-binding region (formed by residues 512–522 and 560–572). Although the surface of this region in Ig domain 6 is electropositive in nature, docking simulations of disaccharides to this region (Figure 2d) resulted in predicted positive free energies of binding; therefore, no binding would be expected. Despite the sulfate-binding region 563–571 in Ig domain 6 showing a high level of identity and a similar structural topology with the template

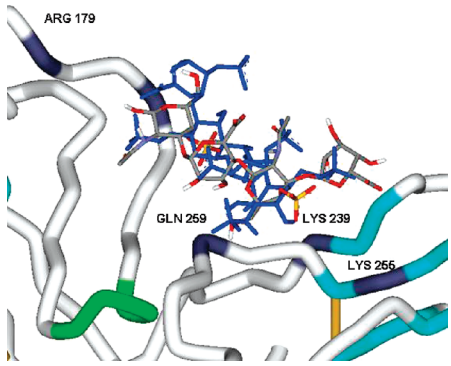

FIGURE 5: Predicted binding mode for a dermatan sulfate tetrasaccharide (shown in sticks and colored by atom type), superimposed on a heparin tetrasaccharide (shown in blue) in Ig domains 2 and 3 of PECAM-1. The amino acids that interact with the fragment are shown in purple.

structure 1MU2 (HIV-2 reverse transcriptase), the presence of Glu 569 in Ig domain 6 of PECAM-1, instead of histidine (as in HIV-2 reverse transcriptase), eliminates the possibility of favorable ionic interactions forming with a sulfate group.

In the case of Ig domain 4, docking of sulfated disaccharides resulted in extremely low binding affinity (dissociation constants in the molar range). A cluster of binding poses for the disaccharides were found to interact with Ser 333, Arg 342, and Asn 344, as predicted by the survey of sulfate-binding motifs; however, no significant affinity was measured. It is likely that the lack of conservation of the protein fold required to coordinate the sulfate on the surface of Ig domain 4, as compared to the templates 2D62 (multiple sugar-binding transport ATP-binding protein) and 1VZO (ribosomal protein S6 kinase α 5), result in a lack of binding affinity of GAGs to this region. The sulfates in the template structures are also coordinated by additional residues from neighboring structural folds. The fold in Ig domain 4 of PECAM-1 lacks this coordination from the neighboring β sheet.

The distribution of amino acids in the predicted sulfate-binding regions of the Ig domains of PECAM-1 was examined in other species. Multiple sequence alignment of the human sequence with *Bos taurus* (bovine), *Mus musculus* (mouse), *Sus scrofa* (pig), and *Rattus norvegicus* (rat) suggests that the Ig domains in other species may bind GAGs with different affinity because of differences in the sequence conservation of sulfate-binding motifs in regions 177–182 and 207–209 in Ig domain 2 and regions 239, 254–258, and 278–286 in Ig domain 3 (Figure S3 in the Supporting Information). The charged residue Arg 179, involved in the interaction of Ig domain 2 of human PECAM-1 with the GAG pentasaccharide, is replaced by Ile 179 in mouse. Arg 209 in the sulfate-binding motif 207–209 is mutated to Gly 209 in the mouse and Asn 209 in the pig. His 239 in human PECAM-1, predicted to interact strongly with the GAG pentasaccharide at slightly acidic pH, is replaced by the more acidic residues Glu 239 in the mouse and Gln 239 in the rat. His 281, found in the predicted sulfate-binding motif of Ig domain 3 of human PECAM-1, is replaced by Thr 281 in the mouse and rat.

A DS tetrasaccharide was docked to Ig domains 2 and 3. The predicted binding mode showed that the hydroxyl groups present in the DS tetrasaccharide contribute to hydrogen bonding with Arg 179 and Gln 259, whereas these residues

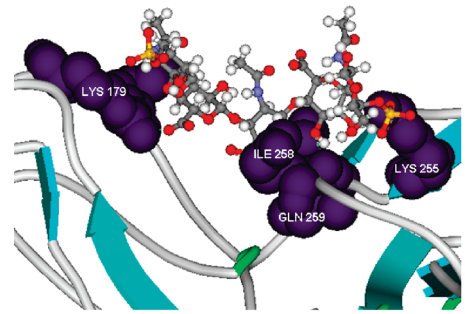

FIGURE 6: Predicted binding modes for chondroitin sulfate pentasaccharides with Ig domains 2 and 3 of PECAM-1. The amino acids that interact with the fragment are shown in purple.

make ionic interactions with the sulfate group of heparin/HS oligosaccharide. The sulfate group of DS makes an ionic interaction with Lys 255, resulting in a free energy of binding and dissociation constant of –6.46 kcal/mol and 18 μM, respectively. The binding modes of DS and heparin tetrasaccharides are shown in Figure 5. It is important to note that residue A at the nonreducing end of the heparin tetrasaccharide is in the $^{1}H_{2}$ form, residues B and D are in the $^{4}C_{1}$ conformation, and residue C of IdoA2S adopts the $^{1}C_{4}$ chair formation, whereas the iduronic acids of DS are in the $^{1}C_{4}$ conformation, resulting in different protein−oligosaccharide interactions. This suggests that the biological activities of heparin and DS are modulated not only by electrostatic interactions but also by the flexibility of the iduronic acid subunits and the presence of optimal van der Waals interactions between the oligosaccharides and the protein.

Docking of a CS pentasaccharide to Ig domains 2 and 3 resulted in a predicted positive free energy of binding because of the presence of a lower degree of sulfation and lower flexibility of saccharide subunits compared to heparin/HS and DS tetrasaccharides, giving rise to fewer interactions with the electropositive regions found on the surface of Ig domains 2 and 3 of PECAM-1 (Figure 6) The charged sulfate and carboxylate groups make ionic interactions with Lys 255 and Lys 176, respectively, and hydrogen bonding is observed between the hydroxyl of CS and the backbone of Ile 258.

These results support the experimental evidence that PECAM-1 cannot bind CS but may bind DS with low affinity (*28*). Our simulations suggest that GAG chains, such as heparin and DS, which contain iduronic acids, can more easily adopt a configuration appropriate for binding to the Ig domains of PECAM-1 because of their greater flexibility, compared to glucuronic acid-rich GAGs, such as CS. Our results also indicate that Ig domains 2 and 3 of PECAM-1 are critical for heparin recognition and binding and are in agreement with a study demonstrating that GAG consensus "LKREKN" (177–182) in the Ig domain 2 of PECAM-1 is involved in GAG binding (*22*).

We have predicted that heparin fragments may have direct electrostatic interactions with positively charged residues located in loops in Ig domains 2 and 3, providing the basis for the existence of a high-affinity GAG-binding region in PECAM-1 at slightly acidic pH. Our docking simulations indicate that the GAG-binding region involves Ig domain 3, residues His 239, Lys 255, and Gln 259, with further contributions from Ig domain 2, residue Arg 179. An additional low-affinity heparin-binding region appears to be located in Ig domain 5, with contributions from Ig domain

6. Importantly, these two putative GAG-binding regions are distinct from regions involved in homo- and heterophilic interactions in Ig domains 1 and 6, as well as the cation-binding sites in Ig domains 5 and 6.

Ligand−protein docking predictions of free energies of binding and dissociation constants do not have the same accuracy as experimental techniques and should be taken as indicative. However, the main reasons behind the differences between the predicted binding affinities and those determined experimentally have to do with differences between what is modeled and what is assayed experimentally. The docking simulations use heparin fragments with IdoA residues in defined ring conformations, while an experimental determination would be an average measure of interactions involving different conformations. In addition, heparin fragments used in experiments may have different patterns of sulfation than the fragment modeled. Not all heparin fragments produced experimentally will contain a pentasaccharide sequence of the repeating trisulfated disaccharide structure that was modeled, but many will because about 60% of heparin consists of this disaccharide structure. Generally, in experimental situations, a mixture of fragments is examined for their ability to bind to proteins and it is unlikely that a heterogeneous population of heparin fragments will give the same binding affinities as a defined structure. Nevertheless, it is expected that heparin fragments longer than those modeled in this study, providing they contain the correct structural motif, will bind to the high-affinity GAG-binding region in Ig domains 2 and 3 of PECAM-1, as is reported in the following paper (*74*).

PECAM-1 has an important role in the diapedesis step of leukocyte extravasation in inflammation (*1*). Our work suggests that PECAM-1 may be capable of mediating cell aggregation through interactions with specific GAGs on adjacent cells, as in the case of NCAM. Thus, it is possible that the interactions of PECAM-1 with HS on the cell surface contribute to the normal function of PECAM-1 during both inflammation and angiogenesis.

ACKNOWLEDGMENT

We gratefully acknowledge the Western Australian Interactive Virtual Environments Centre (IVEC) for access to high-performance computing facilities.

SUPPORTING INFORMATION AVAILABLE

Structural features of human PECAM-1 as described in SWISS-PROT (P16284) (Table S1), predicted secondary-structure and solvent-accessibility compositions (core/surface ratio) for human PECAM-1 using PredictProtein (Table S2), percentage residue composition for human PECAM-1 sequence (Table S3), secondary-structure predictions for the aligned protein sequences of the subunits of human PECAM-1 using PSIPRED (Figure S1), Ramachandran plot of the derived model of PECAM-1 (Figure S2), and multiple sequence alignment of PECAM-1 sequences from various species, including *Homo sapiens*, *Bos taurus*, *Mus musculus*, *Sus scrofa*, and *Rattus norvegicus* (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Newman, P. J. (1997) The biology of PECAM-1. *J. Clin. Invest. 100*, S25–S29.
2. Wang, Y., Su, X., Sorenson, C. M., and Sheibani, N. (2003) Tissue-specific distributions of alternatively spliced human PECAM-1 isoforms. *Am. J. Physiol. Heart Circ. Physiol. 284*, H1008–H1017.
3. Kroll, H., Sun, Q. H., and Santoso, S. (2000) Platelet endothelial cell adhesion molecule-1 (PECAM-1) is a target glycoprotein in drug-induced thrombocytopenia. *Blood 96*, 1409–1414.
4. Newman, P. J. (1994) The role of PECAM-1 in vascular cell biology. *Ann. N.Y. Acad. Sci. 714*, 165–174.
5. Newman, P. J. (1999) Switched at birth: A new family for PECAM-1. *J. Clin. Invest. 103*, 5–9.
6. Moseley, G., and Jackson, D. (2004) The multiple functions of PECAM-1. *Aust. Biochemist 35*, 9–12.
7. Newton, J. P., Hunter, A. P., Simmons, D. L., Buckley, C. D., and Harvey, D. J. (1999) CD31 (PECAM-1) exists as a dimer and is heavily N-glycosylated. *Biochem. Biophys. Res. Commun. 261*, 283–291.
8. Sardjono, C. T., Harbour, S. N., Yip, J. C., Paddock, C., Tridandapani, S., Newman, P. J., and Jackson, D. E. (2006) Palmitoylation at Cys595 is essential for PECAM-1 localisation into membrane microdomains and for efficient PECAM-1-mediated cytoprotection. *Thromb. Haemostasis 96*, 756–766.
9. Jackson, D. E. (2003) The unfolding tale of PECAM-1. *FEBS Lett. 540*, 7–14.
10. Newman, P. J., Berndt, M. C., Gorski, J., White, G. C., II, Lyman, S., Paddock, C., and Muller, W. A. (1990) PECAM-1 (CD31) cloning and relation to adhesion molecules of the immunoglobulin gene superfamily. *Science 247*, 1219–1222.
11. Chothia, C., and Jones, E. Y. (1997) The molecular structure of cell adhesion molecules. *Annu. Rev. Biochem. 66*, 823–862.
12. Holness, C. L., and Simmons, D. L. (1994) Structural motifs for recognition and adhesion in members of the immunoglobulin superfamily. *J. Cell Sci. 107* (part 8), 2065–2070.
13. Newton, J. P., Buckley, C. D., Jones, E. Y., and Simmons, D. L. (1997) Residues on both faces of the first immunoglobulin fold contribute to homophilic binding sites of PECAM-1/CD31. *J. Biol. Chem. 272*, 20555–20563.
14. Zhao, T., and Newman, P. J. (2001) Integrin activation by regulated dimerization and oligomerization of platelet endothelial cell adhesion molecule (PECAM)-1 from within the cell. *J. Cell Biol. 152*, 65–73.
15. Nakada, M. T., Amin, K., Christofidou-Solomidou, M., O'Brien, C. D., Sun, J., Gurubhagavatula, I., Heavner, G. A., Taylor, A. H., Paddock, C., Sun, Q. H., Zehnder, J. L., Newman, P. J., Albelda, S. M., and DeLisser, H. M. (2000) Antibodies against the first Ig-like domain of human platelet endothelial cell adhesion molecule-1 (PECAM-1) that inhibit PECAM-1-dependent homophilic adhesion block in vivo neutrophil recruitment. *J. Immunol. 164*, 452–462.
16. Sun, Q. H., DeLisser, H. M., Zukowski, M. M., Paddock, C., Albelda, S. M., and Newman, P. J. (1996) Individually distinct Ig homology domains in PECAM-1 regulate homophilic binding and modulate receptor affinity. *J. Biol. Chem. 271*, 11090–11098.
17. Sun, J., Williams, J., Yan, H. C., Amin, K. M., Albelda, S. M., and DeLisser, H. M. (1996) Platelet endothelial cell adhesion molecule-1 (PECAM-1) homophilic adhesion is mediated by immunoglobulin-like domains 1 and 2 and depends on the cytoplasmic domain and the level of surface expression. *J. Biol. Chem. 271*, 18561–18570.
18. Yan, H. C., Pilewski, J. M., Zhang, Q., DeLisser, H. M., Romer, L., and Albelda, S. M. (1995) Localization of multiple functional domains on human PECAM-1 (CD31) by monoclonal antibody epitope mapping. *Cell Adhes. Commun. 3*, 45–66.
19. Piali, L., Hammel, P., Uherek, C., Bachmann, F., Gisler, R. H., Dunon, D., and Imhof, B. A. (1995) CD31/PECAM-1 is a ligand for $\alpha_v\beta_3$ integrin involved in adhesion of leukocytes to endothelium. *J. Cell Biol. 130*, 451–460.
20. Deaglio, S., Morra, M., Mallone, R., Ausiello, C. M., Prager, E., Garbarino, G., Dianzani, U., Stockinger, H., and Malavasi, F. (1998) Human CD38 (ADP-ribosyl cyclase) is a counter-receptor of CD31, an Ig superfamily member. *J. Immunol. 160*, 395–402.
21. Prager, E., Sunder-Plassmann, R., Hansmann, C., Koch, C., Holter, W., Knapp, W., and Stockinger, H. (1996) Interaction of CD31 with a heterophilic counterreceptor involved in downregulation of human T cell responses. *J. Exp. Med. 184*, 41–50.
22. DeLisser, H. M., Yan, H. C., Newman, P. J., Muller, W. A., Buck, C. A., and Albelda, S. M. (1993) Platelet/endothelial cell adhesion

molecule-1 (CD31)-mediated cellular aggregation involves cell surface glycosaminoglycans. *J. Biol. Chem. 268*, 16037–16046.
23. DeLisser, H. M., Newman, P. J., and Albelda, S. M. (1993) Platelet endothelial cell adhesion molecule (CD31). *Curr. Top. Microbiol. Immunol. 184*, 37–45.
24. Jackson, D. E., Loo, R. O., Holyst, M. T., and Newman, P. J. (1997) Identification and characterization of functional cation coordination sites in platelet endothelial cell adhesion molecule-1. *Biochemistry 36*, 9395–9404.
25. Albelda, S. M., Muller, W. A., Buck, C. A., and Newman, P. J. (1991) Molecular and cellular properties of PECAM-1 (endoCAM/CD31): A novel vascular cell−cell adhesion molecule. *J. Cell Biol. 114*, 1059–1068.
26. Kasper, C., Rasmussen, H., Kastrup, J. S., Ikemizu, S., Jones, E. Y., Berezin, V., Bock, E., and Larsen, I. K. (2000) Structural basis of cell−cell adhesion by NCAM. *Nat. Struct. Biol. 7*, 389–393.
27. Soroka, V., Kolkova, K., Kastrup, J. S., Diederichs, K., Breed, J., Kiselyov, V. V., Poulsen, F. M., Larsen, I. K., Welte, W., Berezin, V., Bock, E., and Kasper, C. (2003) Structure and interactions of NCAM Ig1-2-3 suggest a novel zipper mechanism for homophilic adhesion. *Structure 11*, 1291–1301.
28. Watt, S. M., Williamson, J., Genevier, H., Fawcett, J., Simmons, D. L., Hatzfeld, A., Nesbitt, S. A., and Coombe, D. R. (1993) The heparin binding PECAM-1 adhesion molecule is expressed by $CD34^+$ hematopoietic precursor cells with early myeloid and B-lymphoid cell phenotypes. *Blood 82*, 2649–2663.
29. Sun, Q. H., Paddock, C., Visentin, G. P., Zukowski, M. M., Muller, W. A., and Newman, P. J. (1998) Cell surface glycosaminoglycans do not serve as ligands for PECAM-1. PECAM-1 is not a heparin-binding protein. *J. Biol. Chem. 273*, 11483–11490.
30. Carter, W. J., Cama, E., and Huntington, J. A. (2005) Crystal structure of thrombin bound to heparin. *J. Biol. Chem. 280*, 2745–2749.
31. Faham, S., Hileman, R. E., Fromm, J. R., Linhardt, R. J., and Rees, D. C. (1996) Heparin structure and interactions with basic fibroblast growth factor. *Science 271*, 1116–1120.
32. Mulloy, B., and Forster, M. J. (2000) Conformation and dynamics of heparin and heparan sulfate. *Glycobiology 10*, 1147–1156.
33. Mulloy, B., Forster, M. J., Jones, C., and Davies, D. B. (1993) NMR and molecular-modelling studies of the solution conformation of heparin. *Biochem. J. 293* (part 3), 849–858.
34. Pellegrini, L., Burke, D. F., von Delft, F., Mulloy, B., and Blundell, T. L. (2000) Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature 407*, 1029–1034.
35. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res. 22*, 4673–4680.
36. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res. 25*, 3389–3402.
37. Rost, B., Yachdav, G., and Liu, J. (2004) The PredictProtein server. *Nucleic Acids Res. 32*, W321–W326.
38. Bryson, K., McGuffin, L. J., Marsden, R. L., Ward, J. J., Sodhi, J. S., and Jones, D. T. (2005) Protein structure prediction servers at University College London. *Nucleic Acids Res. 33*, W36–W38.
39. Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol. 299*, 499–520.
40. Douguet, D., and Labesse, G. (2001) Easier threading through web-based comparisons and cross-validations. *Bioinformatics 17*, 752–753.
41. Pearson, W. R., and Lipman, D. J. (1988) Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A. 85*, 2444–2448.
42. Stockinger, H., Gadd, S. J., Eher, R., Majdic, O., Schreiber, W., Kasinrerk, W., Strass, B., Schnabl, E., and Knapp, W. (1990) Molecular characterization and functional analysis of the leukocyte surface protein CD31. *J. Immunol. 145*, 3889–3897.
43. Simmons, D. L., Walker, C., Power, C., and Pigott, R. (1990) Molecular cloning of CD31, a putative intercellular adhesion molecule closely related to carcinoembryonic antigen. *J. Exp. Med. 171*, 2147–2152.
44. Fiser, A., Do, R. K., and Sali, A. (2000) Modeling of loops in protein structures. *Protein Sci. 9*, 1753–1773.
45. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol. 234*, 779–815.
46. Harding, M. M. (2006) Small revisions to predicted distances around metal sites in proteins. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 62*, 678–682.
47. Coombe, D. R., and Kett, W. C. (2005) Heparan sulfate−protein interactions: Therapeutic potential through structure−function insights. *Cell. Mol. Life Sci. 62*, 410–424.
48. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem. 4*, 187–217.
49. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283–291.
50. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., and Wolfson, H. J. (2005) PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Res. 33*, W363–W367.
51. Schneidman-Duhovny, D., Inbar, Y., Polak, V., Shatsky, M., Halperin, I., Benyamini, H., Barzilai, A., Dror, O., Haspel, N., Nussinov, R., and Wolfson, H. J. (2003) Taking geometry to its edge: Fast unbound rigid (and hinge-bent) docking. *Proteins 52*, 107–112.
52. Imberty, A., Lortat-Jacob, H., and Perez, S. (2007) Structural view of glycosaminoglycan−protein interactions. *Carbohydr. Res. 342*, 430–439.
53. Shao, C., Zhang, F., Kemp, M. M., Linhardt, R. J., Waisman, D. M., Head, J. F., and Seaton, B. A. (2006) Crystallographic analysis of calcium-dependent heparin binding to annexin A2. *J. Biol. Chem. 281*, 31689–31695.
54. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem. 19*, 1639–1662.
55. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., and Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc. 106*, 765–784.
56. Mehler, E. L., and Solmajer, T. (1991) Electrostatic effects in proteins: Comparison of dielectric and charge models. *Protein Eng. 4*, 903–910.
57. Harpaz, Y., and Chothia, C. (1994) Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. *J. Mol. Biol. 238*, 528–539.
58. Caldwell, E. E., Nadkarni, V. D., Fromm, J. R., Linhardt, R. J., and Weiler, J. M. (1996) Importance of specific amino acids in protein binding sites for heparin and heparan sulfate. *Int. J. Biochem. Cell Biol. 28*, 203–216.
59. Fromm, J. R., Hileman, R. E., Caldwell, E. E., Weiler, J. M., and Linhardt, R. J. (1997) Pattern and spacing of basic amino acids in heparin binding sites. *Arch. Biochem. Biophys. 343*, 92–100.
60. Hileman, R. E., Fromm, J. R., Weiler, J. M., and Linhardt, R. J. (1998) Glycosaminoglycan−protein interactions: Definition of consensus sites in glycosaminoglycan binding proteins. *Bioessays 20*, 156–167.
61. Margalit, H., Fischer, N., and Ben-Sasson, S. A. (1993) Comparative analysis of structurally defined heparin binding sequences reveals a distinct spatial distribution of basic residues. *J. Biol. Chem. 268*, 19228–19231.
62. Forster, M., and Mulloy, B. (2006) Computational approaches to the identification of heparin-binding sites on the surfaces of proteins. *Biochem. Soc. Trans. 34*, 431–434.
63. Cardin, A. D., and Weintraub, H. J. (1989) Molecular modeling of protein−glycosaminoglycan interactions. *Arteriosclerosis 9*, 21–32.
64. Silvian, L., Jin, P., Carmillo, P., Boriack-Sjodin, P. A., Pelletier, C., Rushe, M., Gong, B., Sah, D., Pepinsky, B., and Rossomando, A. (2006) Artemin crystal structure reveals insights into heparan sulfate binding. *Biochemistry 45*, 6801–6812.
65. Buckley, C. D., Doyonnas, R., Newton, J. P., Blystone, S. D., Brown, E. J., Watt, S. M., and Simmons, D. L. (1996) Identification of $\alpha_v\beta_3$ as a heterotypic ligand for CD31/PECAM-1. *J. Cell Sci. 109* (part 2), 437–445.
66. Veldkamp, C. T., Peterson, F. C., Pelzek, A. J., and Volkman, B. F. (2005) The monomer−dimer equilibrium of stromal cell-derived factor-1 (CXCL 12) is altered by pH, phosphate, sulfate, and heparin. *Protein Sci. 14*, 1071–1081.
67. Matsumoto, R., Sali, A., Ghildyal, N., Karplus, M., and Stevens, R. L. (1995) Packaging of proteases and proteoglycans in the granules of mast cells and other hematopoietic cells. A cluster of

histidines on mouse mast cell protease 7 regulates its binding to heparin serglycin proteoglycans. *J. Biol. Chem. 270*, 19524–19531.
68. Leitinger, B., and Hogg, N. (2000) From crystal clear ligand binding to designer I domains. *Nat. Struct. Biol. 7*, 614–616.
69. Yang, W., Wilkins, A. L., Ye, Y., Liu, Z. R., Li, S. Y., Urbauer, J. L., Hellinga, H. W., Kearney, A., van der Merwe, P. A., and Yang, J. J. (2005) Design of a calcium-binding protein with desired structure in a cell adhesion molecule. *J. Am. Chem. Soc. 127*, 2085–2093.
70. Babor, M., Greenblatt, H. M., Edelman, M., and Sobolev, V. (2005) Flexibility of metal binding sites in proteins on a database scale. *Proteins 59*, 221–230.
71. Sachs, U. J. H., Andrei-Selmer, C. L., Maniar, A., Weiss, T., Paddock, C., Orlova, V. V., Choi, E. Y., Newman, P. J., Preissner, K. T., Chavakis, T., and Santoso, S. (2007) The neutrophil-specific antigen CD177 is a counter-receptor for platelet endothelial cell adhesion molecule-1 (CD31). *J. Biol. Chem. 282*, 23603–23612.
72. Zhang, C., Vasmatzis, G., Cornette, J. L., and DeLisi, C. (1997) Determination of atomic desolvation energies from the structures of crystallized proteins. *J. Mol. Biol. 267*, 707–726.
73. Capila, I., and Linhardt, R. J. (2002) Heparin−protein interactions. *Angew. Chem., Int. Ed. 41*, 391–412.
74. Coombe, D. R., Stevenson, S. M., Kinnear, B. F., Gandhi, N. S., Mancera, R. L., Osmond, R. I. W., and Kett, W. C. Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 2. Biochemical analyses. *Biochemistry*, in press.

BI702455E