Docking of Glycosaminoglycans to Heparin-Binding Proteins: Validation for aFGF, bFGF, and Antithrombin and Application to IL-8

Wolfgang Bitomsky and Rebecca C. Wade*

Contribution from the European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117-Heidelberg, Germany

Received September 17, 1998. Revised Manuscript Received December 4, 1998

Abstract: Heparin−protein interactions play an important role in many steps of the immune system. Here, we evaluated the search capabilities of three widely used programs—GRID, DOCK, and AutoDock—for heparin binding sites. Because of the weak surface complementarity and the high charge density of the sulfated sugar chain, the docking of heparin to its protein partners presents a challenging task for computational docking. Our protocols were tested on antithrombin and acidic and basic fibroblast growth factor, the only three proteins for which structures of their complexes with heparin are available. With all three programs, the heparin binding site for these test cases was determined correctly. We then used these protocols to predict the heparin binding site on Interleukin-8, a chemokine with a central role in the human immune response. The results indicate that His18, Lys20, Arg60, Lys64, and Arg68 in interleukin-8 bind to heparin.

Introduction

The interaction of glycosaminoglycans (heparin and heparan sulfate) with proteins and peptides[1] plays a role in the regulation of many physiological processes such as hemostasis, growth factor activity, anticoagulation, cell adhesion, and enzyme regulation.[2−5] As such, this interaction is clearly an important potential target for drug design. Glycosaminoglycan chains consist of repeating disaccharides, constructed of alternating uronic acids and glucosamines, in which the uronic acid may be either D-glucuronic acid or L-iduronic acid, and the D-glucosamine residue may be either N-acetylated or N-sulfated. The disaccharide units are O-sulfated to varying degrees at C6 and/or C3 of the glucosamine residues and at C2 of the uronic acids.[6,7]

Given the importance of protein−glycosaminoglycan interactions, considerable effort has been invested in the identification of protein sequences that interact specifically with heparin/heparan sulfate to extract a consensus binding sequence. Cardin *et al.* analyzed the sequences of heparin binding sites, suggesting two different sequence patterns for binding to α-helical and β-strand regions.[8] Margalit and co-workers suggested a spacing of 20 Å between positively charged residues independent of the secondary structure at the binding site.[9] It seems clear that positively charged residues play a critical role in the interaction with the sulfate and carboxylate groups of the glucosaminoglycan (GAG) chain, but an energetic characterization of the bFGF (FGF = fibroblast growth factor) interactions showed an electrostatic contribution of these residues of only 30% to the binding free energy.[10] How proteins discriminate between sulfated sugars is not completely understood, but for some proteins a clear binding motif or selectivity for a specific saccharide composition has been reported.[11−14]

Several models of protein−heparin complexes were developed prior to the experimental determination of structures of protein−GAG complexes.[8,15,16] The first crystal structure was determined by Faham *et al.*[17] for the bFGF−heparin complex, providing a detailed picture of the atomic interactions at the binding site (see Figure 1). In the last 2 years, two more structures of complexes, for aFGF and antithrombin, have been determined and deposited in the Brookhaven protein database.[18,19] In view of the limited amount of structural information available, computational docking methods provide a useful tool to develop models for heparin−protein complexes. The docking of GAGs to their protein partners represents a challenging task

* Address correspondence to this author.

(1) Kjellén, L.; Lindahl, U. *Annu. Rev. Biochem.* **1991**, *60*, 443−475.
(2) Lindahl, U.; Lidholt, K.; Spillman, D.; Kjellén, L. *Thromb. Res.* **1994**, *75*, 1−32.
(3) Yayon, A.; Klagsbrun, M.; Esko, J. D.; Leder, P.; Ornitz, D. M. *Cell* **1991**, *64*, 841−848.
(4) Prestrelski, S.; Fox, G. M.; Arakawa, T. *Arch. Biochem. Biophys.* **1992**, *293*, 314−319.
(5) Bjork, I.; Lindahl, V. *Mol. Cell. Biochem.* **1982**, *48*, 161−182.
(6) Comper, W. D. *Heparin and related polysaccharides;* Gordon & Breach: New York, 1981.
(7) Gallagher, J. T.; Walker, A. *Biochem. J.* **1985**, *230*, 665−674.
(8) Cardin, A. D.; Detemer, D. A.; Weintraub, H. J. R.; Jackson, R. L. *Methods Enzymol.* **1991**, *203*, 556−583.
(9) Margalit, H.; Fischer, N.; Ben-Sasson, S. A. *J. Biol. Chem.* **1993**, *268*, 19228−19231.
(10) Thompson, L. D.; Pantoliano, M. W.; Springer, B. A. *Biochemistry* **1994**, *33*, 3831−3840.
(11) Raggazzi, M.; Ferro, D. R.; Perly, B.; Sinay, P.; Petitou, M.; Choay, J. *Carbohydr. Res.* **1990**, *195*, 169−185.
(12) Witt, D. P.; Lander, A. D. *Curr. Biol.* **1994**, *4*, 394−400.
(13) Maccarana, M.; Casu, B.; Lindahl, U. *J. Biol. Chem.* **1993**, *268*, 23898−23905.
(14) Turnbull, J. E.; Fernig, D. G.; Ke, Y.; Wilkinson, M. C.; Gallagher, J. T. J. *Biol. Chem.* **1992**, *267*, 10337−10341.
(15) Grootenhuis, P. D. J.; van Boeckel, C. A. A. *J. Am. Chem. Soc.* **1991**, *113*, 2743−2747.
(16) Stuckey, J. A.; Charles, R. S.; Edwards, B. F. P. *Proteins: Struct. Funct. Genet.* **1992**, *14*, 277−287.
(17) Faham, S.; Hileman, R. E.; Fromm, J. R.; Linhardt, R. J.; Rees, D. C. *Science* **1996**, *271*, 1116−1120.
(18) DiGabriele, A. D.; Lax, I.; Chen, D. I.; Svahn, C. M.; Jaye, M.; Schlessinger, J.; Hendrickson, W. A. *Nature* **1998**, *393*, 812−817.
(19) Jin, L.; Abrahams, J. P.; Skinner, R.; Petitou, M.; Pike, R. N.; Carrell, R. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 14683−14688.

10.1021/ja983319g CCC: $18.00
Published on Web 03/19/1999

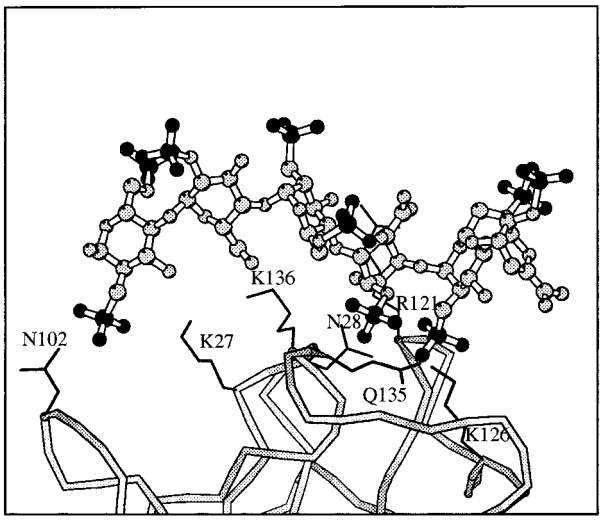

Figure 1. Heparin binding site of bFGF with bound heparin hexasaccharide (PDB-entry: 1bfc).[17] The protein is shown as a ribbon with side chains of important residues at the binding site included in the plot (high affinity site: N28, R121, K126, Q135; low affinity site: K27, N102, K136). The bound hexasaccharide is shown in ball-and-stick, with the atoms of the sulfate groups colored black.

because of the weak surface complementarity, the high charge density of heparin and the heparin binding site, and the highly flexible nature of the GAG chain.

Given the clear importance of electrostatic interactions for heparin binding, we initally conducted an analysis of the electrostatic potentials of the proteins computed by solution of the finite difference Poisson Boltzmann equation. This indicated regions of positive potential on the proteins (putative heparin binding sites) and how potentials were affected by conformational differences between the X-ray and NMR structures of IL-8. However, ligand binding modes could not be obtained by this approach and so we aimed to develop a fully automated protocol for the docking of heparin/heparan sulfate to its protein partners.

In the first step, we tested several programs for their capability to detect heparin binding sites on proteins, without restricting the search to a certain protein region, while in the second step the accuracy of 3D models consisting of acidic or basic FGF and a docked hexasaccharide was evaluated. To treat two of the properties of the protein−heparin interactions mentioned in the previous paragraph that make modeling these interactions challenging, the high charge density and the flexibility, different approaches were used in the search for binding sites. The GRID program[20] was chosen because it has a highly developed force field for the description of protein−ligand interactions. For the representation of flexibility, we chose the AutoDock and DOCK programs,[21,22] because both programs allow the handling of flexible ligands during docking runs. The results of both steps (binding site identification and docking of hexasaccharide) were evaluated by using known crystal structures of protein−heparin complexes.

Our particular interest is focused on the development of an Interleukin-8(IL-8)-heparin model. IL-8 is a member of the chemokine family, which plays an important role in the human immune response and is often upregulated in inflammation. Glycosaminoglycans represent the endothelial binding site for

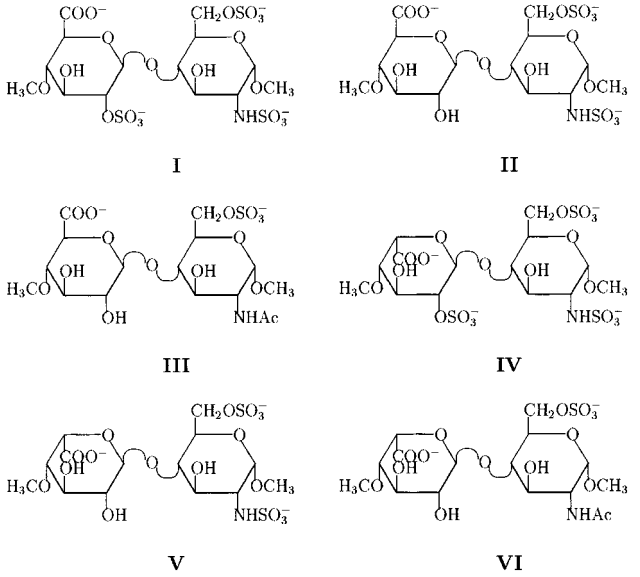

Figure 2. Formulas of the disaccharide probes: (**I**) 1-methoxy-2-deoxy-2-sulfamido-D-glucopyranosyl-6-*O*-sulfate (1→4)4-methoxy-L-glucuronic acid 2-*O*-sulfate; (**II**) 1-methoxy-2-deoxy-2-sulfamido-D-glucopyranosyl-6-*O*-sulfate(1→4)4-methoxy-L-glucuronic acid; (**III**) 1-methoxy-2-deoxy-2-acetamido-D-glucopyranosyl-6-*O*-sulfate(1→4)4-methoxy-L-glucuronic acid; (**IV**) 1-methoxy-2-deoxy-2-sulfamido-D-glucopyranosyl-6-*O*-sulfate(1→4)4-methoxy-L-iduronic acid 2-*O*-sulfate; (**V**) 1-methoxy-2-deoxy-2-sulfamido-D-glucopyranosyl-6-*O*-sulfate(1→4)4-methoxy-L-iduronic acid; (**VI**) 1-methoxy-2-deoxy-2-acetamido-D-glucopyranosyl-6-*O*-sulfate(1→4)4-methoxy-L-iduronic acid. Molecules can be divided into two subtypes: **glp** (glucuronic acid probes, structures **I**−**III**) and **idp** (iduronic acid probes, structures **IV**−**VI**); molecule **III** corresponds to the major repeating unit of heparan sulfate, while molecule **VI** corresponds to that in heparin.

various chemokines and enhance their effects on high-affinity receptors.[23] The C-terminal α-helix of IL-8 has been shown to be important for this interaction.[24] Two possible binding modes have been discussed in the literature: first, a perpendicular orientation between the heparin and the helical axes, which would correspond to the platelet factor 4 (PF4)-heparin model of Stuckey *et al.*,[15] and second, an orientation parallel to the helical axis as suggested by Spillmann *et al.*,[25] causing a horseshoe fashion binding of heparin to the IL-8 dimer.

Methods

The following terminology is used throughout the paper: "global", the whole protein was searched for ligand binding sites; "local", only potential interaction sites, detected in the preceding global searches for ligand binding sites, were searched; glp, disaccharide probes including a glucuronic acid residue (structures **I**−**III** in Figure 2); idp, disaccharide probes including an iduronic acid residue (structures **IV**−**VI** in Figure 2); monomers, all monosaccharide probes with coordinates extracted from the Protein Databank files 1bfc and 1hpn (the latter contains the solution structure of a heparin dodecamer[26]); dimers, all disaccharide probes with coordinates extracted from the 1bfc and 1hpn files.

Software. For the global search for heparin binding sites on each protein studied, the programs GRID (version 15, Molecular Discovery

(20) Goodford, P. J. *J. Med. Chem.* **1985**, *28*, 849−857.
(21) Goodsell, D. S.; Morris, G. M.; Olson, A. J. *J. Mol. Recognit.* **1996**, *9*, 1−5.
(22) Ewing, T. J. A.; Kuntz, I. D. *J. Comp. Chem.* **1997**, *18*, 1175−1189.
(23) Hoogwerf, A. J.; Kuschert, G. S. V.; Proudfoot, A. E. I.; Borlat, F.; Clark-Lewis, I.; Power, C. A.; Wells, T. N. C. *Biochemistry* **1997**, *36*, 13570−13578.
(24) Webb, L. M. C.; Ehrengruber, M. U.; Clark-Lewis, I.; Baggiolini, M.; Rot, A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7158−7162.
(25) Spillmann, D.; Witt, D.; Lindahl, U. *J. Biol. Chem.* **1998**, *273*, 15487−15493.
(26) Mulloy, B.; Forster, M. J.; Jones, C.; Davies, D. B. *Biochem. J.* **1993**, *293*, 849−858.

Ltd.),[20,27−29] AutoDock (version 2.4),[21] and DOCK (version 4.0)[22] were used. Subsequent docking of heparin hexamers was performed with the AutoDock and DOCK programs. The InsightII package (BIOSYM Technologies) was used for modeling substituents on the GAGs and the addition of methoxy groups at the GAG chain ends. The SYBYL program package (TRIPOS inc.) was used to generate missing hydrogen atoms for protein and probe structures used in the AutoDock (polar hydrogens only) and DOCK (all hydrogens) runs. The figures were made with the InsightII and Molscript software.[30] The hydrogen bonds between the docked hexasaccharides and the proteins were calculated with use of the LIGPLOT program (version 4.0).[31]

Protein Coordinates. The following Brookhaven database entries were used: bFGF, 1bfc;[17] aFGF, the A chain of 2axm;[18] antithrombin III (ATIII), the I-chain of 2ant;[32] IL-8, 3il8[33] (crystal), and 1il8[34] (NMR) structure. So our test set includes two structures taken from protein−heparin complexes (aFGF and bFGF) and one unbound protein (the structure of the antithrombin−heparin complex was not available at the time of this study). In the case of bFGF, the missing side chain of residue 61 was modeled with use of the WHATIF program.[35] Prior to the docking runs, all protein structures were overlaid on antithrombin to allow the same starting coordinates to be used for all the docking runs.

GAG Coordinates. Starting geometries of the disaccharide probes (Figure 2) were taken from the bFGF−heparin complex (residues IDU304 and SGN305). The coordinates of the C atoms of the methoxy end groups correspond to the coordinates of the C1 and C4 atoms of the respectively preceding or following sugar ring. For the GRID program, additional mono- and disaccharide probes (monomers, dimers) were used for the representation of structural flexibility. For the hexasaccharide docking runs, the initial coordinates were taken from 1bfc for the bFGF and IL-8 trials and from 2axm for docking to aFGF. Prior to the docking runs, all probes were placed arbitrarily at a distance of 30 Å from the protein surfaces of antithrombin.

Partial Charges. For AutoDock and DOCK, the partial charges for the protein atoms were taken from the AMBER force field[36,37] (united atoms for AutoDock, all atoms for DOCK). The following charges were used for the probe atoms for all programs: C 0.05e; C(COO^-) 0.500e; H (bound to C atom) 0.080e; H (hydroxyl group) 0.400e; H (bound to N) 0.250e; N −0.735e; S 1.305e; O (sulfate group) −0.650e; O (carboxylate group) −0.700e; O (hydroxyl group) −0.500e; O (C−O−S) −0.550e; O (ether group) −0.300e. For Autodock runs, these charges were modified so that the charges of the nonpolar hydrogen atoms were assigned to the atom to which the hydrogen atom is bonded.

Protocols. GRID/GROUP. The following single atom probes were used to represent the polar groups of the mono- and disaccharide molecules: O:: (sp^2 carboxy oxygen), O= (sulfate oxygen), O1 (alkyl hydroxyl oxygen), OC2 (ether oxygen), N1: (sp^3 NH with lone pair), H (hydrogen atom), and OH2 (water). Interaction energy maps for these probes were computed on a grid with 1 Å spacing. This grid spacing is required (as are maps for the last two probes) for running the GROUP module which docks ligands by optimizing their location on the interaction energy maps.

AutoDock: Grids of probe atom interaction energies were computed first−50 Å side grids with a spacing of 0.5 Å for aFGF, bFGF, and IL-8, while 65 Å side grids were used for antithrombin. The ligand probes were then docked by simulated annealing according to the protocol of Coutinho *et al.*,[38] except that the ratio of accepted/rejected steps was increased to 3000. All rotatable torsion angles were allowed to rotate freely, and the 100 lowest energy structures were stored for further analysis.

DOCK: A 1.4 Å radius probe was used to compute the protein molecular surface that was subsequently represented by sets of overlapping spheres with use of the sphgen program from the DOCK software package (radmin 1.4 Å; radmax 4.0 Å; dotlim -0.5). All generated clusters were used for the calculation of grids of interaction energies for probe atoms which had a spacing of 0.5 Å and used the all-atom model implemented in the DOCK program package. Ligands were docked by optimizing overlap with the protein spheres and computing interaction energies with the grids. The maximum number of orientations and conformations of the probe was limited to 1000, and the 100 lowest energy structures were saved.

Hexasaccharide Docking. Prior to all docking runs, the protocols were tested by performing rigid-body docking of heparin hexamers to aFGF and bFGF with the DOCK and AutoDock programs. The influence of the crystalline environment was evaluated for bFGF by performing rigid body docking with all crystallographic neighbors included in the calculations. Then docking of monomers and dimers was performed: these were flexible for DOCK and Autodock runs and rigid for GRID/GROUP runs. After this, docking of flexible hexasaccharides to aFGF, bFGF, and IL-8 was performed with DOCK and AutoDock. For local docking with the AutoDock program, the center of the grid map was evaluated by simply averaging the coordinates of the docked dimers at a certain interaction site, and then a 35 Å side grid was constructed around this point (grid spacing 0.5 Å). For local docking with the DOCK program, all residues with an atom−atom distance of less than 4 Å from any docked structure were included in the molecular surface calculation. For IL-8, only local docking runs were performed for the IL-8 monomer. Both the crystal and the NMR structures of IL-8 monomer were used, whereas for the IL-8 dimer, only the NMR structure was used.

Because test runs showed some interaction energy differences between local and global searches for the crystal structure complexes with the 0.5 Å spacing grids, all interaction energies were recalculated with a 0.2 Å spacing grid to reduce interpolation errors. A complete search with the smaller grid spacing was not possible due to computational limits.

Analysis. The following protocol was used for the analysis of the global mono- and disaccharide docking runs. In the GRID/GROUP runs, all output docked ligand structures are within a prescribed energy cutoff of the lowest energy docked structure. On the other hand, the output from the AutoDock and DOCK programs is the 100 lowest energy structures according to their force-field scoring functions. Therefore, to obtain subsets of AutoDock and DOCK structures that could be compared with those obtained with GRID, for each program only structures within a chosen energy cutoff of the lowest energy structure were used. A different energy cutoff was used for each program. The energy cutoff was chosen to correspond to the smallest of the energy ranges for the 100 structures obtained with each of the 6 disaccharide probes (Figure 2). Having obtained a set of structures for each probe that lie in the same energy interval, the total number of contacts per residue was calculated with use of a 4 Å cutoff. This number was normalized per probe (according to the number of structures) and afterward the average and standard deviations of this "interaction probability" were calculated per probe subtype (glp, idp, monomer, and dimers−cf.Tables 1−7). For the comparison with the crystal structures, the same 4 Å cutoff was used for the identification of the interacting residues in aFGF[18] and bFGF.[17,39] For antithrombin, the identity of the interacting residues was taken from the literature[15,19] because the coordinates of the complex with heparin were not available

(27) Boobbyer, D. N. A.; Goodford, P. J.; McWhinnie, P. M.; Wade, R. C. *J. Med. Chem.* **1989**, *32*, 1083−1094.
(28) Wade, R. C.; Clark, K. J.; Goodford, P. J. *J. Med. Chem.* **1993**, *36*, 140−147.
(29) Wade, R. C.; Goodford, P. J. *J. Med. Chem.* **1993**, *36*, 148−156.
(30) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.
(31) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. *Protein Eng.* **1995**, *8*, 127−134.
(32) Skinner, R.; Abrahams, J.-P.; Whisstock, J. C.; Lesk, A. M.; Carrell, R. W.; Wardell, M. R. *J. Mol. Biol.* **1997**, *266*, 601−609.
(33) Baldwin, E. T.; Weber, I. T.; Charles, R. S.; Xuan, J.-C.; Appella, E.; Yamada, M.; Matsushima, K.; Edwards, B. F. P.; Clore, G. M.; Gronenborn, A. M.; Wlodawer, A. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 502−506.
(34) Clore, G. M.; Appella, E.; Yamada, M.; Matsushima, K.; Gronenborn, A. M. *Biochemistry* **1990**, *29*, 1689−1696.
(35) Vriend, G. *J. Mol. Graph.* **1990**, *8*, 52−56.
(36) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comp. Chem.* **1986**, *7*, 230-252.
(37) Weiner, S. J.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1984**, *106*, 765−784.
(38) Coutinho, P. M.; Dowd, M. K.; Reilly, P. J. *Proteins: Struct. Funct. Genet.* **1997**, *28*, 162−173.
(39) Ornitz, D. M.; Herr, A. B.; Nilsson, M.; Westman, J.; Svahn, C.-M.; Waksman, G. *Science* **1995**, *268*, 432−436.

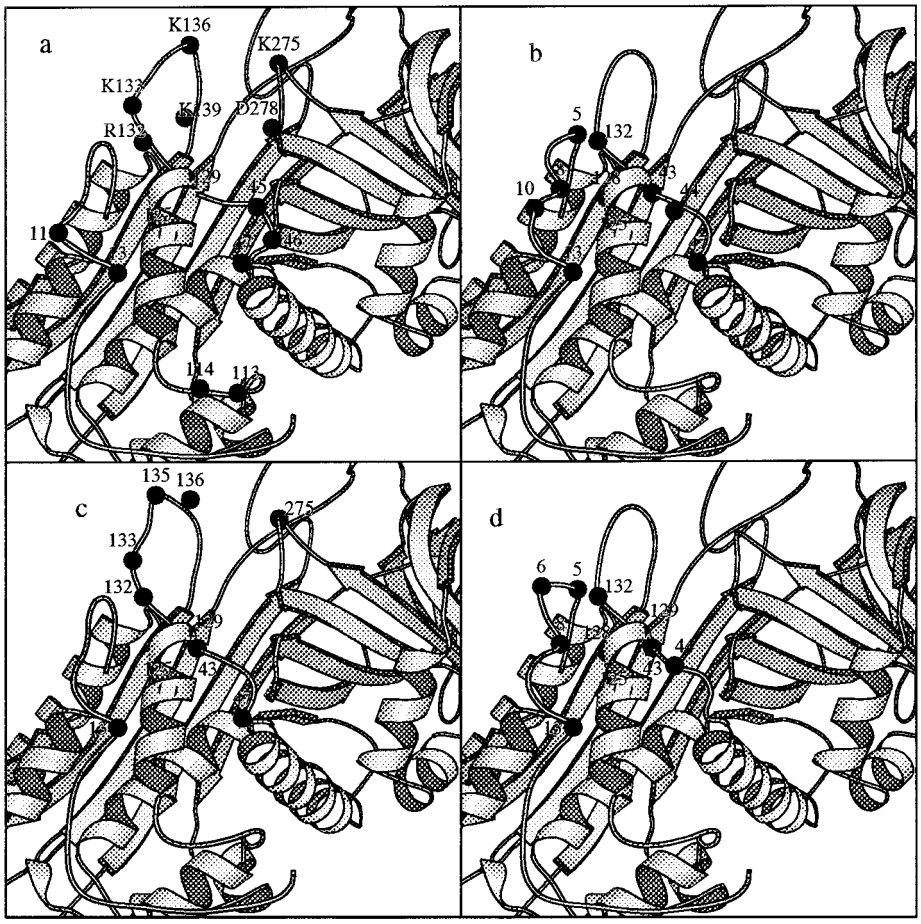

Figure 3. Heparin binding site of antithrombin. Interacting residues are labeled and their Cα atoms are shown as spheres: (a) crystal structure—residues identified by Jin *et al.*[19] in the crystal structure (res. 11, 13, 45, 46, 47, 113, 114, 125, 129) and by Grootenhuis *et al.*[15] in their modeled structure (res. R132, K133, K136, K139, K275, D278); in the case of the docking runs—(b) GRID; (c) AutoDock; (d) DOCK—only the 10 residues with the highest interaction probabilities are labelled.

Table 1. Interaction Probabilities for Selected Residues of Antithrombin[a]

	residue number														
run/probe	K11*	R13*	N45*	R46*	R47*	E113*	K114*	K125*	R129*	R132	K133	K136	K139	K275	D278
GRID															
monomers		**0.18**	0.05	0.11	0.16	0.08	0.05	**0.18**	0.08	0.11	0.05	0.03	0.03		
dimers		**0.57**	0.03	0.05	0.40	0.05	0.05	0.50	0.12	0.15	0.03		0.10		
glp		0.44	0.14		0.43			**0.52**	0.10	0.38	0.33	0.17	0.33		
idp		**0.70**	0.11	0.04	0.32	0.04		**0.70**	0.44	0.44	0.07				
AutoDock															
glp	0.09	**0.22**	0.09	0.02	0.16		0.01	0.18	0.14	0.18	0.16	0.17	0.10	0.17	0.04
idp	0.08	0.13	0.04	0.06	0.07		0.01	0.07	0.08	0.10	**0.15**	**0.15**	0.05	0.11	0.02
DOCK															
glp		**0.49**	0.16		0.04	0.01	0.01	0.43	0.43	0.27	0.18		0.10		0.05
idp	0.01	0.54	0.09	0.01	0.01			0.53	**0.65**	0.56	0.38	0.01	0.04	0.01	0.01

[a] Only the residues of the binding site in the experimental structures (res. 11, 13, 45, 46, 47, 113, 114, 125, 129—marked with asterisks)[19] and in the model of Grootenhuis *et al.* (res. 132, 133, 136, 275, 278)[15] are listed. The maximum standard deviation is 0.23 for GRID, 0.01 for AutoDock, and 0.20 for DOCK results (see Methods section for details). The highest probabilities per probe are shown in bold.

at the time of this study. For IL-8, the 10 residues with the highest interaction probabilities are shown (Table 6). For the hexasaccharide docking runs, the 100 lowest energy structures, based on the force field scoring, were used for the analysis. The interacting residues were calculated in the same way as for the mono- and disaccharide docking runs, and then the docked structures were clustered by using a 5 Å RMSD cutoff for all non-hydrogen atoms.

Results and Discussion

Rigid-Body Docking of Heparin Hexamers. For the preliminary testing of the docking protocols, rigid body docking runs were performed on bFGF and aFGF. For bFGF, the low-energy structures show small deviations from the crystal structure coordinates (DOCK 1.3 Å; AutoDock 2.5 Å). Larger deviations were observed in the aFGF case (DOCK 7.1 Å; AutoDock 7.5 Å). This might be because binding is less specific to aFGF than to bFGF (see below) or because a second aFGF molecule was not included in the docking target. In the crystal structure of aFGF, the hexasaccharide is sandwiched between two aFGF molecules and makes hydrogen bonds to both protein molecules. In the crystal structure of the complex of bFGF, only the position of the last three residues of the bound hexasaccharide is influenced by crystal contacts. The inclusion of the

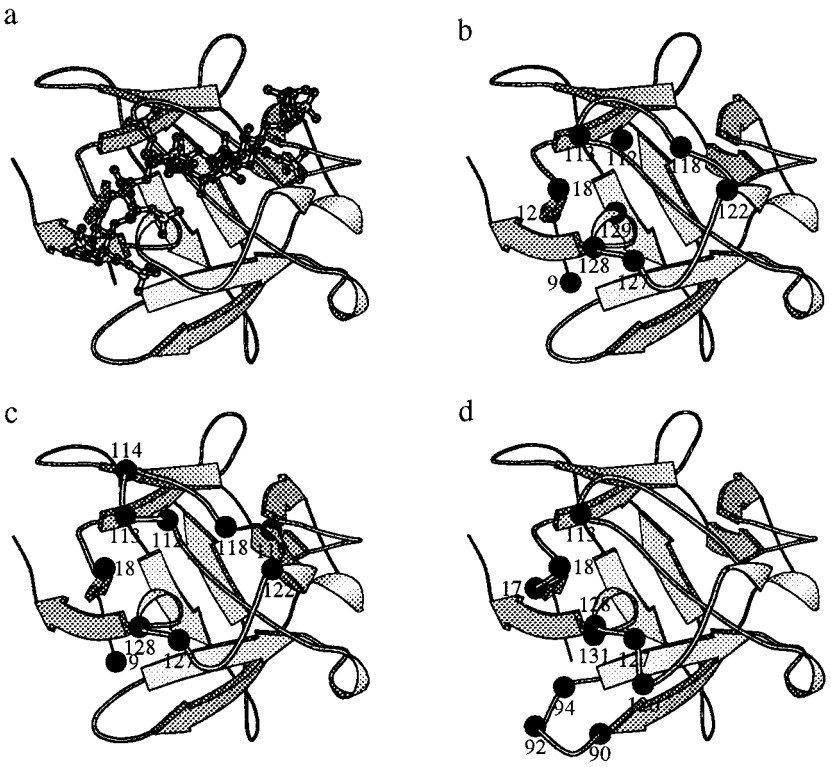

Figure 4. Heparin binding site of aFGF: (a) crystal structure of aFGF[18] with bound hexasaccharide; in the case of the docking runs—(b) GRID; (c) AutoDock; (d) DOCK—only the 10 residues with the highest interaction probabilities are shown (Cα atoms shown as spheres and labeled).

Table 2. Interaction Probabilities for Selected Residues of aFGF[a]

	residue number									
run/probe	N18	L111	K112	K113	K118	R122	G126	Q127	K128	A129
GRID										
monomers	0.38	0.12	**0.58**	0.42	0.52	0.27	0.02	0.18	0.22	0.25
dimers	0.38	0.04	0.56	0.40	**0.62**	0.54		0.29	0.15	0.19
glp	**1.00**		0.67	0.67	**1.00**	0.67		**1.00**	0.67	0.67
idp	0.23	0.04	**0.40**	0.23	0.36	0.32	0.04	0.27	0.18	0.13
AutoDock										
glp	0.35	0.02	**0.48**	0.46	0.46	0.41	0.05	0.23	0.27	0.16
idp	0.15		0.25	**0.27**	0.18	0.25	0.04	0.11	0.17	0.06
DOCK										
glp	0.77	0.01	**0.98**	0.80	0.96	0.88		0.32	0.17	0.30
idp	0.60		**0.96**	0.72	0.85	0.56	0.01	0.35	0.45	0.32

[a] Only residues with atoms within a 4 Å cutoff of the bound saccharide in the experimental structures[18] are listed. The maximum standard deviation is 0.11 for the GRID, 0.03 for AutoDock, and 0.04 for DOCK results. The highest probabilities per probe are shown in bold.

symmetry-related nearest neighbors for aFGF and bFGF did not improve the docking results, but this may be because sampling was reduced due to the greater computational requirements for docking to a larger target.

Docking of Mono- and Disaccharides. (a) Antithrombin: Antithrombin has a single binding site for sulfated saccharides but docking is complicated by a conformational change in the protein that occurs upon binding.[40] Most of the residues identified in the global docking runs (Figure 3) correspond to those in the experimentally determined binding site[19] and the residues suggested in the modeling study of Grootenhuis *et al.*[15] The additional residues that are identified in the GRID and DOCK runs (Val5, Asp6, Cys8, Ala10) are in close proximity to the upper part of the D-helix (residues 113−130) that is part of the experimentally determined binding site. A more extensive analysis of the residues of the binding site (Table 1) revealed Arg13, Lys125, and Arg129 as important interaction partners, while Lys114, which forms three hydrogen bonds in the crystal structure,[19] shows only a low interaction probability with the docked ligands. Residues 132 to 139 and 275 and 278, suggested by the modeling study of Grootenhuis and co-workers, show only low interaction probabilities. The small interaction probabilities for these residues might be due to the size of our probes (mono- and disaccharides), as these residues might show more interactions with GAG chains longer than 5 residues.[19]

(b) aFGF: The structure of aFGF complexed with heparin was recently determined by DiGabriele *et al.*[18] and consists of a fully active dimer of FGF promoters, in which each monomer is linked by a heparin decasaccharide, without a direct protein−protein contact. As for the antithrombin case, all three programs were able to localize the heparin binding site (Figure 4). In the GRID and AutoDock runs, most of the identified residues correspond to those in the experimentally determined binding site, while in the DOCK case, the preferred site is shifted in the direction of the loop region between residues 89 and 95. The interaction probabilities are shown in Table 2, suggesting

(40) Evans, D. L.; Marshall, C. J.; Christey, P. B.; Carrell, R. W. *Biochemistry* **1992**, *31*, 12629−12642.

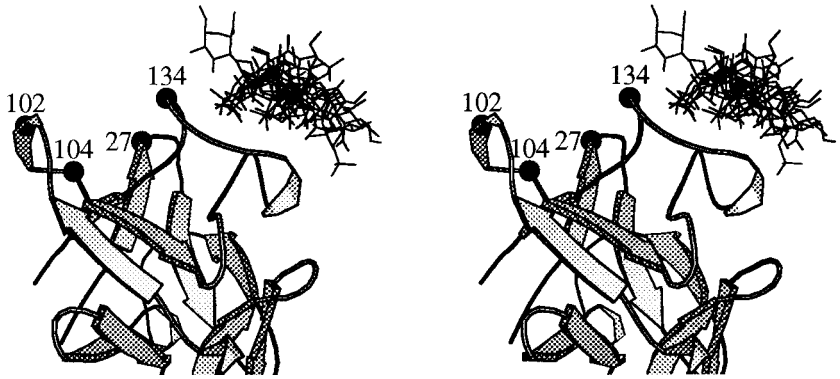

Figure 5. bFGF with 10 disaccharide structures docked with the AutoDock probram. All the structures are located at the "high preference" binding site of bFGF. The residues of the "low preference" binding site are labeled and their Cα atoms are shown as spheres.

Table 3. Interaction Probabilities for Selected Residues of bFGF (1st binding site)[a]

	residue number											
run/probe	K27	N28	N102	T104	K120	R121	K126	K130	G134	Q135	K136	A137
GRID												
monomers		0.58			**0.64**	0.55	0.68	0.42	0.04	0.36	0.30	0.47
dimers		0.23			**0.55**	0.32	0.47	0.43		0.13	0.11	0.13
glp		**0.67**			**0.67**	**0.67**	**0.67**	**0.67**		0.33	0.33	0.33
idp		0.50			**0.80**	0.57	0.60	0.47		0.03		0.50
AutoDock												
glp	0.33	0.61	0.03	0.09	0.49	0.66	0.60	0.37	0.17	0.49	**0.70**	0.37
idp	**0.14**	0.08	0.06	0.07	0.10	0.12	0.06	0.10	0.05	0.04	**0.14**	0.02
DOCK												
glp	0.15	0.62	0.08	0.13	0.54	0.65	0.66	0.54	0.05	0.48	**0.67**	0.37
idp	0.21	0.47	0.08	0.12	0.49	0.52	0.52	0.32	0.07	0.34	**0.55**	0.24

[a] Only residues with atoms within a 4 Å cutoff of the bound saccharides in the experimental structures (res. 27, 28, 102, 104, 120, 121, 126, 135, 136, 137)[17] and (res. 130, 134)[39] are listed. The maximum standard deviation is 0.22 for the GRID, 0.03 for AutoDock, and 0.08 for DOCK results. The highest probabilities per probe are shown in bold.

that residues 112, 113, and 118 play a very important role in the binding interactions. Two of this residues, Lys112 and Lys118, take part in the formation of two important sulfate group binding sites in the crystal structure.[18]

(c) bFGF: In contrast to the preceding proteins, two binding sites have been identified for sulfated saccharides on bFGF. The first site is occupied in the tetra-/hexasaccharide structure solved by Faham and co-workers,[17] while the second binding site has been reported only for trisaccharides.[39] All three programs show a strong preference for the first binding site (Figure 5), while only a few docked probes are located around the second site (data not shown). A more extensive analysis of the interactions at the first binding site (Table 3) suggested the splitting of this site into two patches—a "high preference" and a "low preference" patch. The high preference patch includes the residues 28, 120, 121, 126, 135, and 136, while the "low preference" one consists of residues 27, 102, 104, and 134. This distinction between the two binding regions corresponds with the one made by Faham and co-workers for high and low affinity binding sites based on their crystal structure.[17] The importance of some of the "high preference" residues (residues 28, 121, 126, and 135) was demonstrated by Thompson *et al.*,[39] who showed that the contribution to the binding free energy of each of these residues was more than 1 kcal/mol.

Docking of GAG Hexasaccharides. The interaction probabilities for the hexasaccharides are shown in Tables 4 and 5.

(a) aFGF: In the case of aFGF, residues 112 and 113 show the highest probabilities, which corresponds to the results of the global docking runs with smaller probes (Table 2). In the crystal structure of the aFGF complex, three sulfate groups (residues 301 to 303) are pointing toward the protein surface, and two of them (302 and 303) occupy two important sulfate binding sites formed by Asn18, Lys118, and Gln127.[18] In the minimum energy structure from the AutoDock run (colored yellow in Figure 6a), three sulfate groups come near these sites, but their sulfur atoms are 2 to 4 Å apart from those in the crystal structure. In addition, the location of all the GAG residues is shifted about two residues so that the sulfate group of IDU303 comes close to the sulfate group of SGN301 of the crystal structure, resulting in an RMSD of 9.3 Å for the non-hydrogen atoms. The minimum energy structure from the DOCK run (colored pink in Figure 6a) shows a similar conformation to that from the AutoDock run, but the sulfate groups are a bit closer to those in the crystal structure, also resulting in an RMSD of 9.3 Å for the non-hydrogen atoms. These deviations from the crystal structure show similarity to the structural variations found in the crystal structures of the aFGF−heparin complexes (PDB entries: 1axm and 2axm).[18] Overlaying the three dimer structures from the asymmetric unit of the orthorhombic crystal (1axm) on the monomer A of the hexagonal crystal (2axm) revealed that only two (residues 302 and 303) out of the three sulfate positions mentioned above are observed in all structures, but these positions show deviations of up to 1.5 Å for the sulfur atoms. The position of the third sulfate group in monomer A (residue 301) is not always observed in the other structures because of a shift of the GAG chain by about one saccharide unit (dimer B) and the inversion of the chain orientation (dimer C).

(b) bFGF: In the crystal structure of bFGF, the sulfate groups of residues 302 and 303 interact with the high affinity binding region of the protein.[17] The minimum energy structure from the AutoDock run (colored yellow in Figure 6b) shows a similar conformation for the first three residues (301 to 303) to the crystal structure, with the sulfur atoms of the O2 groups of

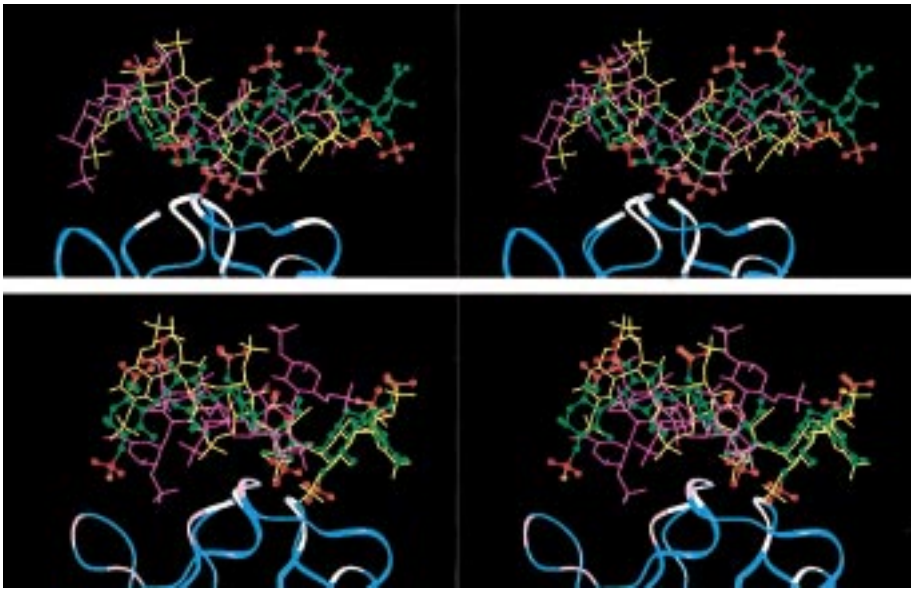

Figure 6. (a, top) Hexasaccharide structures at the binding site of aFGF. The crystal structure of the bound hexasaccharide is green with sulfate groups shown in red and represented as ball-and-sticks, the AutoDock minimum energy structure is yellow, and the DOCK minimum energy structure is magenta. The protein is represented as a ribbon and the residues of the binding site (res. 18, 111, 112, 113, 118, 122, 126, 127, 128, 129) are white. (b, bottom) Hexasaccharide structures at the binding site of bFGF. The crystal structure of the bound hexasaccharide is green with sulfate groups in red and represented as ball-and-sticks, the AutoDock minimum energy structure is yellow, and the DOCK minimum energy structure is magenta. The protein is represented as a ribbon and the residues of the high binding preference region (res. 28, 120, 121, 126, 135, 136) are white, and those of the low preference binding region (res. 27, 102, 104, 134) are pink.

Table 4. Interaction Probabilities for aFGF (hexasaccharide runs)[a]

	residue number											
run/prog	N18	L111	K112	K113	K118	R122	G126	Q127	K128	A129	E_ref	E_min
AutoDock												
global	0.40	0.04	0.48	**0.60**	0.42	0.46	0.10	0.30	0.25	0.15	-52.9	−81.8
local	0.50	0.02	**0.86**	0.80	0.80	0.85	0.12	0.46	0.28	0.24	-52.6	−68.0
DOCK												
global	0.26		**0.94**	0.57	0.46	0.61		0.15	0.15	0.12	-63.2	−96.5
local	0.78	0.01	0.56	**0.86**	0.15	0.29	0.06	0.44	0.70	0.44	-64.5	−99.6

[a] Only residues with atoms within a 4 Å cutoff of the bound saccharide in the experimental structures[18] are listed. E_ref = interaction energy of the crystal structure, E_min = minimum interaction energy after docking; energies are given in kcal/mol and were calculated with use of a 0.2 Å spacing grid; the highest probabilities for each run are shown in bold.

Table 5. Interaction Probabilities for the 1st Binding Site of bFGF (hexasaccharide runs)[a]

	residue number													
run/prog	K27	N28	N102	T104	K120	R121	K126	K130	G134	Q135	K136	A137	E_ref	E_min
AutoDock														
global	0.15	0.30	0.05	0.08	0.66	**0.70**	0.45	0.62	0.10	0.33	0.31	0.11	−85.9	−87.7
local	0.39	0.51		0.01	0.54	**0.94**	0.76	0.65	0.30	0.65	0.26	0.26	−85.9	−103.3
DOCK														
global	0.61	0.25	0.28	0.43	0.21	0.50	0.19	0.18	0.13	0.15	**0.66**	0.06	−67.5	−93.1
local	0.92	0.22	0.02	0.49		0.71	0.05	0.02	0.38	0.11	**0.99**	0.01	−66.7	−83.9

[a] Only the residues with atoms within a 4 Å cutoff of the bound saccharides in the experimental structures (res. 27, 28, 102, 104, 120, 121, 126, 135, 136, 137)[17] and (res. 130, 134)[39] are listed. E_ref = interaction energy of the crystal structure, E_min = minimum interaction energy after docking; energies are given in kcal/mol and were calculated with use of a 0.2 Å spacing grid; the highest probabilities per probe are shown in bold.

IDS302 and SGN303 very close to the crystal positions (deviations ca. 0.6 Å). One more sulfate group (O6 of SGN304) is oriented toward the protein surface, forming a hydrogen bond with Lys27. This changes the orientation of the following residue, so that the last two residues point away from the protein surface, without forming any further hydrogen bonds and deviating from the crystal structure position. In the crystal structure SGN306, rather than SGN304, makes a hydrogen bond to Lys27 but the position of SGN306 (and thus of 304 and 305) is also influenced by crystal contacts to Arg34 and His36 of a symmetry related molecule. In contrast to the good agreement of the AutoDock structure (RMSD = 4.1 Å), the minimum energy structure from the DOCK run (colored pink in Figure 6b) shows a completely different orientation. The whole chain is rotated about −140° so that residue 306 of the DOCK structure comes close to residue 301 of the crystal structure. This inversion of the chain ordering produces an RMSD of 14.2 Å between the crystal structure and the minimum energy structure.

Assessment of Docking Protocols. In all of the docking runs, the crystal structure does not represent the most favorable conformation with respect to the force fields of the docking programs (Tables 4 and 5). In two cases, the global docking

runs reveal better interaction energies than the local ones, which might reflect incomplete sampling during the docking runs. Given the highly flexible nature and large size of the ligands (41 rotatable bonds in the case of the hexasaccharide of bFGF and 42 in the aFGF case), the sampling in both variants, the local and the global runs, might be insufficient for detecting the global energy minimum. This is also indicated by the energy distribution found in the docking results. For the AutoDock local run with bFGF, the minimum energy structure has a total interaction energy of −103.3 kcal/mol, while the second ranked structure has an energy of −89.1 kcal/mol, and is followed by a dense energy distribution for all higher ranked conformations. For the global docking run, the lowest energy structure has an interaction energy of −87.7 kcal/mol, and is followed by a dense energy distribution for all higher conformations. A significant increase of the number of docking trials is prohibitive given that a 10-fold increase in sampling would require a corresponding increase in computation time, raising the calculation time from a day (ca. 30 CPU hours on a SGI Origin 2000 for a local docking of the hexasaccharide) to weeks for each docking run.

One way of solving the sampling problem might be the reduction of the number of rotatable bonds, but the choice of which bonds to rigidify is not necessarily obvious. The glycosidic bonds would seem to be good candidates for fixing, but this would inhibit larger structural changes of the hexasaccharide during the docking runs. Thus although bonds could be rigidified when experimental structures are known, we preferred not to introduce *a priori* bias by doing this.

The dense energy distribution for the docked structures, despite considerable structural variation, means that multiple binding modes must be considered in analyzing the results. This may reflect the physical nature of the binding site or it may point to the need to improve the energy scoring functions used. One possibility would be to improve the electrostatic model by computing electrostatic binding free energies by using a continuum dielectric model.[41,42] However, this is considerably more demanding computationally than computation of the scoring functions used here and further parametrization and calibration would be necessary. Computation of binding energies might also benefit from explicit consideration of water molecules which could mediate protein−GAG interactions and play an important role in determining the orientation and conformation of the bound hexasaccharides. For example, water molecules have been observed to mediate sugar binding in arabinose-binding protein[43] and the maltodextrin-binding protein.[44] Analysis of the available protein−GAG complexes revealed several water molecules that participate in hydrogen bond networks between the protein and the saccharide, but none of these positions seems to be well conserved. We therefore did not consider individual water molecules and their effects on binding in our docking protocols. However, improvement of the treatment of solvent and solvation effects in the docking protocols and scoring functions is desirable and should improve the docking results.

Analysis of the tables (Tables 2 and 4 for aFGF and Tables 3 and 5 for bFGF) also allows comparison between the two docking strategies. In both global docking runs, with disaccharides and with hexasaccharides, the same residues were identified as possible interaction partners. So what are the advantages of one or the other docking strategy? A global search with disaccharides is fast, 2 h of computational time, compared with the 50 h needed for a global search with a hexasaccharide. So the combination of a global disaccharide run with a local hexasaccharide run saves computational time, improves the sampling in important binding regions, and allows the usage of multiple ligands, representing different features of the large probe, in the global search.

Docking Runs with IL-8. Having applied the docking protocols to protein−heparin complexes of known structure, we searched for the heparin binding region of IL-8. Because IL-8 shows some structural variability, both the crystal and the NMR structures were chosen as target structures for all docking runs. Two different regions of IL-8 were suggested for the interaction with the receptor and GAG.[24,45−48] The receptor binding site should include the N-terminal E−L−R motif (residues 4 to 6) and a hydrophobic patch, both of which are required for receptor binding. The location of the heparin binding site has been suggested to be at the C-terminal α -helix by a truncation study.[24] The global docking runs (Table 6) suggest two possible interaction regions for sulfated sugars. The first one is located near the N-terminus and includes the residues 5, 6, 11, 12, 13, 36, and 49, while the second one includes residues of the C-terminal α-helix (residues 64 and 68) and the proximal loop region (residue 18). Both regions include a number of basic residues and are spatially separated from each other (Figure 7a). The C-terminal region corresponds to the region identified by Kuschert and co-workers in their NMR shifting experiment.[49] They used various sulfated disaccharides for the determination of heparin binding residues of IL-8. Their measurement of binding constants identified two trisulfated probes (with the same substituent patterns as disaccharide probes **I** and **II** (Figure 2)) as having the strongest interactions. A preference for binding of these disaccharides could not be detected in our docking runs (data not shown).

Considering the available biological and spectroscopic data about the heparin binding site of IL-8, we only performed local docking runs with hexasaccharide for the second binding site (C-terminal α-helix and proximal loop region). The results of the docking runs are shown in Table 7. His18, Lys20, Lys64, Lys67, and Arg68 are identified as the most important interaction partners—all of these residues have interaction probabilities higher than 70%. Two different orientations were observed for the docked structures: The first orientation contains a binding mode in which the heparin axis is orientated perpendicular to the α -helical axis (Figure 7b). The bound hexasaccharide interacts with the proximal loop region (His18 and Lys20) and the C-terminal helix. All of the low-energy structures of the AutoDock and DOCK runs show this orientation. In the second mode, the heparin axis is orientated parallel to the helix axis (Figure 7c), but this mode was only observed in some of the high-energy structures.

The first binding mode is similar to the heparin binding mode of the PF4 dimer,[16,50] in which the heparin axis shows a perpendicular orientation and the chain bridges the gap between the two monomers. The second orientation corresponds to the model suggested by Spillmann *et al.*,[25] in which the heparin

(41) Jackson, R. M.; Sternberg, M. J. E. *J. Mol. Biol.* **1995**, *20*, 258−275.
(42) Olson, M. A.; Cuff, L. *J. Mol. Recognit.* **1997**, *10*, 277−289.
(43) Vyas, N. K.; Vyas, M. N.; Quiocho, F. A. *J. Biol. Chem.* **1991**, *266*, 5226−5237.
(44) Spurlino, J. C.; Lu, G. Y.; Quiocho, F. A. *J. Biol. Chem.* **1991**, *266*, 5202−5219.
(45) Hebert, C. A.; Vitangcol, R. V.; Baker, J. B. *J. Biol. Chem.* **1991**, *266*, 18989−18994.
(46) Clark-Lewis, I.; Schumacher, C.; Baggiolini, M.; Moser, B. *J. Biol. Chem.* **1991**, *266*, 23128−23134.
(47) Clubb, R. T.; Omichinski, J. G.; Clore, G. M.; Gronenborn, A. M. *FEBS Lett.* **1994**, *338*, 93−97.
(48) Williams, G.; Borkakoti, N.; Bottomley, G. A.; Cowan, I.; Fallowfield, A. G.; Jones, P. S.; Kirtland, S. J.; Price, G. J.; Price, L. *J. Biol. Chem.* **1996**, *271*, 9579−9586.

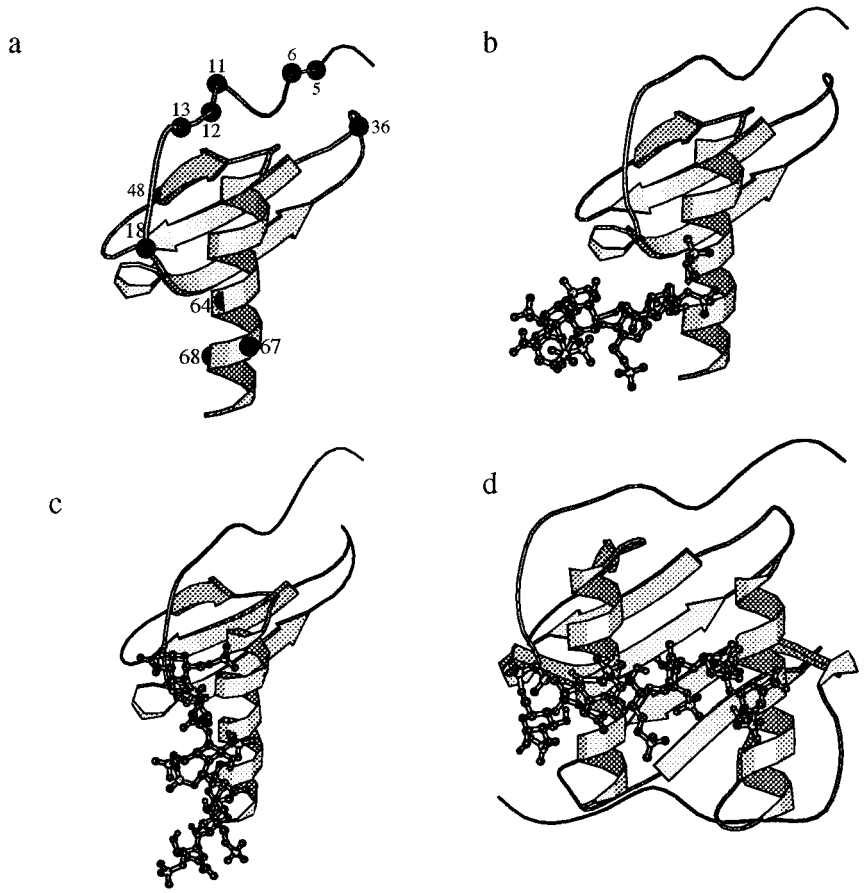

Figure 7. Heparin binding sites of IL-8: (a) the two potential binding sites of IL-8 identified in the mono- and disaccharide docking runs—first site (res. L5, R6, K11, T13, N36, E48), second site (res. H18, K64, K67, R68); (b and c) two orientations of the docked hexasaccharide, (b) is the minimum energy structure from the AutoDock run, while (c) represents a high energy structure from the same run; and (d) minimum energy conformation from the AutoDock run with the IL-8 dimer. The NMR structure of IL-8 was used for all plots.

Table 6. Interaction Probabilities for IL-8 (mono- and disaccharide runs)[a]

GRID										
monomers:	L5(0.52),	**Q8**(0.64),	K11(0.21),	K23*(0.21),	I28(0.50),	H33(0.50),	E48(0.18),	L49(0.18),	C50(0.20),	R68*(0.24)
dimers:	I10(0.22),	**K11**(0.72),	T12(0.50),	Y13(0.54),	S14(0.41),	K15(0.28),	R47(0.49),	E48(0.46),	L49(0.54),	C50(0.38)
glp:	L5(0.27),	R6(0.26),	C7(0.43),	I10(0.28),	**K11**(0.48),	T12(0.43),	Y13(0.30),	P32(0.26),	H33(0.27),	N36(0.45)
idp:	K3(0.21),	E4(0.21),	C9(0.24),	I10(0.31),	**K11**(0.56),	T12(0.47),	Y13(0.51),	S14(0.33),	K15(0.23),	D52(0.25)
AutoDock										
glp:	P16(0.15),	F17(0.18),	H18*(0.38),	P19*(0.16),	K20*(0.33),	F21*(0.18),	R60*(0.20),	**K64***(0.42),	K67*(0.18),	R68*(0.33)
idp:	Y13(0.15),	K15(0.18),	P16(0.11),	F17(0.12),	H18*(0.19),	K20*(0.17),	K23*(0.17),	R60*(0.16),	**K64***(0.24),	R68*(0.15)
DOCK										
glp:	L5(0.27),	R6(0.30),	C7(0.25),	C9(0.25),	**I10**(0.34),	K11(0.29),	E48(0.25),	C50(0.25),	K64*(0.28),	R68*(0.25)
idp:	**L5**(0.38),	R6(0.34),	C7(0.23),	I10(0.31),	K11(0.26),	H33(0.35),	E48(0.21),	C50(0.25),	K64*(0.26),	R68*(0.25)

[a] The 10 residues with the highest number of contacts for each probe and program used are shown, with the normalized number of contacts given in parentheses. Maximum standard deviation is 0.40 for GRID, 0.01 for AutoDock, and 0.07 for the DOCK results. For each probe, the residue with the highest probability is shown in bold. Residues identified as GAG binding in the NMR shifting study by Kuschert *et al.*[49] are marked with asterisks.

Table 7. Interaction Probabilities of IL-8 (hexasaccharide run)[a]

	residue number	E_min
X-ray		
AutoDock	F17(0.46), **H18***(0.90), K20*(0.73), K64*(0.70), K67*(0.55)	−51.5
DOCK	H18*(0.78), **K20***(0.92), K64*(0.81), K67*(0.84), R68*(0.83)	−42.9
NMR		
AutoDock	H18*(0.78), K20*(0.67), R60*(0.52), **K64***(0.81), R68*(0.74)	−74.2
DOCK	H18*(0.72), K20*(0.55), **K64***(0.98), K67*(0.76), R68*(0.89)	−62.8

[a] Only the 5 residues with the highest probabilities are reported; the residues with the highest probability are shown in bold. The normalized number of contacts is given in parentheses. E_min = minimum interaction energy after docking; energies are given in kcal/mol and were calculated with use of a 0.2 Å spacing grid. Residues identified as GAG binding in the NMR shifting study by Kuschert *et al.*[49] are marked with asterisks.

chain binds in a horseshoe fashion. This binding mode increases the required length for binding up to an 18-mer, so that a loop between helices can be formed, although the actual interaction site can be as small as a pentasaccharide.

Having the results for the IL-8 monomer, we also performed docking runs with the IL-8 dimer to answer the question whether both binding modes could be detected in the IL-8 dimer. All the docked structures from the AutoDock and DOCK runs show an orientation similar to that shown in Figure 7d. The docked hexasaccharide has a perpendicular orientation and bridges the gap between the two antiparallel α-helices. Because of the length of the hexasaccharide, only interactions with one of the proximal loop regions were observed. Interactions with the other proximal loop caused a shift of all GAG residues by two residues, which would increase the minimum length for bridging the distance between both proximal loops to eight or nine residues, which corresponds to the minimum length for binding of a heparin chain under low ionic strength in the study of Spillmann *et al.*[25]

Conclusion

In the current work, we tested the reliability of fully automated docking protocols for the detection of heparin binding sites and applied these protocols for the prediction of the binding mode of heparin to IL-8. Our findings may be summarized as follows:

(1) All three docking programs (GRID, AutoDock, and DOCK) were able to correctly localize the heparin binding sites on our test case proteins (aFGF, bFGF, and Antithrombin) with sulfated mono- and disaccharides as probes. The suggested interacting residues are in good agreement with the experimentally determined interactions.

(2) The combination of a global search for binding sites with sulfated mono- and disaccharides with a subsequent local docking of a hexameric GAG appears to be the most promising docking strategy. This protocol reduces the required computational time compared to a global docking of the hexasaccharide. It also improves sampling in important binding regions and allows the use of different probes in the global disaccharide docking stage, which improves the reliability of binding site identification. This docking procedure can be realized with either DOCK or AutoDOCK alone or one of these programs in combination with the GRID program. The use of more than one program is recommended as no single program produced dramatically better results than the others although quite different results were obtained with the different programs in some cases, e.g., global docking to Antithrombin.

(3) In the case of bFGF, a strong preference for the first binding site was observed and only a few structures were localized at the second binding site. Considering the hexasaccharide runs in the aFGF and bFGF cases, our results are in good agreement with the experimental results for the heparin binding sites of FGFs. The interaction site of bFGF seems to be more specific with regard to the location of the sulfate groups, while more structural variation may occur in the case of aFGF.

(4) Despite the structural differences between the crystal and the NMR structures of IL-8, a consistent set of residues was identified as possible interaction partners for heparin in the di- and hexasaccharide runs. This set includes the residues His18, Lys20, Lys64, Lys67, and Arg68. Two possible orientations were observed for a hexasaccharide docked to the IL-8 monomer—perpendicular and parallel to the helical axis—with the perpendicular orientation representing the low-energy structures. In the case of the IL-8 dimer, all docked structures show a perpendicular orientation in which the hexasaccharide bridges the gap between the α-helices.

Acknowledgment. We thank Dr. Peter J. Goodford and Dr. Todd Ewing for helpful discussion concerning GRID and DOCK, respectively, and Dr. Waksman and Dr. Herr for their bFGF-trisaccharide structure. This project was supported in part by the Supercomputing Resource for Molecular Biology at EMBL, Heidelberg (an EU Large Scale Facility, contract No. ER BCHGE CT940062).

JA983319G

(49) Kuschert, G. S. V.; Hoogewerf, A. J.; Proudfoot, A. E. I.; Chung, C.; Cooke, R. M.; Hubbard, R. E.; Wells, T. N. C.; Sanderson, P. N. *Biochemistry* **1998**, *37*, 11193−11201.

(50) Mayo, K. H.; Ilyina, E.; Roongta, V.; Dundas, M.; Joseph, J.; Lai, C. K.; Maione, T.; Daly, T. J. *Biochem. J.* **1995**, *312*, 357−365.