School of Biomedical Sciences

Molecular Modelling of the Interactions of Complex Carbohydrates with Proteins

Neha Sureshchandra Gandhi

This thesis is presented for the degree of

Doctor of Philosophy

of

Curtin University

September 2011

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature: …….

Date: ……...

ABSTRACT

Glycosaminoglycans (GAGs) are ubiquitous complex carbohydrate molecules present on the cell surfaces and in extracellular matrices (ECM) of vertebrate and invertebrate tissues. The interactions of sulphated GAGs such as heparin and heparan sulphate (HS) with numerous immunologically-relevant proteins is now attracting considerable interest as a source of new therapeutics for the treatment of infectious diseases, inflammation and allergies, and cancers. Various computational molecular modelling methods are being employed to determine the nature of the interactions of heparin oligosaccharides with various proteins in order to establish the structural requirements that determine their binding specificity and selectivity.

The first part of this research focused on understanding the inflammatory cytokine CXCL-8 (Interleukin-8 or IL-8) and its interactions with GAGs. A variety of molecular modelling methods were used to investigate the binding of complex carbohydrates to this protein. A number of consensus heparin/HS binding motifs were predicted to be required for the binding of monomeric or oligomeric structures of CXCL-8. Bioinformatics analyses showed that the basic residues in the heparin binding site are highly variable within the CXC family and amongst various CXCL-8 species. Drug-like carbohydrate mimetic molecules (cyclitols) that bind optimally to CXCL-8 were identified and characterised. It was established that both an optimum number of sulphates and a certain length of alkyl spacers are required for the interaction of cyclitol inhibitors with the dimeric form of CXCL-8. Furthermore, explicit solvent molecular dynamics simulations of dimeric CXCL-8 showed how its two anti-parallel helices exhibit domain movements that can bring them in closer proximity. In addition, these simulations revealed shearing movements in the C-terminal helices in the CXCL-8 dimer. This inherent flexibility of the CXCL-8 dimer can be exploited in drug design as it plays an important role in the understanding of the interactions of molecules such as sulphated cyclitols with the two monomers.

Structural bioinformatics and molecular modelling methods were used to generate and analyse a three-dimensional model of heparanase, an enzyme involved in metastasis and angiogenesis in cancer, in order to gain insight into its protein sequence, and its structural and functional relationships. The interactions of heparanase with disaccharide substrates and GAG mimetics were modelled to investigate the structural determinants of their protein binding specificity and selectivity. The choice of structural template for modelling the binding site of heparanase is very critical. Analyses of active-site similarity across groups of homologous template structures revealed that these bound oligosaccharides can block ligand

binding to the catalytic and heparin binding sites of heparanase. Ligand-protein docking simulations further revealed the existence of a large binding site extending at least two saccharide units beyond the cleavage site (towards the non-reducing end) and at least three saccharides towards the reducing end (towards heparin-binding site 2). Extensive modelling of substrate and inhibitor interactions with the catalytically-active glutamic acids and the two binding sites for heparan sulphate of heparanase provides information useful for future drug discovery efforts focused on the identification of novel inhibitors of this enzyme.

Free energy calculations of the binding of sGAGs to GAG-binding proteins were pioneered with the proteins PECAM-1 (platelet endothelial cell adhesion molecule) and Annexin, using the molecular mechanics/Poisson Boltzmann surface area (MM/PBSA) method. Analysis of the free energy of binding components revealed that the major contributors to complex stability are electrostatic interactions, with equally important contributions from van der Waals interactions and vibrational entropy changes, against a large, unfavourable desolvation penalty due to the high charge density of sGAG. The calculated absolute free energies of binding of the molecules investigated were found to be inaccurate compared to experimental values, but the method performed well in discriminating weak and strong binding.

A final focus of this research was to investigate the conformational properties of sulphated iduronic acid (IdoA2S), a hexopyranose present in heparin. IdoA2S adopts more than one conformation (skew boat and chair) when internally positioned within an oligosaccharide or when it interacts with proteins. The influence of the solvent on the flexibility and conformations of this saccharide is of significant interest given that its biomolecular interactions occur in an aqueous environment. Therefore, molecular dynamics simulations were used to investigate the ability of the GROMOS force field and the GLYCAM carbohydrate parameter set in the presence of explicit solvent to successfully predict rotamer populations for this ring system. Calculations of theoretical proton NMR coupling constants showed that the GROMOS96 force field can predict the skew-boat to chair conformational ratio in good agreement with experiment; however, the accuracy of the GLYCAM force field in representing ring conformer populations during unconstrained molecular dynamics simulations is still debatable.

ACKNOWLEDGEMENTS

I am overwhelmed by the tremendous support given by the people who made this thesis possible.

I would like to express my sincere appreciation to my supervisor Prof. Ricardo Mancera, who has supported and encouraged me throughout this project. He was always there to meet, listen and advice about my ideas and to ask me good questions to help me throughout the course of this research work. His patience in reading draft after draft of every paper (torture at times!), proposals and ideas I wrote up continues to amaze me. I am very grateful to him for introducing me to the field of glycosaminoglycans and for his interest in my future career. His mentorship has been invaluable throughout the research. It is difficult to overstate my gratitude to him.

I would like to take this opportunity to thank Prof. Linda Kristjanson in her former role of Deputy Vice-Chancellor- Research and Development for her confidence and willingness to help me achieve my goals in my Ph.D. as well as for her generosity in helping to support my future postdoctoral position. Prof. Erik Helmerhorst deserves special thanks as the chairperson of my thesis committee and as an advisor. I would also like to thank Mrs. Sandra Spears and Dr. Geraldine Pinto in the School of Biomedical Sciences for their friendship during this study.

I am grateful for the award of an Endeavour International Postgraduate Research Studentship (EIPRS) and a Curtin University Postgraduate Scholarship (CUPS) that made it possible for me to do my PhD research. I also gratefully acknowledge the Western Australian Interactive Virtual Environments Centre (IVEC) and the National Computational Infrastructure (NCI) facility for access to high-performance computing.

My most sincere thanks go to my mentors, the Lohray's - Dr. Braj Bhushan and Dr. Vidya. I thank them for introducing me to the wonders and frustrations of drug discovery research and giving me opportunities in the field of Bioinformatics and Molecular Modelling. I thank Dr. Vidya for teaching me important lessons in project management, leadership and communication during my tenure with them and for all her moral support *via* email. I also value the support from my colleagues at Zydus Research Centre who helped me to learn in the fields of biotechnology and medicinal chemistry.

I am indebted to my parents Suresh and Saroj Gandhi for their unconditional love and for inculcating in me the dedication and discipline to do whatever I undertake. All my life my

father has been my inspiration and my mummy has been a special and best friend who is always there to listen patiently and help me fulfill my aspirations. On a more personal note, I would like to thank my brother Samir, my sister-in-law Meghna and my cousin Prashant, who support and encourage me on all of my endeavours, and gave me strength to move forward to something better in life. Without the support and constant encouragement of my friends this work would not have been possible. I owe a debt of gratitude to my friend and colleague Jestin Mandumpal, who taught me so much and whom I miss every day. I wish to thank my best friends from school (Manisha and Shabana) for helping me to get through the difficult times, and for all the emotional support and caring they provided. Grateful acknowledgement goes to my friends Pradeep Shukla, Miguel Renteria, Daniel Hernandez, Mario Lausberg Leena Sharma and Deepa Sharma, who never give up in giving their support and care to me in all aspects of life and trusting my capabilities. Special thanks go to Michael Connaughton for getting me interested in good food and cars, and for defending me against all odds. I also appreciate the timely help from my colleagues Zak Hughes, Zhili Zuo and Cara Kreck, as well as present and past members of the group. I also acknowledge my house owners Denise and Louise at Walpole St. for making my stay in Bentley memorable. I was pleased to have housemates like Shamila, Priyanka, Elvis and Mahmood, who were a constant source of entertainment. My genuine appreciation goes to my teachers from school who laid a solid foundation of techniques, knowledge, discipline and respect. My friends on networking sites also deserve acknowledgement for keeping in touch to date.

List of publications included as part of this thesis

REFFERED PUBLICATIONS

Neha S. Gandhi and Ricardo L. Mancera, The structure of glycosaminoglycans and their interactions with proteins. *Chemical Biology and Drug Design,* 2008, vol. 72, pp. 455-482. Front cover of the issue, most downloaded and cited paper of 2009.

Neha S. Gandhi and Ricardo L. Mancera, Free energy calculations of glycosaminoglycan-protein interactions. *Glycobiology*, 2009, vol. 19, pp. 1103-1115.

Neha S. Gandhi and Ricardo L. Mancera, Can current force fields reproduce ring puckering in 2-*O*-sulfo-α-L-iduronic acid? A molecular dynamics simulation study. *Carbohydrate Research,* 2010, vol. 345, pp. 689-695.

Neha S. Gandhi and Ricardo L. Mancera, Heparin/heparan sulphate-based drugs. *Drug Discovery Today*, 2010, vol. 15, pp. 1058-1069.

Neha S. Gandhi and Ricardo L. Mancera, Molecular dynamics simulations of CXCL-8 and its interactions with heparin fragments and sulphated linked cyclitols. *Journal of Chemical Informatics and Modeling*, 2011, vol. 51, pp. 335-358.

Neha S. Gandhi, Craig Freeman, Christopher Parish and Ricardo L. Mancera, Computational analyses of the catalytic and heparin binding sites and their interactions with glycosaminoglycans in glycoside hydrolase family 79 endo-β-D-glucuronidase (Heparanase). *Glycobiology*, 2011, in press.

I warrant that I have obtained, where necessary, permission from the copyright owners to use any third party copyright material reproduced in the thesis (e.g. questionnaires, artwork, unpublished letters), or to use any of my own published work (e.g. journal articles) in which the copyright is held by another party (e.g. publisher, co-author).

Statement of Contribution of Co-Authors

To Whom It May Concern

I, Neha S. Gandhi, conceived, developed, implemented and carried out all computational analyses to the paper/publication entitled below:

Neha S. Gandhi and Ricardo L. Mancera, The structure of glycosaminoglycans and their interactions with proteins. *Chemical Biology and Drug Design,* 2008, vol. 72, pp. 455-482. Front cover of the issue, most downloaded and cited paper of 2009.

Neha S. Gandhi and Ricardo L. Mancera, Free energy calculations of glycosaminoglycan-protein interactions. *Glycobiology*, 2009, vol. 19, pp. 1103-1115.

Neha S. Gandhi and Ricardo L. Mancera, Can current force fields reproduce ring puckering in 2-*O*-sulfo-α-L-iduronic acid? A molecular dynamics simulation study. *Carbohydrate Research,* 2010, vol. 345, pp. 689-695.

Neha S. Gandhi and Ricardo L. Mancera, Heparin/heparan sulphate-based drugs. *Drug Discovery Today*, 2010, vol. 15, pp. 1058-1069.

Neha S. Gandhi and Ricardo L. Mancera, Molecular dynamics simulations of CXCL-8 and its interactions with heparin fragments and sulphated linked cyclitols. *Journal of Chemical Informatics and Modeling*, 2011, vol. 51, pp. 335-358.

The author conducted the research, analysed the data, implemented the research plan and wrote the papers in consultation with Prof. Ricardo L. Mancera.

(Signature of Candidate)

I, as a Co-Author, endorse that this level of contribution by the candidate indicated above is appropriate.

Prof. Ricardo L. Mancera (Signature of Co-Author)

Statement of Contribution of Co-Authors

To Whom It May Concern

I, Neha S. Gandhi, conceived, developed, implemented and carried out all computational analyses to the paper/publication entitled below:

Neha S. Gandhi, Craig Freeman, Christopher Parish and Ricardo L. Mancera, Computational analyses of the catalytic and heparin binding sites and their interactions with glycosaminoglycans in glycoside hydrolase family 79 endo-β-D-glucuronidase (Heparanase). *Glycobiology*, 2011, in press.

The author[1] analysed data, guided the research plan and wrote the papers in consultation with Prof. Ricardo L. Mancera. Craig Freeman and Christopher Parish suggested this work and provided advice.

(Signature of Candidate)

I, as a Co-Author[1], endorse that this level of contribution by the candidate indicated above is appropriate.

Prof. Ricardo L. Mancera (Signature of Co-Author[1])

We, as Co-Authors[2,3], endorse that this level of contribution by the candidate indicated above is appropriate.

Dr. Craig Freeman (Signature of Co-Author[2])

Prof. Christopher Parish (Signature of Co-Author[3])

List of additional publications published during my PhD

REFEREED PUBLICATIONS

Zhili Zuo, **Neha S. Gandhi** and Ricardo L. Mancera, Calculations of the free energy of interaction of the cJun-cFos coiled-coil: Effect of the solvation model and the inclusion of polarisation effects. *J Chem Inf Model*, 2010, 50, 2201-2212.

Miguel E. Renteria, **Neha S. Gandhi**, Pablo Vinuesa, Erik Helmerhorst and Ricardo L. Mancera, A comparative structural bioinformatics analysis of the insulin receptor family ecto-domain based on phylogenetic information. *PLoS ONE*, 2008, e3667.

Neha S. Gandhi and Ricardo L. Mancera, Computational approaches for the prediction of the structure and interaction of coiled-coil peptides. *Curr Bioinform*, 2008, 3, 149-161.

Neha S. Gandhi, Kathy Young, John Warmington and Ricardo L. Mancera, Characterisation of sequence and structural features of the Candida krusei enolase. *In Silico Biol*, 2008, 8, 0034.

ORAL PRESENTATIONS

Neha S. Gandhi, Molecular modelling of polyanionic carbohydrates and their interactions with proteins, 2010, *MolSim Winter School*, Curtin University of Technology, Perth, Australia.

Neha S. Gandhi, Molecular dynamics simulations of CXCL-8 chemokine and sulphated GAG mimetics, 2009, *6th international Conference on Proteoglycan*, Aix-les-Bains, France.

Neha S. Gandhi, Molecular modelling of the interactions of complex carbohydrates with proteins, 2009, *WABRI symposia*, Curtin University of Technology, Perth, Australia.

POSTER PRESENTATIONS

Neha S. Gandhi and Ricardo L. Mancera, Free energy calculations of glycosaminoglycan-protein interactions, 2009, *6th international Conference on Proteoglycan*, Aix-les-Bains, France.

Neha S. Gandhi, Craig Freeman, Christopher R. Parish and Ricardo L. Mancera, Molecular dynamics simulations of CXCL-8 and its interactions with sulphated linked cyclitols, 2009, *6th international Conference on Proteoglycan*, Aix-les-Bains, France.

Neha S. Gandhi and Ricardo L. Mancera, Molecular modelling of the interactions of glycosaminoglycans with PECAM-1/CD31, 2009, *Combined Biological Sciences Meeting*, Perth, Australia.

Neha S. Gandhi and Ricardo L. Mancera, Interactions of complex carbohydrates with proteins. A molecular modelling approach, 2008, *ComBio*, Canberra, Australia.

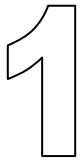
I warrant that I have obtained, where necessary, permission from the copyright owners to use any third party copyright material reproduced in the thesis (e.g. questionnaires, artwork, unpublished letters), or to use any of my own published work (e.g. journal articles) in which the copyright is held by another party (e.g. publisher, co-author).

Contents

ABBREVIATIONS

2-deoxy-2-sulfamido-α-D-glucopyranosyl	GlcNS
2-deoxy-2-sulfamido-α-D-glucopyranosyl-6-O-sulfate	GlcNS6S
2-O-sulfo-α-L-iduronic acid	IdoA2S
Alveolar macrophage chemotactic factor-1	AMCF-1
Alzheimer's disease	AD
Anti-thrombin	AT
AT-binding domain	ABD
Atrial fibrillation	AF
Beta-amyloid peptide	Aβ
Biomolecular ligand energy evaluation protocol	BLEEP
Brain natriuretic peptide	BNP
Chondroitin sulphate	CS
Deep vein thrombosis	DVT
Dermatan sulphate	DS
Duffy antigen/receptor for chemokines	DARC
Endothelial cell neutrophil-activating peptide	ENAP
Factor Xa	Fxa
Fibroblast growth factor	FGF
Generalised Born	GB
Glycosaminoglycans	GAGs
GLYcosidic bonds Genetic Algorithm	GLYGAL
Granulocyte chemotactic peptide	GCP
Granulocyte macrophage colony stimulating factor	GM-CSF
GROningen MOlecular Simulation	GROMOS
Heparan sulphate	HS
Hyaluronic acid	HA
Interleukin-8	IL-8
Keratan sulphate	KS
Lamarckian genetic algorithm	LGA
Leukocyte adhesion inhibitor	LAI
Linear interaction energy	LIE
LInear Notation for Unique description of Carbohydrate Sequences	LINUCS
Low molecular weight heparin	LMWH
Lung carcinoma-derived chemotaxin	LUCT
Lymphocyte-derived neutrophil-activating peptide	LYNAP
Macrophage inflammatory protein 1 alpha	MIP-1α
Molecular dynamics	MD
Monocyte-derived neutrophil chemotactic factor	MDNCF

Monocyte-derived neutrophil-activating peptide	MONAP
Monte carlo multiple minima	MCMM
Multiple sequence alignments	MSA
Neutrophil-activating factor	NAF
Neutrophil-activating protein 1	NAP-1
Normal modes analysis	NMA
Nuclear magnetic resonance	NMR
Nuclear Overhauser enhancement	NOE
Particle Mesh Ewald	PME
Platelet endothelial cell adhesion molecule 1	PECAM-1
Platelet factor 4	PF4
Poisson Boltzmann	PB
Restrained ElectroStatic Potential fit	RESP
Saturation transfer difference	STD
Scaling of Electrostatic Interactions	SEI
solvent accessible surface area	SASA
Stromal cell-derived factor-1	SDF-1
T-cell chemotactic factor	TCF
Thrombin-binding domain	TBD
time-averaged restrained molecular dynamics	TARMD
Transferable intermolecular potential 3 point	TIP3P
Unfractionated heparin	UFH
van der Waals	VDW
Vascular endothelial growth factor	VEGF
Venous thromboembolism	VTE
α-L-iduronic acid	IdoA
β-L-glucuronic acid	GlcA



1 Introduction and Overview

1.1 Motivation, Background, Aims and Scope of the Present Work

A large number of carbohydrate-based derivatives are used as therapeutics or in diagnostics. Sulphated GAGs (sGAGs) such as "heparin-like agonists" have important uses, such as anti-coagulants. Current research efforts are aimed at designing potent and selective antagonists that can bind to the heparan sulphate (HS)-binding site of a protein and block HS–protein interactions. Several methods have been developed to study sGAG-protein interactions, including affinity chromatography, analytical ultracentrifugation, electrophoretic mobility shift assays, competition experiments, mass spectrometry-based approaches, isothermal titration calorimetry and surface plasmon resonance. To date, only a few molecular modelling methods have been specifically developed to investigate HS-protein complexes.

In this thesis various recent advances in the field of structural bioinformatics and molecular modelling and simulation of sGAGs are described. Conformational studies of sulphated iduronic acid-containing GAGs in the presence of water shed light on the ability of these methods to describe GAG structures. Structural bioinformatics, molecular modelling and molecular dynamics simulations were used to characterise in detail the interactions of heparin fragments and their mimetics with proteins of therapeutic importance: the cytokine CXCL-8, the enzyme heparanase, the cell adhesion molecule PECAM-1 and the calcium-dependent phospholipid-binding protein Annexin A2. The roles and importance of CXCL-8 and heparanase and the properties of saccharides are briefly described below, while a more detailed discussion is provided in the corresponding chapters in this thesis. The findings of the molecular modelling studies reported are aimed at providing a rationale for the future development of new selective, potent drug-like carbohydrate-mimic molecules that can target CXCL-8 and heparanase and which may be useful for the treatment of inflammatory diseases and angiogenesis.

CXCL-8 is a potent neutrophil chemotactic and activating factor and is a primary inflammatory cytokine produced by many cells in response to proinflammatory stimuli [1].

Its function is partly to attract neutrophils to the site of inflammation and to activate them. In humans, the CXCL-8 protein is encoded by the IL-8 (interleukin-8) gene mapped to a region on chromosome 4q which includes several other family members (α-chemokines) [2-5]. The CXCL-8 cDNA sequence is a protein of 99 amino acids (Uniprot: P10145) that includes a signal peptide of 22 residues [6] (Figure 1).

CXCL-8 was initially described as a monocyte-derived neutrophil chemotactic factor (MDNCF) [6]. It has also been referred to as neutrophil-activating factor (NAF) [7], neutrophil-activating protein 1 (NAP-1) [8, 9], granulocyte chemotactic peptide (GCP) [10], leukocyte adhesion inhibitor (LAI) [11], 3-10C protein [12], lung carcinoma-derived chemotaxin (LUCT) [13, 14], lymphocyte-derived neutrophil-activating peptide (LYNAP) [15], alveolar macrophage chemotactic factor-1 (AMCF-1) [16, 17], monocyte-derived neutrophil-activating peptide (MONAP) [18], endothelial cell neutrophil-activating peptide (ENAP-beta) [19], K60 [20], T-cell chemotactic factor (TCF) [21], to name a few. CXCL-8 (1-77) is referred to as Interleukin-8 (IL-8) and it is differentially processed into four forms encompassing 69, 72, 77, and 79 amino acids. These variants include IL-8 (5-77), IL-8 (6-77), IL-8 (7-77), IL-8 (8-77) and IL-8 (9-77). The 77 amino acid variant (1-77) is predominant in endothelial cells, whilst the 72 amino acid variant (5-77) is found in other leukocytes [22]. IL-8 (6-77) has 5-10-fold higher neutrophil activation, IL-8 (5-77) has increased activity for neutrophil activation and IL-8 (7-77) has a higher affinity for receptors CXCR1 and CXCR2 compared to IL-8 (1-77) [9, 22].

CXCL-8 is a member of the neutrophil-specific CXC (the two N-terminal cysteines are separated by one amino acid "X") subfamily of chemokines [23] and hence shares the characteristic chemokine secondary structure comprising a NH_2-terminal loop including conserved disulphide bridges, three anti-parallel β-strands connected by loops and a COOH-terminal α-helix (Figure 1). CXCL-8 can form non-covalent dimers in solution, especially at high concentrations, and the quaternary structure consists of a six-stranded anti-parallel β-sheet and two anti-parallel helices lying across the β-sheet.

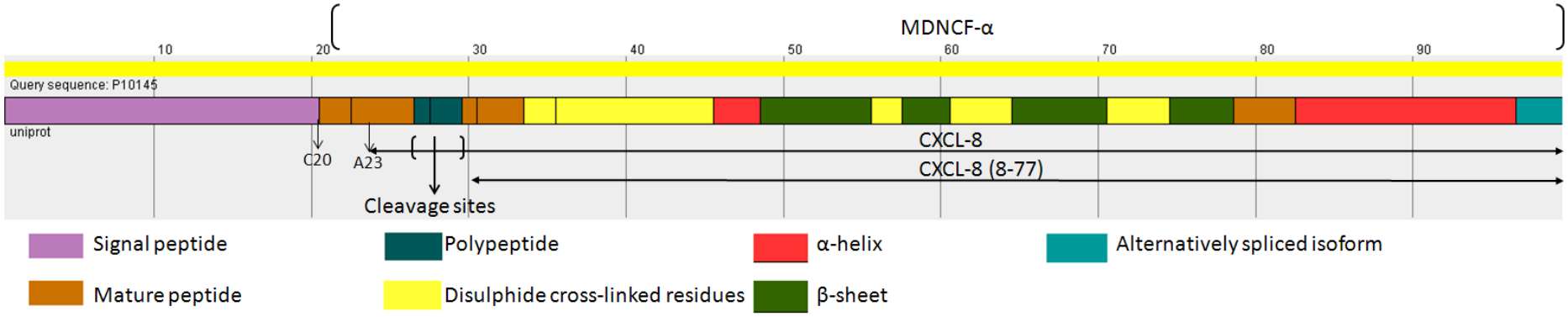


Figure 1. Schematic representation of the primary structure of CXCL-8. The CXCL-8 cDNA sequence is a protein of 99 amino acids (Uniprot: P10145). Removal of the 22-residue signal peptide generates a mature protein of 77 amino acids. Further proteolysis of the N-terminal end leads to a variant form with 72 amino acids. Full activation of CXCL-8 may require cleavage to the 72 amino acid form. Secondary structure representation on the primary structure for CXCL-8 shows disulphide bond connections represented by yellow. The green blocks represent the first, second and third β-strands, respectively. The red block represents the 3_{10}-helix present before the first β-strand and the C-terminal α-helix.

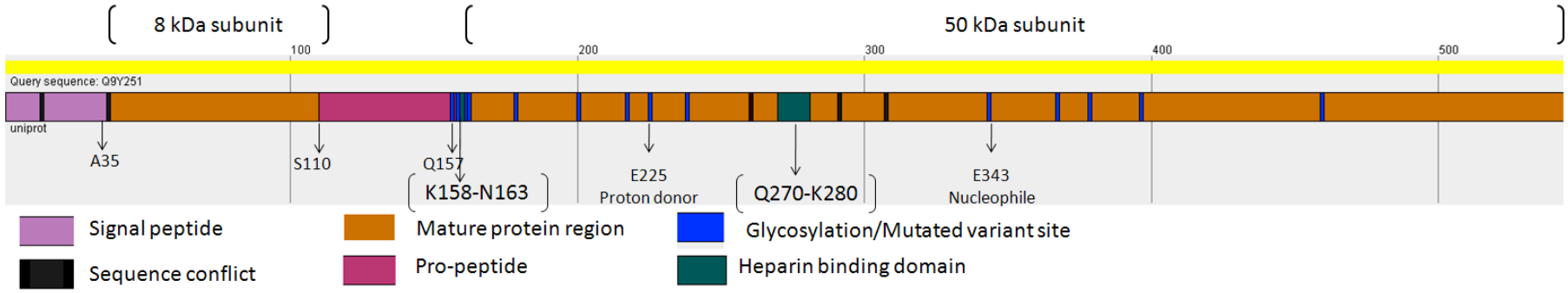


Figure 2. Schematic representation of the primary structure of human heparanase. Signal peptidase cleavage of the pre pro-enzyme (Met1-Ala35) generates the 65 kDa latent form (pro-heparanase). The latent form of heparanase is schematically represented as a contiguous fusion of 8 (Gln36-Glu109) and 50 (Lys158-Ile543) kDa subunits connected by an intervening linker region (Ser110-Gln157). The mature and active forms of the enzyme (heterodimer) is subsequently produced by proteolysis of the pro-enzyme and excision of the intervening 6 kDa linker region (S110–Q157). The active residues (Glu225 and Glu343) required for HS cleavage are both contained within the 50 kDa subunit. The heparin-binding domains 1 (HBD-1, residues Lys158-Asn162) and 2 (HBD-2, residues Gln270-Lys280) are in close proximity to the catalytic residues, thus making it an important target for the development of heparanase inhibitors.

While CXCL-8 is one of the major mediators of the inflammatory response, increased heparanase activity contributes to metastasis and angiogenesis. The human heparanase gene is located on chromosome 4 at band 4q21.3 [24]. The cDNA encodes a pre-propolypeptide of the 543-amino acid long heparanase [25-27] (HPSE, HPA1, HEP, HPR1, HPSE1) with 6 potential N-glycosylation sites (Figure 2). Heparanase is synthesised as a latent enzyme (pro-heparanase) of 65 kDa where the NH_2 terminal peptide (Met1-Ala35) directs its secretion. Proteolytic cleavage removes a 6 kDa linker segment (Ser110- Ser110-Gln157), yielding an 8 kDa at the amino terminus (Gln36-Glu109) and a non-covalently linked 50 kDa catalytic domain (Lys158-Ile543) that heterodimerises to form the active enzyme [28].

Heparanase is an endo-β-D-glucuronidase that catalyses the hydrolytic cleavage of the β-1,4-glycosidic bond between a D-glucuronic acid and a D-glucosamine in heparin/HS. Pronounced correlation between heparanase gene expression and different stages of tumour progression in humans has been established in clinical studies [29, 30]. Overexpression of heparanase has been observed in tumours of the head and neck [31], pancreatic tumours [32], hepatocellular carcinoma [33], colon cancer [34], cancer of the bladder [35], esophageal carcinoma [36] and cultured human tumour cell lines and tissues [37, 38]. Besides the well-documented catalytic activity of the enzyme, heparanase has also been reported to have non-enzymatic functions [39], including enhanced cell adhesion mediated by activation of signalling molecules such as Akt, Src, EGFR, and Rac in HS-dependent and -independent manners, cross-talk with matrix metalloproteinases, and induction of p38 and Src phosphorylation associated with vascular endothelial growth factor (VEGF) and tissue factor gene induction.

The human heparanase gene is located in chromosome 4 at band 4q21.3, and its cDNA is composed of 14 exons and 13 introns [24]. The heparanase gene is alternatively spliced, producing two isoforms: HPSE 1a (5 kb) and HPSE 1b (1.7 kb) [24, 40]. The purification of heparanase and the characterisation of its activity have been reported by various groups [40-45], suggesting that the heparanase *cDNA* encodes a polypeptide of 543 amino acids that

appears as a ~65 kDa protein. HPSE acts *via* a retaining reaction mechanism that involves general acid/base catalysis by Glu225 and nucleophilic catalysis by Glu343 (Figure 2). The crystal structure of heparanase has not been resolved yet and hence molecular modelling of the interactions of amino residues Glu225 and Glu343 and heparin/HS binding domains [46] with the substrates is the only way to try to perform the rational design of inhibitors of angiogenesis and metastasis.

The key objectives of performing the molecular modelling of GAG binding proteins CXCL-8 and heparanase are summarised below:

- Predict heparin/HS binding sites for CXCL-8 and heparanase based upon their secondary structures and a survey of sulphate binding motifs from public databases.
- Build a homology model of heparanase.
- Characterise the structural determinants of heparin-protein binding specificity and selectivity using well established molecular modelling methods, focusing on the chemokine CXCL-8 and the HS-cleaving enzyme heparanase.
- Predict the interactions and binding affinities of heparin/HS fragments with CXCL-8 and heparanase.
- Validate these binding affinity estimates using sophisticated free energy methods like MM/PBSA.
- Characterise the conformations in aqueous solution of various carbohydrate/heparin fragments using molecular dynamics simulations.

Monosaccharides can be coupled to more than one monosaccharide through glycosidic bonds to form chained and branched structures. The glycosidic linkage is a covalent bond between the hydroxyl of the anomeric carbon of a monosaccharide and a hydroxyl group of another saccharide. Glycosidic linkages can be either α or β based on the relative stereochemistry of the anomeric position and the stereochemical centre furthest from C1 in the saccharide. Glycoside hydrolases (or glycosidases) are enzymes that break glycosidic bonds. For

example, heparanase acts on β-glycosidic bonds. Glycosidic linkages accounts for most of the flexibility of oligosaccharide structures. The conformation assumed by a glycosidic linkage can be expressed by means of the torsion angles defined by its rotatable bonds. Most glycosidic linkages have two torsions, namely phi (φ) and psi (ψ). In the present work, torsions are measured using the IUPAC definition (Figure 3), which defines the angles of the (1→X) linkage as φ = O5-C1-O-C'$_X$ and ψ = C1-O-C'$_X$-C'$_{X+1}$.

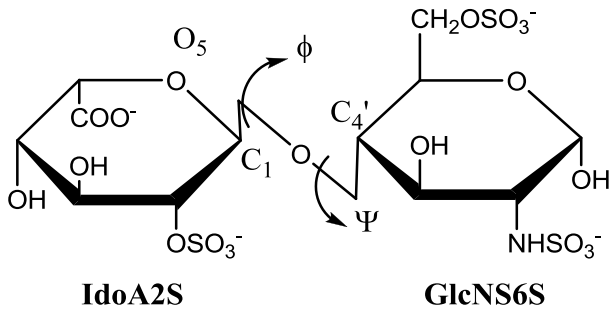


Figure 3. Saccharides are joined through a glycosidic bond to form polysaccharides. The conformation of the torsional angles around the glycosidic linkage imparts flexibility to the polysaccharide. The orientations of these torsions, typically phi (ϕ) and psi (ψ), fall within an expected range of values.

Aqueous solvation effects play an important role in determining the conformation of GAGs. The hydroxyl groups within the sugar ring have a tendency to orient themselves in chains due to hydrogen bonding following a clockwise or anti-clockwise arrangement. These conformations are generally energetically optimal, although the aqueous solvent may favour other stable configurations involving hydrogen bonds with water or due to entropic effects [47]. A generalised notation to represent these conformations is used wherein the approximate conformation of the ring is indicated with an italic capital letter, which designates the ring shape, and numerals, which distinguish between the variant forms of each shape. For example, $^1C_4$ denotes a chair conformation in which the C1 lies above the plane and the C4 below the plane of pyranose ring. Alternatively, these conformations are represented through the Cremer-Pople ring puckering formalism [48] by means of using the

reduced Cartesian coordinates q, θ and φ_2. The θ and φ_2 parameters reveal the relative orientations of the puckering about the ring and the magnitude of puckering (the mean deviation of the atoms from the plane) is given by q. The parameters θ and φ_2 are angles in the range $0° \leq \theta \leq 180°$ and $0° \leq \varphi_2 \leq 360°$ respectively.

Hexopyranoses have 38 basic conformations: 2 chairs, 6 boats, 6 skew-boats, 12 half-chairs and 12 envelopes [49]. Canonical chair structures $^4C_1$ and $^1C_4$ for hexopyranose rings correspond to values of θ = 0° and 180°, respectively, and hence are at the north and south poles, respectively, of the Cremer-Pople sphere while the half-chair and envelope are positioned between the poles and the equator (Figure 4). In solution, there is a dynamic interconversion between the equatorial forms, which is referred to as pseudorotation. Therefore, at the equator of the Cremer-Pople sphere, where θ = 90°, there exist six boat ($^{3,0}B$, $B_{1,4}$, $^{2,5}B$, $B^{3,0}$, $^{1,4}B$ and $B_{2,5}$) and six skew-boat or twist-boat ring conformers ($^3S_1$, $^5S_1$, ^2S_o, $^lS_3$, $^1S_5$ and $^oS_1$) with different φ_2 values. For IdoA2S, the ^2S_o twist boat has a φ_2 value of 150°. The last aim of this research is to consider the conformational equilibrium of sulphated α-L-iduronic acid (IdoA2S), which is the major uronic acid component of heparin. Given their inherent flexibility, the conformational analysis of IdoA2S is more complicated than comparable studies with other pyranosides, as the interconversion between different ring conformations takes place on the microsecond time scale.

Together with ring puckering and glycosidic linkages, *J-coupling* (dipolar coupling) is another important observable effect for all solution structures accessible to a molecule at the temperature of the measurement and averaged on the NMR timescale. *J-coupling* is the coupling between two nuclear spins due to the influence of bonding electrons on the magnetic field running between the two nuclei. It contains information about dihedral angles, which can be estimated using the Karplus equation [50]. The *J-coupling* constants computed from snapshots collected from MD trajectories are indicative of average ring torsional motion corresponding to $^4C_1$, $^2S_0$, and $^1C_4$ geometries. It should be noted that if motion occurs close to a Karplus maximum or minimum, deviation from ideality and increased

monosaccharide ring flexibility result in decreased magnitude of the observed coupling, whereas if the Karplus equation is close to zero, deviation and motion generally result in an increase in coupling magnitude. Comparion of experimental *J-coupling constants* with those obtained from MD simulations also allows parameterisation of force fields.

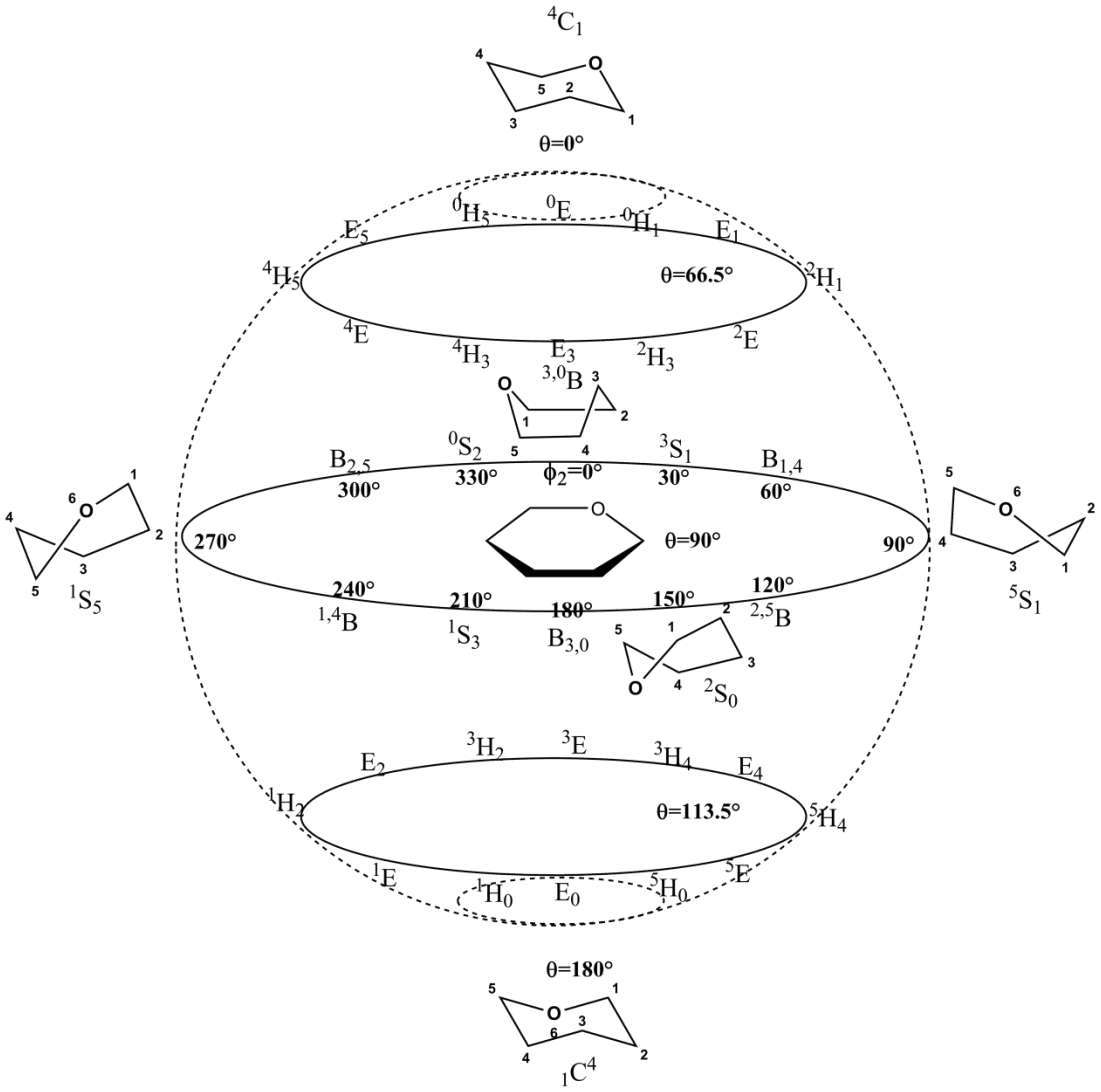


Figure 4. Ring puckering parameters for the chair, boat, and twist-boat forms of the hexopyranose ring as defined by Jeffrey and Yates [51] according to Cremer and Pople [48]. The parameters θ and ϕ_2 are angles in the range 0°≤ θ≤180° and 0°≤ϕ_2 ≤360°, where the chair conformations ($^4C_1$ and $^1C_4$) are found at opposite poles (θ = 0° and 180°, respectively) and boat and skew-boat forms are found around the equator (θ = 90°). Interconversion between the equatorial forms is referred to as pseudorotation.

1.2 Organisation of this thesis

This thesis has been organised as a compilation of an exegesis and several research papers published during the doctoral research period that describe work aimed at addressing the above-defined objectives. Each chapter outlines its own objectives and has well defined abstract, methods, results and discussion and conclusion sections. This format makes it easier to read any chapter with minimal referral to previous chapters. The following chapters are organised in accordance to the objectives defined above.

Chapter 2: This chapter is a comprehensive compilation of literature providing historical and general insights on the structure of GAGs, GAG-binding proteins and their clinical significance. The review paper focuses on understanding the complexity of GAG–protein interactions and mechanisms using classical examples such as growth factors, anti-thrombin, cytokines and cell adhesion molecules. Furthermore, the nature of GAG-protein interactions and the experimental and computational approaches that are used to address them is discussed.

Chapter 3: This chapter provides a broad review of the field of heparin-based therapeutics with a focus on the most important, recent advances in the field of heparin-mimetic oligosaccharides and other small molecules in pre-clinical development and clinical trials for treating cancer, inflammation and infections.

Chapter 4: This chapter describes explicit solvent molecular dynamics simulations that capture and contrast the the motions of terminal helices in the X-ray and NMR structures of CXCL-8. The structures of CXCL-8 are analysed in detail, including the prediction of its likely heparin/HS binding regions. This chapter further describes molecular docking simulations used to model the interactions of GAGs and its mimetics (cyclitols) with CXCL-8. In addition, the calculations of relative free energies of binding of cyclitols with CXCL-8 using MM-GB/PBSA simulations are described.

Chapter 5: This chapter describes the construction of a three-dimensional model of heparanase using homology modelling and threading methods. The model of heparanase is analysed in detail, including the prediction of its likely heparin/HS binding regions. Molecular docking simulations used to investigate the interactions of substrates and inhibitors (GAGs and their derivatives) with heparanase are also described. This publication is currently in press.

Chapter 6: This chapter reports the use and effectiveness of free energies calculation methods using continuum solvation methods to estimate the binding affinities of GAGs with representative protein molecules. It has been found experimentally and through molecular docking studies that PECAM-1 exhibits different binding affinities for heparin oligosaccharides of different size, sulphation and ring conformations. To provide mechanistic insight into these GAG–protein interactions, molecular dynamics simulations using the MM/PB(GB) SA approach. Similar investigation was carried out for Annexin A2, which is another protein known to bind heparin with high affinity and in a calcium-dependent manner. Molecular dynamics simulations were also used to analyse the fluctuations in glycosidic linkages of bound and free GAGs in solution.

Chapter 7: This chapter describes the investigation of the accuracy of commonly used force fields to reproduce the torsional potential parameters (Cremer-Pople) that determine the ring conformations of monosaccharides (IdoA2S) in solution by comparing the predictions of explicit solvent molecular dynamics simulations with known experimental data. The GLYCAM and GROMOS force fields are studied given their widespread use to simulate hexopyranose-based carbohydrates.

Chapter 8: This chapter summarises the implications and limitations derived from the various studies described in previous chapters and provides a discussion of the likely direction of future work in this field.

1.3 References

1. Baggiolini, M. and I. Clark-Lewis, *Interleukin-8, a chemotactic and inflammatory cytokine.* FEBS Lett., 1992. **307**(1): p. 97-101.
2. Modi, W.S. and Z.Q. Chen, *Localization of the human CXC chemokine subfamily on the long arm of chromosome 4 using radiation hybrids.* Genomics, 1998. **47**(1): p. 136-139.
3. Modi, W.S., et al., *Monocyte-derived neutrophil chemotactic factor (MDNCF/IL-8) resides in a gene cluster along with several other members of the platelet factor 4 gene superfamily.* Hum. Genet., 1990. **84**(2): p. 185-187.
4. O'Donovan, N., M. Galvin, and J.G. Morgan, *Physical mapping of the CXC chemokine locus on human chromosome 4.* Cytogenet. Cell Genet., 1999. **84**(1-2): p. 39-42.
5. Modi, W.S., et al., *Chromosome mapping and RFLP analyses of monocyte-derived neutrophil chemotactic factor (MDNCF/IL-8).* Cytogenet. Cell Genet., 1989. **51**: p. 1046.
6. Matsushima, K., et al., *Molecular cloning of a human monocyte-derived neutrophil chemotactic factor (MDNCF) and the induction of MDNCF mRNA by interleukin 1 and tumor necrosis factor.* J. Exp. Med., 1988. **167**(6): p. 1883-1893.
7. Walz, A., et al., *Purification and amino acid sequencing of NAF, a novel neutrophil-activating factor produced by monocytes.* Biochem. Biophys. Res. Commun., 1987. **149**(2): p. 755-761.
8. Schroder, J.M., et al., *Purification and partial biochemical characterization of a human monocyte-derived, neutrophil-activating peptide that lacks interleukin 1 activity.* J. Immunol., 1987. **139**(10): p. 3474-3483.
9. Baggiolini, M., A. Walz, and S.L. Kunkel, *Neutrophil-activating peptide-1/interleukin 8, a novel cytokine that activates neutrophils.* J. Clin. Invest., 1989. **84**(4): p. 1045-1049.
10. Van Damme, J., et al., *Purification of granulocyte chemotactic peptide/interleukin-8 reveals N-terminal sequence heterogeneity similar to that of beta-thromboglobulin.* Eur. J. Biochem., 1989. **181**(2): p. 337-344.
11. Gimbrone, M.A., Jr., et al., *Endothelial interleukin-8: a novel inhibitor of leukocyte-endothelial interactions.* Science, 1989. **246**(4937): p. 1601-1603.
12. Schmid, J. and C. Weissmann, *Induction of mRNA for a serine protease and a beta-thromboglobulin-like protein in mitogen-stimulated human leukocytes.* J. Immunol., 1987. **139**(1): p. 250-256.
13. Hotta, K., et al., *Coding region structure of interleukin-8 gene of human lung giant cell carcinoma LU65C cells that produce LUCT/interleukin-8: homogeneity in interleukin-8 genes.* Immunol. Lett., 1990. **24**(3): p. 165-169.
14. Suzuki, K., et al., *Purification and partial primary sequence of a chemotactic protein for polymorphonuclear leukocytes derived from human lung giant cell carcinoma LU65C cells.* J. Exp. Med., 1989. **169**(6): p. 1895-1901.
15. Gregory, H., et al., *Structure determination of a human lymphocyte derived neutrophil activating peptide (LYNAP).* Biochem. Biophys. Res. Commun., 1988. **151**(2): p. 883-890.
16. Goodman, R.B., et al., *Identification of two neutrophil chemotactic peptides produced by porcine alveolar macrophages.* J. Biol. Chem., 1991. **266**(13): p. 8455-8463.
17. Goodman, R.B., et al., *Molecular cloning of porcine alveolar macrophage-derived neutrophil chemotactic factors I and II; identification of porcine IL-8 and another intercrine-alpha protein.* Biochemistry, 1992. **31**(43): p. 10483-10490.

18. Schroder, J.M., U. Mrowietz, and E. Christophers, *Identification of different charged species of a human monocyte derived neutrophil activating peptide (MONAP).* Biochem. Biophys. Res. Commun., 1988. **152**(1): p. 277-284.
19. Schroder, J.M. and E. Christophers, *Secretion of novel and homologous neutrophil-activating peptides by LPS-stimulated human endothelial cells.* J. Immunol., 1989. **142**(1): p. 244-251.
20. Sick, C., et al., *Novel chicken CXC and CC chemokines.* Cytokine, 2000. **12**(3): p. 181-186.
21. Larsen, C.G., et al., *The neutrophil-activating protein (NAP-1) is also chemotactic for T lymphocytes.* Science, 1989. **243**(4897): p. 1464-1466.
22. Hebert, C.A., et al., *Endothelial and leukocyte forms of IL-8. Conversion by thrombin and interactions with neutrophils.* J. Immunol., 1990. **145**(9): p. 3033-3040.
23. *Chemokine/chemokine receptor nomenclature*, in *Cytokine*. 2003, IUIS/WHO Subcommittee on Chemokine Nomenclature. p. 48-49.
24. Dong, J., et al., *Genomic organization and chromosome localization of the newly identified human heparanase gene.* Gene, 2000. **253**(2): p. 171-178.
25. Hulett, M.D., et al., *Cloning of mammalian heparanase, an important enzyme in tumor invasion and metastasis.* Nat. Med., 1999. **5**(7): p. 803-809.
26. Toyoshima, M. and M. Nakajima, *Human heparanase. Purification, characterization, cloning, and expression.* J. Biol. Chem., 1999. **274**(34): p. 24153-24160.
27. Vlodavsky, I., et al., *Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis.* Nat. Med., 1999. **5**(7): p. 793-802.
28. Fairbanks, M.B., et al., *Processing of the human heparanase precursor and evidence that the active enzyme is a heterodimer.* J. Biol. Chem., 1999. **274**(42): p. 29587-29590.
29. Barash, U., et al., *Proteoglycans in health and disease: new concepts for heparanase function in tumor progression and metastasis.* FEBS J., 2010. **277**(19): p. 3890-3903.
30. Vlodavsky, I., et al., *Heparanase: one molecule with multiple functions in cancer progression.* Connect. Tissue Res., 2008. **49**(3): p. 207-210.
31. Simizu, S., et al., *Expression of heparanase in human tumor cell lines and human head and neck tumors.* Cancer Lett., 2003. **193**(1): p. 83-89.
32. Koliopanos, A., et al., *Heparanase expression in primary and metastatic pancreatic cancer.* Cancer Res., 2001. **61**(12): p. 4655-4659.
33. El-Assal, O.N., et al., *The clinicopathological significance of heparanase and basic fibroblast growth factor expressions in hepatocellular carcinoma.* Clin. Cancer Res., 2001. **7**(5): p. 1299-1305.
34. Friedmann, Y., et al., *Expression of heparanase in normal, dysplastic, and neoplastic human colonic mucosa and stroma. Evidence for its role in colonic tumorigenesis.* Am. J. Pathol., 2000. **157**(4): p. 1167-1175.
35. Gohji, K., et al., *Heparanase protein and gene expression in bladder cancer.* J. Urol., 2001. **166**(4): p. 1286-1290.
36. Mikami, S., et al., *Loss of syndecan-1 and increased expression of heparanase in invasive esophageal carcinomas.* Jpn. J. Cancer Res.. 2001. **92**(10): p. 1062-1073.
37. Hong, X., et al., *Increased chemotactic migration and growth in heparanase-overexpressing human U251n glioma cells.* J. Exp. Clin. Cancer Res., 2008. **27**: p. 23.
38. Mikami, S., et al., *Expression of heparanase in renal cell carcinomas: implications for tumor invasion and prognosis.* Clin. Cancer Res., 2008. **14**(19): p. 6055-6061.
39. Levy-Adam, F., N. Ilan, and I. Vlodavsky, *Tumorigenic and adhesive properties of heparanase.* Semin. Cancer Biol.. 2010. **20**(3): p. 153-160.

40. Toyoshima, M. and M. Nakajima, *Human heparanase. Purification, characterization, cloning, and expression.* J Biol Chem, 1999. **274**(34): p. 24153-24160.
41. Freeman, C. and C.R. Parish, *Human platelet heparanase: purification, characterization and catalytic activity.* Biochem. J., 1998. **330** (Pt 3): p. 1341-1350.
42. Gonzalez-Stawinski, G.V., et al., *Partial sequence of human platelet heparitinase and evidence of its ability to polymerize.* Biochim. Biophys. Acta, 1999. **1429**(2): p. 431-438.
43. Goshen, R., et al., *Purification and characterization of placental heparanase and its expression by cultured cytotrophoblasts.* Mol. Hum. Reprod., 1996. **2**(9): p. 679-684.
44. Kussie, P.H., et al., *Cloning and functional expression of a human heparanase gene.* Biochem. Biophys. Res. Commun., 1999. **261**(1): p. 183-187.
45. Vlodavsky, I., et al., *Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis.* Nat Med, 1999. **5**(7): p. 793-802.
46. Levy-Adam, F., et al., *Identification and characterization of heparin/heparan sulfate binding domains of the endoglycosidase heparanase.* J. Biol. Chem., 2005. **280**(21): p. 20457-20466.
47. Rao, V.S.R., et al., *Conformation of carbohydrates*. 1998: Harwood Academic Publishers. 409.
48. Cremer, D. and J.A. Pople, *General definition of ring puckering coordinates.* J. Am. Chem. Soc., 1975. **97**(6): p. 1354-1358.
49. Stoddart, J.F., *Stereochemistry of carbohydrates*. 1971: Wiley-Interscience. 249.
50. Karplus, M., *Vicinal proton coupling in nuclear magnetic resonance.* J. Am. Chem. Soc., 1963. **85**(18): p. 2870-2871.
51. Jeffrey, G.A. and J.H. Yates, *Stereographic representation of the cremer-pople ring-puckering parameters for pyranoid rings.* Carbohydr. Res., 1979. **74**(1): p. 319-322.

2

2 The Structure of Glycosaminoglycans and their Interactions with Proteins

Chem Biol Drug Des 2008; **72:** 455–482

Review Article

© 2008 The Authors
Journal compilation © 2008 Blackwell Munksgaard

doi: 10.1111/j.1747-0285.2008.00741.x

The Structure of Glycosaminoglycans and their Interactions with Proteins

Neha S. Gandhi[1,2] and Ricardo L. Mancera[1,2,3,*]

[1]Western Australian Biomedical Research Institute, Curtin University of Technology, GPO Box U1987, Perth, WA 6845, Australia
[2]School of Biomedical Sciences, Curtin University of Technology, GPO Box U1987, Perth, WA 6845, Australia
[3]School of Pharmacy, Curtin University of Technology, GPO Box U1987, Perth, WA 6845, Australia
**Corresponding author: Ricardo L. Mancera*
r.mancera@curtin.edu.au

Glycosaminoglycans (GAGs) are important complex carbohydrates that participate in many biological processes through the regulation of their various protein partners. Biochemical, structural biology and molecular modelling approaches have assisted in understanding the molecular basis of such interactions, creating an opportunity to capitalize on the large structural diversity of GAGs in the discovery of new drugs. The complexity of GAG–protein interactions is in part due to the conformational flexibility and underlying sulphation patterns of GAGs, the role of metal ions and the effect of pH on the affinity of binding. Current understanding of the structure of GAGs and their interactions with proteins is here reviewed: the basic structures and functions of GAGs and their proteoglycans, their clinical significance, the three-dimensional features of GAGs, their interactions with proteins and the molecular modelling of heparin binding sites and GAG–protein interactions. This review focuses on some key aspects of GAG structure–function relationships using classical examples that illustrate the specificity of GAG–protein interactions, such as growth factors, anti-thrombin, cytokines and cell adhesion molecules. New approaches to the development of GAG mimetics as possible new glycotherapeutics are also briefly covered.

Key words: carbohydrates, chondroitin sulphate, dermatan sulphate, glycosaminoglycans, heparan sulphate, heparin, hyaluronic acid, keratan sulphate, proteoglycans

Received 17 September 2008, revised and accepted for publication 19 October 2008

Carbohydrates can exist as simple sugars and as complex conjugates known as glycans. Glycans mediate a wide variety of events in cell–cell and cell–matrix interactions that are crucial to the development and function of complex multicellular organisms. Glycomic technologies for exploring the structure of complex sugar molecules have emerged in the past two decades, opening up a new frontier which has been called 'glycobiology' (1). This review provides an introduction to the structural properties of the linear chain glycans called glycosaminoglycans (GAGs) and their interactions with proteins.

Basic Features and Functions of GAGs

Glycosaminoglycans are large complex carbohydrate molecules that interact with a wide range of proteins involved in physiological and pathological processes (2,3). Glycosaminoglycans are sometimes known as mucopolysaccharides because of their viscous, lubricating properties, as found in mucous secretions. These molecules are present on all animal cell surfaces in the extracellular matrix (ECM), and some are known to bind and regulate a number of distinct proteins, including chemokines, cytokines, growth factors, morphogens, enzymes and adhesion molecules (2,4). The key properties of GAGs are summarized in Table 1.

Glycosaminoglycans in aqueous solution are surrounded by a shell of water molecules, which makes them occupy an enormous hydrodynamic volume in solution (5). When a solution of GAGs is compressed, the water is squeezed out and the GAGs are forced to occupy a smaller volume. When the compression is removed, GAGs regain their original hydrated volume because of the repulsion arising from their negative charges (5).

Classification of GAGs

Glycosaminoglycans are linear, sulphated, negatively charged polysaccharides that have molecular weights of roughly 10–100 kDa. There are two main types of GAGs. Non-sulphated GAGs include hyaluronic acid (HA), whereas sulphated GAGs include chondroitin sulphate (CS), dermatan sulphate (DS), keratan sulphate (KS), heparin and heparan sulphate (HS). Glycosaminoglycans chains are composed of disaccharide repeating units called disaccharide repeating regions (Table 2). The repeating units are composed of uronic acid (D-glucoronic acid or L-iduronic acid) and amino sugar (D-galactosamine or D-glucosamine). Hence, GAGs differ according to the type of hexosamine, hexose or hexuronic acid unit that they contain, as well as the geometry of the glycosidic linkage between these units. Chondroitin sulphate and DS, which contain galactosamine, are called galactosaminoglycans, whereas heparin and HS, which contain glucosamine, are called glucosaminoglycans. The amino sugar may be sulphated on carbons 4 or

Table 1: Key properties of glycosaminoglycans

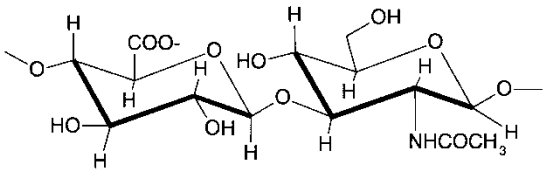
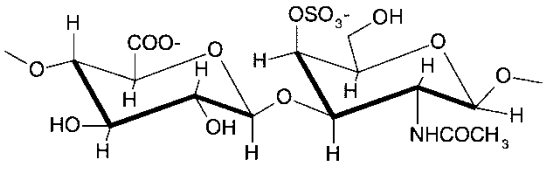
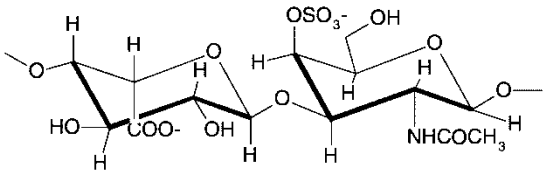
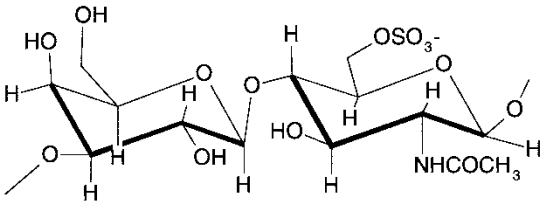
Physico-chemical properties of GAGs: negatively charged, viscous, lubricating, unbranched polysaccharides, contain repeating disaccharide units, bind large amounts of water, low compressibility.
Classification of GAGs: chondrotin sulphates, keratan sulphate, dermatan sulphate, hyaluronan, heparin and heparan sulphate.
Function of GAGs: cell adhesion, cell growth and differentiation, cell signalling anticoagulation.

6 or on the non-acetylated nitrogen; however, the sugar backbone of GAGs can be sulphated at various positions. As a result, a simple octasaccharide can have over 1 000 000 different sulphation sequences (6). Glycosaminoglycans also vary in the geometry of the glycosidic linkage (α or β). At physiological pH, all carboxylic acid and sulphate groups are deprotonated, giving GAGs very high negative charge densities (heparin has the highest negative charge density of any known biomolecule) (7).

Nomenclature of GAG fragments

The names of the monosaccharides present in GAGs are frequently abbreviated. The most common are three-letter abbreviations for simple monosaccharides (e.g. Gal for galactose, Glc for glucose, Xyl for xylose, Man for mannose). Most of the monosaccharides are assumed to be in the D-configuration, except for iduronic acid (IdoA). All monosaccharides are assumed to be in the pyranose (*p*) form (six-membered ring), while all glycosidic linkages are assumed to originate from the anomeric C1 hydroxyl group. These monosac-

Table 2: Repeating disaccharide units of various glycosaminoglycans

Glycosaminoglycan	Disaccharide units	Features
Hyaluronic acid	 D-GlcA- $\beta(1\rightarrow 4)$-D-GlcNAc-$\alpha(1\rightarrow 4)$	Molecular weight 4–8000 kDa Non-sulphated non-covalently attached to proteins in the ECM; also found in bacteria Usually found in synovial fluid, vitreous humour, ECM of loose connective tissue Excellent lubricators and shock absorbers
Chondroitin sulphate	 D-GlcA-$\beta(1\rightarrow 3)$-D-GalNAc4S-$\beta(1\rightarrow 4)$	Molecular weight 5–50 kDa Most abundant GAG in the body Found in cartilage, tendon, ligament, aorta Bind to proteins (like collagen) to form proteoglycan aggregates
Dermatan sulphate	 L-IdoA-$\alpha(1\rightarrow 3)$-D-GalNAc4S-$\beta(1\rightarrow 4)$	Molecular weight 15–40 kDa Found in skin, blood vessels, heart valves
Keratan sulphates I and II	 D-Gal-$\beta(1\rightarrow 4)$-D-GalNAc6S-$\beta(1\rightarrow 3)$	Molecular weight 4–19 kDa Most heterogeneous GAG KS I is found in the cornea KS II is found in cartilage aggregated with CS

charides are further classified based on the substitution position and the substituent (such as *O*-sulphate). If sulphate is attached to the C2 carbon of iduronic acid, it is referred to as IdoA2S. Similarly, sulphation at positions 2 and 6 in *N*-acetylglucosamine can be written as GlcNS6S. The glycosidic linkage between the monosaccharides is in either α or β configuration, involving the anomeric hydroxyl of one monosaccharide and any available hydroxyl group in a second monosaccharide. For example, $\alpha(1 \rightarrow 4)$ refers to the α linkage in a disaccharide between the anomeric carbon of the first monosaccharide and the hydroxyl at position 4 of the other monosaccharide. The name begins at the non-reducing end (which is usually highly sulphated) proceeding towards the reducing end (the sugar with the free anomeric carbon that can be oxidized). The GAG nomenclature system (8) is based on Roman numerals and acronyms. The simple heparin pentasaccharide with $\Delta^{4,5}$-unsaturated uronic acid at the non-reducing end can be written in this system as Δ UA2S(1 $\rightarrow$ 4)GlcNS6S(1 $\rightarrow$ 4)IdoA2S(1 $\rightarrow$ 4)GlcNS6S(1 $\rightarrow$ 4)IdoA2S.

Recently, a disaccharide structural code, a new shorthand nomenclature, has been introduced for designating the disaccharide subunit structure of all GAGs (9). Disaccharide structural code assigns each GAG disaccharide a four-character code. The first character assigns the stereochemistry of the uronic acid: U, G, I or D for an undesignated uronic acid, glucuronic acid, iduronic acid or $\Delta^{4,5}$-unsaturated uronic acid, respectively, or g for galactose. The second character is used to define the location of sulphate groups: 0, 2, 3 or 6, representing no sulphate, sulphate at C2 or C3 of uronic acid, or sulphate at C6 of galactose, respectively. The third character designates the type of hexosamine (upper case for glucosamine and lower case for galactosamine) and the N substituent (H, A, S and R for free amine, acetate, sulphate or some other substituent, respectively). The fourth character in the code identifies the pattern of sulphation on the hexosamine (0 for no sulphate, 3 for sulphate at C3, 6 for sulphate at C6, 9 for disulphated glucosamine at C3 and C6, and 10 for disulphated galactosamine at C4 and C6). The pentasaccharide described above can be described as D2-S6I2-S6I2, with all of the isomeric information specified in the code.

A linear, canonical description of carbohydrates, Linear Notation for Unique description of Carbohydrate Sequences (LINUCS) (10), has been implemented[a] to describe carbohydrate structures according to the IUPAC nomenclature (8), which includes GAGs. LinearCode (11) is another approach developed that allows the compact and unambiguous description of complex structures. It uses a simple one- to two-letter representation of monosaccharide units and linkages and a description of branches using a look-up table.

Clinical significance of GAGs

The simplest non-sulphated GAG HA has many important functional roles (12), including signalling activity during embryonic morphogenesis (13), pulmonary and vascular diseases (14) and wound healing (15). Hyaluronic acid also acts as in the lubrication of synovial joints and joint movement, and its function has been described as space filler, wetting agent, flow barrier within the synovium and protector of cartilage surfaces (16). The influence of HA on cancer progression has been well described recently (14,17). The main receptor for HA is CD44 (18), which is expressed on the surface of virtually all stem cells including cancer stem cells. CD44/HA interactions can mediate leukocyte rolling and extravasation in some tissues and changes in CD44 expression contribute to tumour growth (19). Many cells also express receptor for hyaluronan and motility (RHAMM), which is a major HA-binding membrane protein (20), and this RHAMM pathway induces focal adhesions and signals the cytoskeletal changes, hence elevating cell motility seen in tumour progression, invasion and metastasis.

Anti-coagulation was the first described function for sulphated GAGs (21). Heparin was first discovered in 1917 because of its capacity to prolong the process of blood clotting, an effect due to its potentiating interaction with the natural inhibitor of thrombin, antithrombin III (AT-III), with only about one-third of all heparin chains possessing the structures required for AT binding (21). Heparin is mainly used in pharmaceutical products as an anti-coagulant for the treatment of thrombosis, thrombophlebitis and embolism. Pharmaceutical heparin is usually derived from bovine lung or porcine intestinal mucosa (22). Its name is derived from the fact that it was originally isolated from canine liver cells (from the Greek *hepar* for liver) (23,24). It has different molecular weights due to variations in chain length and is structurally heterogeneous.

Glycosaminoglycans play a major role in cell signalling and development, angiogenesis (25), axonal growth (26), tumour progression (27,28), metastasis (27,29) and anti-coagulation (30,31). Uncontrolled progenitor cell proliferation leads to malignant tissue transformation and cancer (32,33). Glycosaminoglycans and proteoglycans (PGs, see Proteoglycans) are believed to play a very important role in cell proliferation because they act as co-receptors for growth factors of the fibroblast growth factor (FGF) family. Indeed, members of the FGF family need to interact with both a heparin/HS chain and their high affinity receptor to realize their full signalling potential (25,34). Overexpression of these growth factors may contribute to tumour progression.

Sulphated GAGs are a common constituent in many different types of amyloid, playing an important role in the pathology of amyloid diseases such as amyloid A-amyloidosis, Alzheimer's disease, type-2 diabetes, Parkinson's disease and prion diseases (35). These diseases are characterized by the deposition in tissues of fibrillar aggregates of polypeptides. Heparan sulphate is known to bind amyloidogenic peptides *in vitro* and *in vivo,* promoting fibril formation and enhancing the disease condition. Heparan sulphate may sometimes be present within the amyloid β-containing deposits in Alzheimer's diseased brains (36).

Diseases such as rheumatoid arthritis, inflammatory bowel disease and microbial infections are associated with inflammatory responses. Many proteins play a role in the inflammation cascade that leads to the activation of leukocytes and endothelial cells, and ultimately to the extravasation of leukocytes and leukocyte migration into the inflamed or diseased tissue. Glycosaminoglycans such as heparin have important roles in these processes, as adhesion ligands in leukocyte extravasation and carriers/presenters of chemokines and growth factors (37).

Glycosaminoglycans are also known to promote microbial pathogenesis and invasion (38,39) by interacting with several microbial pathogens on cell surfaces and in the ECM. Many pathogenic micro-organisms such as bacteria (e.g. *Helicobacter pylori, Bordetella pertussis, Mycobacterium tuberculosis* and *Chlamydia trachomatis*), viruses (e.g. herpes simplex) and protozoa (e.g. *Plasmodium* and *Leishmania*) express proteins capable of binding to HS, DS and CS on cell surfaces, and these interactions appear to mediate infection (39). Dengue and foot-and-mouth viruses interact with cell surface heparan sulphate proteoglycans (HSPGs) and promote the concentration of virus particles at the cell surface after subsequent binding to integrin receptors. Heparin is known to exert its anti-HIV-1 activity by binding to the viral surface glycoprotein, gp120 (40,41), thus blocking HIV-1 entry into cells. Similarly, heparin/HS acts as the primary receptor for herpes simplex virus (HSV) (42,43). Heparan sulphate present on cell surfaces interacts with glycoproteins gB, gC and gD (44,45), which are known to be present on the viral envelope and enhance infection (46).

Proteoglycans

In nature, all GAG chains with the exception of HA are covalently linked to a core protein (Figure 1) to give a PG. The linkage of GAGs to the protein core involves a specific trisaccharide composed of two galactose (Gal) residues and a xylose (Xyl) residue (GAG-GalGalXyl-*O*-CH_2-protein). These saccharide residues are coupled to the protein core through an *O*-glycosidic bond to a serine residue. Some forms of KSs are linked to the protein core through an *N*-asparaginyl bond.

Virtually all mammalian cells produce PGs and either secrete them into the ECM, insert them into the plasma membrane, or store them in secretory granules. Examples of large PGs are aggrecan, the major PG in cartilage, and versican, which is present in many adult tissues including blood vessels and skin. A variety of core proteins have been shown to carry HS chains in the ECM and at the cell surface. Some membrane PGs such as sydecan-1 are hybrid structures known to contain both HS and CS (47).

Proteoglycans exhibit tremendous structural variation due to a number of factors, such as the differential expression of genes encoding core proteins, different exon usage of these genes and variations in the length and types of GAG chains (48). Different numbers of GAG chains with different saccharide sequences can be attached to the various serine residues present in the core protein. There are two major types of HSPGs (49): the syndecans and the glypicans. The core protein of each family differs: the syndecans are composed of an integral membrane protein whereas the glypicans have a glycosylphosphatidylinositol-anchored protein as their core protein. A HSPG with a core protein of a completely different structure is formed in the ECM (50). Thus, HSPGs are formed both on the cell surface and in ECMs. Proteoglycans are known to have affinity to a variety of ligands, including growth factors, cell adhesion molecules, matrix components, enzymes and enzyme inhibitors (4).

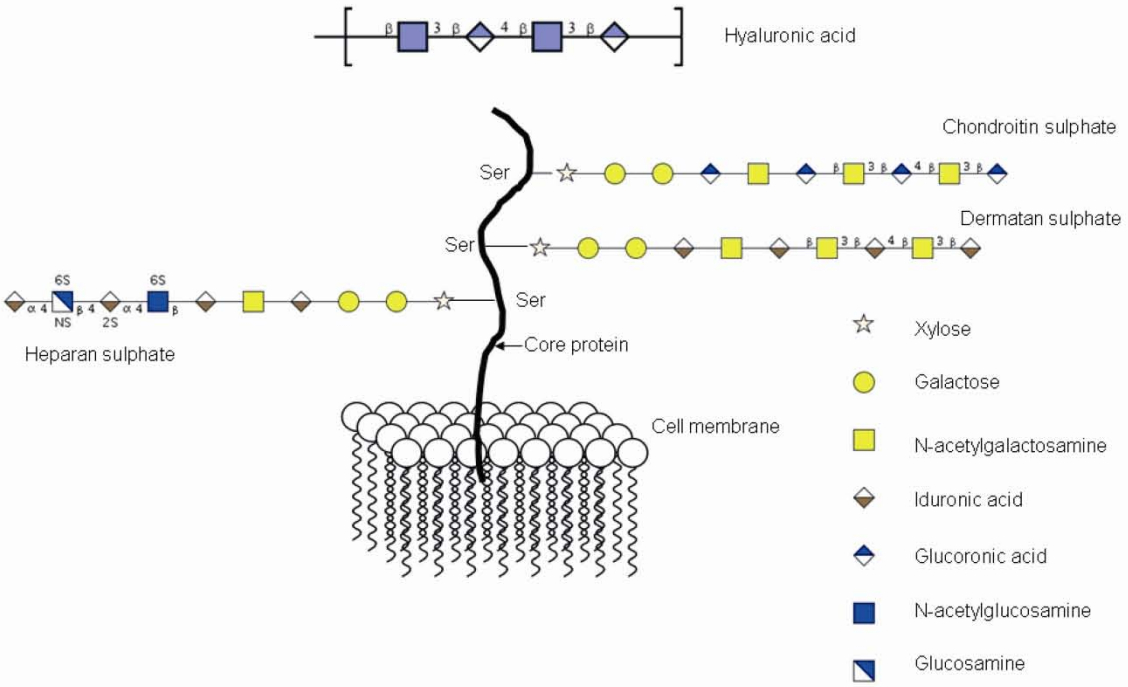


Figure 1: Structure of the linkage of GAGs to proteins in proteoglycans. Heparan sulphate, dermatan sulphate and chondroitin sulphate are linked to a protein via a serine residue. Note that hyaluronic acid is not found linked to any protein core. The image has been created using GlycanBuilder (232).

Sulphated GAGs: Heparin/Heparan Sulphate

The highly sulphated analogues heparin and HS have been studied extensively (7) due to their well understood functions in anticoagulation. Heparin is known to be highly evolutionarily conserved with similar structures found in a broad range of vertebrate and invertebrate organisms (51), such as turkey (52), whale (53), camel (54), mouse (55), human (56), lobster (57), shrimp (58), mussel (59), marine clam species (60) and crab (61). The difference between HS and heparin is quantitative and not qualitative (62), as can be seen in Table 3. Heparan sulphate contains a higher level of acetylated glucosamine and is less sulphated than heparin (63). Heparin is synthesized by and stored exclusively in mast cells, whereas HS is expressed on cell surfaces and in the ECM as part of a PG (48).

Heparin consists of repeating units of 1 → 4 linked pyranosyluronic acid and 2-amino-2-deoxyglucopyranose (glucosamine) residues. The uronic acid residues typically consist of 90% L-idopyranosyluronic acid (L-iduronic acid) and 10% D-glucopyranosyluronic acid (D-glucuronic acid). The amino group of the glucosamine residue may be substituted with an acetyl or sulphate group, or remain unsubstituted. The 3- and 6-positions of the glucosamine residues can either be substituted with an *O*-sulphate group or remain unsubstituted. The uronic acid, which can either be L-iduronic or D-glucuronic acid, may also contain a 2-*O*-sulphate group. Heparan sulphate is structurally related to heparin but is much less substituted with sulphate groups than heparin. Like heparin, HS is a repeating linear copolymer of an uronic acid 1 → 4 linked to glucosamine. D-glucuronic acid predominates in HS, but HS can also contain substantial amounts of L-iduronic acid. Heparan sulphate generally contains about one sulphate group per disaccharide, but its sulphate contents tend to vary (48). On the cell surface, the *O*-sulphonate and *N*-sulphonate groups are deprotonated in HS and attract positively charged counter ions to form a salt under physiological conditions.

Heparan sulphate chains often also contain domains of extended sequences having low sulphation compared with heparin, as illustrated in Figure 2. The non-sulphated regions that have a GlcA-GlcNAc (acetylated glucosamine) sequence are the most common in the HS chain, with IdoA-containing sulphated regions (called S-domains) usually of about 5–10 disaccharides (64). There are also relatively minor proportions of mixed sequences, which contain both $GlcNSO_3$ and GlcNAc (called NA-domain). A substantial proportion of the HS chain may consist of alternating GlcA–GlcNAc residues with no sulphate substitution.

The sulphated–acetylated–sulphated domain of HS has been subsequently found to be recognized by a number of chemokines, such as interleukin-8 (IL-8) (65), platelet factor 4 (PF4) (66) and macrophage inflammatory protein 1 alpha (MIP-1α) (67). The IL-8 dimer consists of two α-helical monomers lying on top of two β-sheets forming basic clusters on one face of the dimer. The two S-domains, each consisting of five to six saccharides in HS, accelerate the rate of dimer formation in IL-8 (65,68). The flexibility in the N-acetyl–rich 'spacer' or NA-domain (six to seven saccharides) in HS allows more conformational freedom for the simultaneous interactions of two S-domains and brings the monomers of IL-8 in close proximity to form a dimer in an anti-parallel arrangement (Figure 3). On the other hand, interferon-gamma (IFN-γ) does not significantly interact with isolated S-domains (69), in contrast to many other heparin binding proteins. Similarly, basic residues are clustered on both faces of the tetramer of PF4, requiring 21 saccharides in HS to form a more extended binding site on the charged surface of PF4. Heparin is assumed to be an analogue of the S-domains of HS, consisting mainly of sequences of sulphated disaccharides with IdoA2S (iduronic acid sulphated at C-2) and GlcNS6S (2,6-disulphoglucosamine).

Heparin and HS can often be structurally distinguished through their sensitivity towards microbial GAG-degrading enzymes, the heparin lyases. Three major polysaccharide lyases, heparin lyases I, II and III, isolated from *Flavobacterium heparinum*, are capable of cleaving linkages present in heparin and HS chains (70) by a beta elimination mechanism and each has three distinct substrate specificities (71,72). These three enzymes share very little homology at the DNA, protein or structural level, which imparts specificity towards the substrates. Heparin lyase I is involved in heparin binding whereas heparin lyase III exhibits strong specificity for HS. Heparin lyase II is believed to act on heparin and as well as on HS through two distinct active sites. The degrading enzymes work on the non-reducing ends leaving the region at the reducing end of HS usually unmodified (73). Distinct from bacterial heparinases, HS degradation by mammalian endoglycosidic enzyme heparanase has also been described in human placenta and rat liver hepatocytes. Heparanase cleaves the glycosidic bond through a hydrolytic mechanism, yielding HS fragments of appreciable size (10–20 sugar units), suggesting that the enzyme recognizes a particular HS structure (74).

Heparan sulphate proteoglycans are the major component of ECM in mammals (75). The structural heterogeneity of HS with respect to the size of the polysaccharide chain, the ratio of IduA to GlcA units, and the amount and distribution of sulphate groups along the carbohydrate backbone is the result of variations in the biosynthesis of HSPGs. The fine structure of the chains depends on the regulated expression and action of multiple biosynthetic enzymes, such as glycosyltransferases, sulphotransferases and an epimerase, which are arrayed in the lumen of the Golgi apparatus. The reactions catalysed by these enzymes do not go to completion, yielding individual chains whose sequences are likely to be distinct from all other chains (76).

Conformation of Heparin

Heparin is a linear, unbranched polysaccharide that tends to have an extended conformation in solution because of its highly hydrophilic nature arising from its extensive degree of sulphation. Analysis of the conformations of individual sugars within heparin (Figure 4) indicates that unsubstituted IdoA residues exist predominantly in the $^4C_1$ or $^1C_4$ chair form, whereas IdoA residues, when bearing a sulphate group at position 2 (IdoA2S), exist in equilibrium between a number of different conformations, the most important being the chair ($^2C_4$) and skew-boat ($^2S_0$) forms (77). Solution NMR

Table 3: Key differences between heparan sulphate and heparin

Property	Heparan sulphate	Heparin
Sulphate versus hexosamine content	0.8 1.8	1.8 2.4
2-deoxy-2-sulphamido-D-glucopyranosyl content	40–60%	>85%
L-iduronic acid content	30–50%	>70%
Site of synthesis	Extracellular component found in the basement membrane and as a ubiquitous component of cell surfaces	Intracellular component of mast cells, especially in the liver, lungs and skin
Mass	10–70 kDa	10–12 kDa
Major disaccharide repeating units	D-GlcA-β(1→4)-D-GlcNAc-α(1-4)	L-IdoA2S-α(1→4)-D-GlcNS6S-α(1→4)
Minor (variable) disaccharide repeating units (X = H or SO_3^-; Y = Ac, SO_3^- or H)		

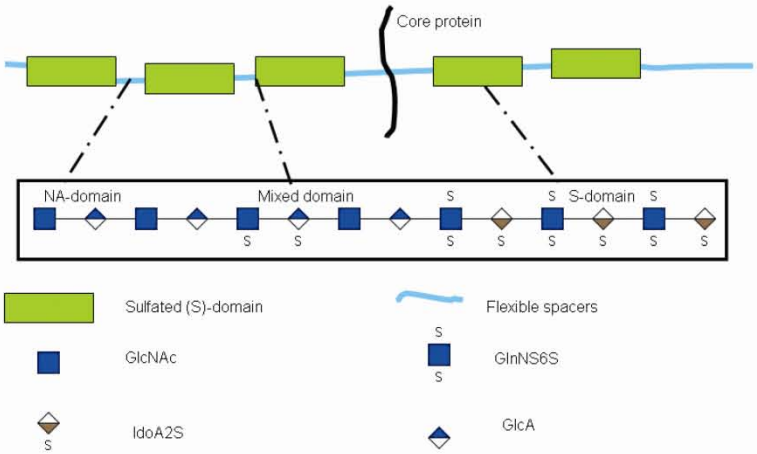


Figure 2: Multidomain structure of HS. The distributed sulphated domains of HS are separated by flexible spacers of low sulphation. The mixed sequences define transition zones between the S-domains and the unmodified *N*-acetyl-rich regions. Several monomeric or oligomeric proteins can bind to GAGs, often by recognizing different structural features of the domain. The figure has been created using GlycanBuilder (232).

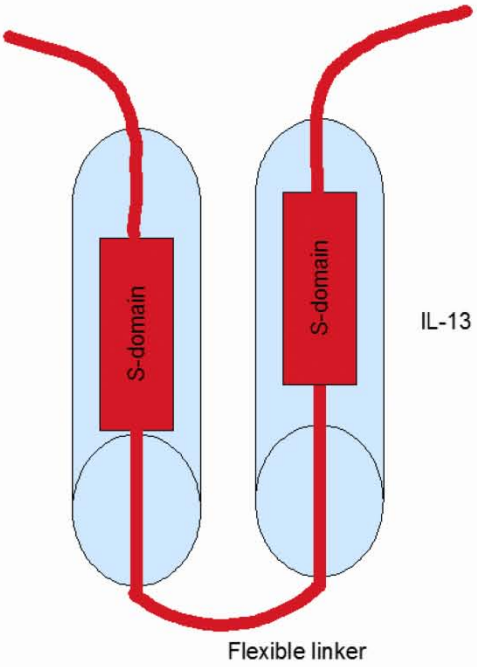


Figure 3: Schematic representation of the dimerization of IL-8 by HS. HS is coloured red and the IL-8 monomer is shown in blue. *N*-sulphated stretches of HS (6 sugar units) are known to interact with the α-helical heparin/HS-binding domains of each IL-8 monomer. The rate of dimer of formation of IL-8 is accelerated by the S-domains, whereas the flexible spacer sequence (NA-domain with 12–14 monosaccharide units) allows appropriate folding of the monomer in an anti-parallel arrangement.

studies suggest that IdoA2S prefers the $^{2}S_{0}$ and $^{1}C_{4}$ conformation, whereas glucosamine sulphated at the *N* and *O* positions (GlcNS6S) prefers the $^{4}C_{1}$ conformation (78,79). It seems that glucosamine and its derivatives are stable in the $^{4}C_{1}$ chair conformation irrespective of substitution (78,80,81).

Heparin oligosaccharides sometimes contain a non-reducing terminal 4-deoxy-L-threo-2-sulphohex-4-enopyranosyluronic acid (unsaturated Δ^{4}-uronic acid, ΔUA2S) residue arising from heparin lyase cleavage of an HS chain. Based on the conformation of the 4,5-double bond, ΔUA2S can exist in either the $^{2}H_{1}$ or $^{1}H_{2}$ conformations (Figure 4) and the equilibrium between these two conformations is controlled by their substitution pattern. The solution structures of heparin-derived oligosaccharides determined by NMR spectroscopy suggest that the terminal ΔUA2S residue exists predominantly in the $^{1}H_{2}$ form, with a minor contribution from the $^{2}H_{1}$ conformation (82).

The solution structure of a heparin dodecasaccharide composed of six GlcNS6S–IdoA2S repeat units has been determined using a combination of NMR spectroscopy and molecular modelling techniques (79). These two structures (Figure 5) have been deposited in the protein data bank (PDB) under code 1HPN. One structure has all IdoA2S residues in the $^{2}S_{0}$ conformation (Figure 5A) and the other one has all IdoA2S residues in the $^{1}C_{4}$ conformation (Figure 5B). The three-dimensional structure of heparin is thus complicated by the fact that iduronic acid may be present in either of two low energy conformations when internally positioned within an oligosaccharide. This conformational equilibrium can be influenced by the sulphation state of adjacent glucosamine sugars (83). The $^{2}S_{0}$ form appears to be slightly favoured in terms of conformational stability, as it tends to minimize the unfavourable 1,3 diaxial non-bonded interactions that are expected in the $^{1}C_{4}$ form, where four of the substituents are axially oriented and only the carboxylate group is equatorial (84). Whilst the spatial orientation of the 2-*O*-sulphate group in the IdoA2S residues is altered during $^{1}C_{4}$–$^{2}S_{0}$ interconversion, no significant conformational change can be seen in the backbone of the polysaccharide chain in the NMR structures. In these NMR structures heparin adopts a helical conformation, the rotation of which places clusters of sulphate groups at regular intervals of about 17 Å on either side of the helical axis.

The iduronate ring can adopt either the $^{2}S_{0}$ or $^{1}C_{4}$ forms in the protein-bound state, which enables it to make specific electrostatic interactions with the electropositive surface regions of a protein. Nonetheless, the helical parameters of heparin oligosaccharides are conserved in spite of the conformational flexibility of the

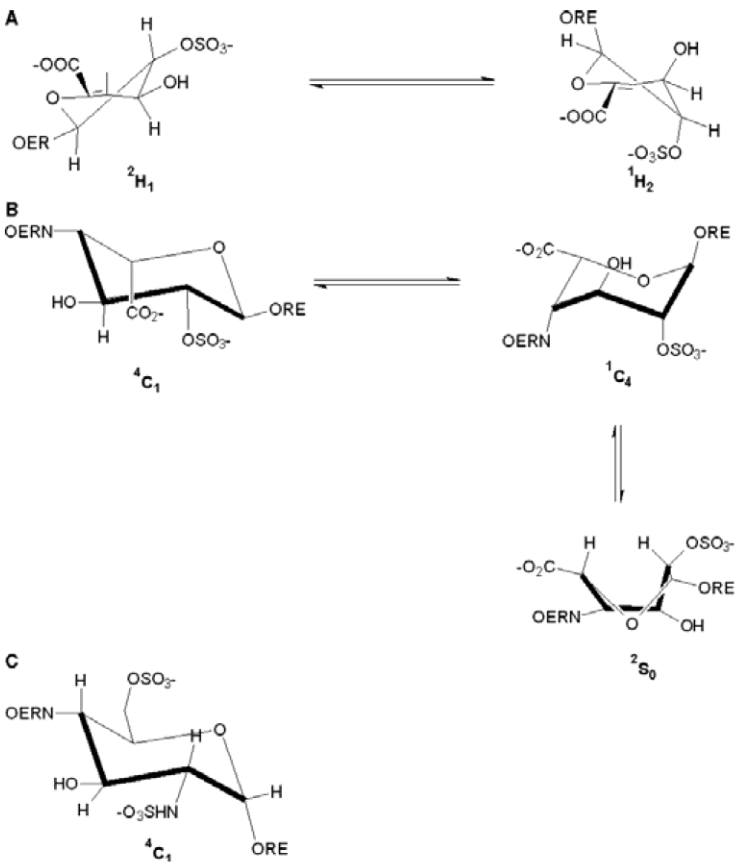


Figure 4: Conformations of Δ^4-uronic acid, sulphated iduronate and glucosamine: (A) $^2H_1$ and $^1H_2$ conformations of unsaturated Δ^4-uronic acid. (B) $^4C_1$, $^1C_4$ and $^2S_0$ conformations of IdoA2S. (C) The predominant $^4C_1$ conformation of GlcNS6S. RE and NRE correspond to the reducing and the non-reducing end of a saccharide sequence, respectively.

L-iduronate residues. NMR studies of a series of modified heparins with systematically altered substitution patterns indicate that all derivatives in the unbound form, regardless of the sulphation pattern, exhibit similar glycosidic bond ψ and φ conformations (85). The conformations of these glycosidic linkages are also observed in the X-ray structures of heparin fragments in complex with proteins such as acidic FGFs (86) (PDB codes 1AXM, 2AXM, 1E0O, 1FQ9) and many other heparin structures bound to proteins, such as AT (PDB codes 1AZX, 1E03), bFGF (PDB code 1BFC), annexin V (PDB code 1G5N) and foot and mouth disease virus (PDB code 1QQP).

Depending on the local sequence, the conformation of heparin/HS oligosaccharides may be affected by the degree of flexibility in the disaccharide subunits and the surrounding water and cations (87). A recent theoretical study (88) determined the stable conformations of 1-OMe IdoA2SNa$_2$ ($^2H_1$ and $^1H_2$ conformations), 1-OMe GlcNS6SNa$_2$, 1,4-DiOMe GlcNa, 1,4-DiOMe GlcNS3S6SNa$_3$, 1,4-DiOMe IdoA2SNa$_2$ ($^4C_1$, $^1C_4$ and $^2S_0$ conformations) and 1,4-DiOMe GlcNS6SNa$_2$ monomers and their ionized forms both in the gas phase and in the presence of solvent and cations. In the gas phase, the $^2H_1$ conformation of the uronate residue is more stable than the $^1H_2$ form observed in the presence of water. The most stable structure was observed to be the 1,4-DiOMe GlcN-S6SNa$_2$ monomer in the skew-boat $^2S_0$ conformation in water. The $^1C_4$ conformation is the most stable form in the presence of anions. In general, the results indicate that the relative stability of cation-heparin ionic interactions is considerably diminished in aqueous solution. Another study using similar saccharides have revealed that only two negatively charged oxygen atoms in the SO_3^- group are involved in co-ordination of the sodium cation (89).

Various studies of heparin conformations have revealed similar, well-defined molecular structures in terms of overall chain conformation, both in the solid state and in solution, as a result of the flexibility of the pyranose ring of iduronic acid, which results in either the $^1C_4$ or $^2S_0$ conformations (90). However, variations in the primary sequence of GAGs and the degree of sulphation can result in different binding modes with proteins that can affect their activity.

Conformation of Heparin Fragments Bound to Proteins

Iduronate may exist in skew-boat ($^2S_0$), chair ($^1C_4$) and intermediate ring (a mixture of $^1C_4$ and $^2S_0$) conformations in heparin–protein complex crystal structures depending on either receptor specificity or predominance in solution. The central iduronate in the crystal structure of a heparin pentasaccharide with the foot and mouth virus (PDB

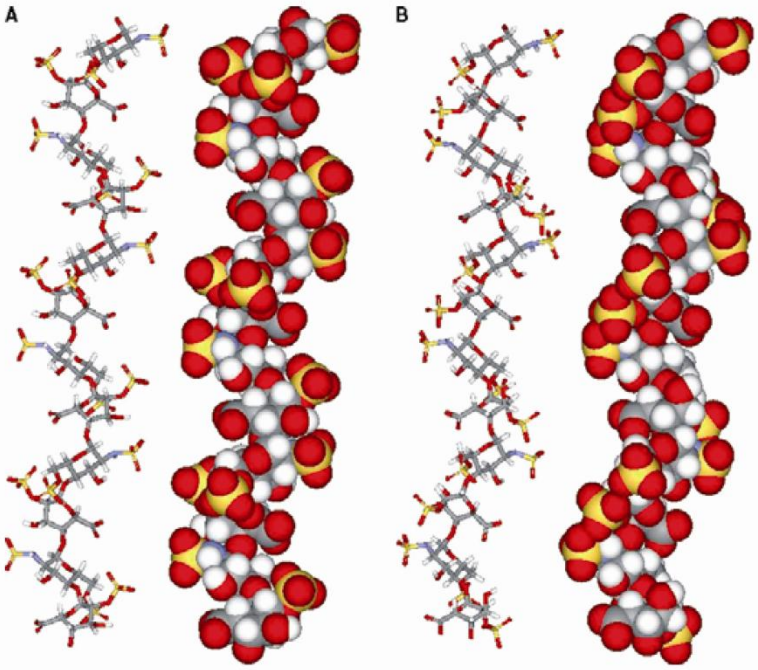


Figure 5: Solution structure of a heparin dodecasaccharide (PDB code 1HPN) in which all IdoA2S are in the $^{2}S_{0}$ conformation (A) or in the $^{1}C_{4}$ conformation (B). The stick and CPK models were generated using DS Visualizer 2.0[c]).

code 1QQP) has an intermediate conformation, whereas the outer iduronates are in the $^{1}C_{4}$ and $^{2,5}B$ conformations (38). The third iduronate ring in the crystal structure of the complex of a heparin hexasaccharide with bFGF (PDB code 1BFC) adopts a $^{1}C_{4}$ chair conformation and the other at the fifth position adopts a $^{2}S_{0}$ skew boat conformation (91). However, in the crystal structure of a heparin tetrasaccharide with annexin V the IdoA2S residue that interacts with the protein adopts a $^{2}S_{0}$ skew conformation, whereas the non-interacting IdoA2S residue is in the $^{1}C_{4}$ conformation (92). These structures suggest that when heparin binds to a protein, a change in the conformation of the IdoA2S residue may be induced, resulting in a better fit and enhanced binding, whilst the conformation of the less flexible GlcNS6S residue remains unaltered (i.e. $^{4}C_{1}$ conformation).

Studies of the conformation and dynamics of a heparin pentasaccharide (denoted as AGA*IA) have investigated its high affinity interactions with AT, both in the crystal (93) and solution states (94). The protein-bound pentasaccharide has a conformation roughly similar to one of the conformations predicted by molecular mechanics calculations for the pentasaccharide, wherein the iduronate residue adopts a conformation between the $^{2}S_{0}$ skew-boat and $^{2,5}B$ conformations (95). NMR studies of the complex of a heparin tetrasaccharide with AT reveal a distinct change in conformation of the glycosidic linkage upon binding and a stabilization of the $^{1}C_{4}$ chair conformation of the iduronate residue (96), which is not observed in the solid state structure. A conformational change in the geometry around the glycosidic linkage between the non-reducing end glucosamine and the adjacent glucuronic acid residue is also observed upon binding of heparin to AT, as compared with the solution state (97,98). Similarly, NMR and simulation studies of heparin octasaccharides containing the AT binding pentasaccharide sequence (AGA*IA) indicated that the non-sulphated IdoA residue preceding AGA*IA can exist in the $^{1}C_{4}$ conformation when bound to AT and in the $^{2}S_{0}$ conformation in the absence of AT (99). These altered conformations of IdoA do not affect the binding affinity for AT (99).

The conformation of heparin oligosaccharides bound to growth factors in solution has been studied using NMR. Studies of heparin tetrasaccharides in the presence of fibroblast growth factors aFGF and bFGF indicate that FGF binding stabilizes $^{1}C_{4}$ conformation of the IdoA2S residue directly involved in binding (98) whereas the solid state structure of the complex of a heparin hexasaccharide with bFGF revealed that one of the IdoA2S residues adopts the $^{1}C_{4}$ chair, and the other the $^{2}S_{0}$ skew boat (91). NMR studies also confirmed the crucial role of the 6-*O*-sulphate group on at least one GlcNS6S residue required for the formation of the complex with aFGF, but not with bFGF (98). This also indicates that GAG-binding site specificity varies among family members. These examples suggest that the complexation of GAGs with proteins induces a change in conformation of IdoA2S, resulting in a better binding mode and high affinity, whilst the conformation of GlcNS6S remains unaltered.

Interactions of Heparin/Heparan Sulphate with Proteins

Numerous studies have identified common structural features in the heparin/HS binding sites of proteins. Different structural (NMR spectroscopy and X-ray crystallography) and molecular modelling approaches have been used to elucidate the three-dimensional features and structure-activity relationships of GAG–protein interactions (100). A list of the different proteins that have been

crystallized in complex with heparin oligosaccharides and their characteristics, such as the optimal length required for binding and their binding affinities can be found in Table 4. The crystal structure of some of these proteins, such as IL-8, PF-4 and NCAM (Neural Cell Adhesion Molecule), are not available in complex with GAGs. A full list of GAG-binding proteins is available in the GAGPROT database[5].

The nature of GAG–protein interactions

Structure and sequence-based statistical analyses indicate that Asn, Asp, Glu, Gln, Arg, His and Trp are more likely to make up the binding sites for non-sulphated carbohydrates than other amino acids (101–103). The aromatic residue Trp has a significantly higher mean solvent accessibility in carbohydrate binding locations, whereas aliphatic residues Ala, Gly, Ile and Leu, hydrophobic residues which are usually buried inside proteins, do not appear to participate in sugar binding. The aromatic ring in Trp can pack against the hydrophobic face of a sugar molecule.

Strong ionic interactions are expected between GAGs and proteins. Clusters of positively charged basic amino acids on proteins form ion pairs with spatially defined negatively charged sulphate or carboxylate groups on heparin chains. Glycosaminoglycans interact with residues that are prominently exposed on the surface of proteins. The main contribution to binding affinity comes from ionic interactions between the highly acidic sulphate groups and the basic side chains of arginine, lysine and, to a lesser extent, histidine (104). The relative strength of heparin binding by basic amino acid residues has been compared and arginine has been shown to bind 2.5 times more tightly than lysine. The guanidino group in arginine forms more stable hydrogen bonds as well as stronger electrostatic interactions with sulphate groups. The ratio of these two residues is said to define, in part, the affinity of a binding site in a protein for GAGs (105).

The interactions of GAGs with proteins also involve a variety of different types of interactions, including van der Waals (VDW) forces, hydrogen bonds and hydrophobic interactions with the carbohydrate backbone. It has also been observed that heparin-binding domains contain amino acids such as asparagine and glutamine which are capable of hydrogen bonding. The affinity of heparin-binding proteins for heparin/HS was also enhanced due to the presence of polar residues with smaller side chains like serine and glycine. These residues provide minimal steric constraints and good flexibility for the interaction with GAGs (106). Ionic interactions contributed 30% to the free energy of binding of heparin to bFGF and non-ionic forces such as hydrogen bonding and hydrophobic interactions also contributed to the affinity of low molecular weight heparin (LMWH) heparin to bFGF (107). Studies of the interaction of heparin with the brain natriuretic peptide (BNP) revealed that only a small portion of the free energy of binding arises from ionic interactions (6%), whereas the major contribution arises from hydrogen bonding (94%) between polar amino acids on BNP and heparin (108). Hydrophobic interactions can also play an important role in heparin-protein interactions. NMR studies reveal that a tyrosine residue in a synthetic AT peptide makes specific hydrophobic interactions with the *N*-acetyl group of a GAG pentasaccharide from porcine mucosal heparin (108).

Structural studies of the complex between a heparin pentasaccharide and AT have shown that basic amino acids participate in five to six ionic interactions, contributing 40% of the binding energy, whereas non-ionic interactions are responsible for the remaining 60% of the binding energy (109). Two aromatic residues, Phe 121 and Phe 122, which lie near basic amino acids in the heparin-binding domain, make direct contact with the pentasaccharide. Phe 121 and Phe 122 were mutated to Ala and Leu, respectively, resulting in decreased affinity of heparin for AT. These residues thus appear to play a critical role in heparin binding and AT activation through hydrophobic and VDW interactions (109). The positively charged basic amino acid residues Arg 47, Lys 114, Lys 125, and Arg 129 have been identified as the most important in the heparin

Table 4: Characteristics of some of the known complexes between heparin or heparan sulphate fragments and proteins

Name of protein	Type of protein	PDB code	Size of oligosaccharide	K_d	References
NCAM	Adhesion protein	–	5-mer	52 nM	195,196
Fibronectin	Adhesion protein	1FNH	8- to 14-mer	μM	122,197
IL-8	Chemokine	–	18- to 20-mer	6 μM	47
PF-4	Chemokine		12-mer	nM	48
RANTES	Chemokine	1U4L, 1U4M	16- to 18-mer	32 nM	198
Annexin V	Extracellular protein	1G5N	8-mer	20 nM	199
Annexin A2	Extracellular protein	2HYU, 2HYV	4- to 5-mer	30 nM	121
Amyloid P (AP)	Glycoprotein	–	4-mer	μM	200
Basic fibroblast growth factor (bFGF)	Growth factor	1BFB, 1BFC	4- to 6-mer	nM	72
Acidic fibroblast growth factor (aFGF)	Growth factor	1AXM, 2AXM	4- to 6-mer	nM	68
Heparin binding growth associated molecule (HB-GAM)	Growth factor		16- to 18-mer	10 nM	201
aFGF/ecto-domain of FGF receptor 2 (FGFR2)	Growth factor/receptor	1E00	12-mer	nM	108
bFGF/ecto-domain of FGF receptor 1 (FGFR1)	Growth factor/receptor	1FQ9	12-mer	nM	202
Thrombin	Protease	1XMN	8-mer	7 μM	203
Secretory leukocyte protease inhibitor (SLPI)	Protease	–	12- to 14-mer	6 nM	204
AT-III	Serpin	1AZX, 1E03, 1NQ9	5-mer (synthetic pentasaccharide)	20 nM	74
AT-III/factor Xa	Serpin/protease	2GD4	5-mer (Fondaparinux)	100–200 nM	177
Cardiotoxin A3, A5, M4 and M1	Toxin	1XT3	5- to 7-mer	μM	205
HIV-1-gp120	Viral pathogen	–	10-mer	0.3 μM	28,206,207
Dengue viral envelope protein	Viral pathogen	–	10-mer	15 nM	208

Table 5: List of heparin-binding proteins that contain the Cardin–Weintraub consensus sequence

Protein	XBBBXXBX
Annexin II	KIRSEFKKKYGKSLYY
Vitronectin	QRFRHRNRKGYRSQRG
ApoB	KFIIPSPKRPVKLLSG
bFGF	GHFKDPKRLYCKNGGF
NCAM	DGGSPIRHYLIKYKAK
Protein C inhibitor	GLSEKTLRKWLKMFKK
AT-III	KLNCRLYRKANKSSKL
ApoE	SHLRKLRKRLLRDADD
Fibrin	GHRPLDKKREEAPSLR
FGFR-1	AAPVAHLKKEMK
B-thromboglobulin	PDAPRIKKIVQKKLAG
Insulin-like growth factor-binding protein-3	DKKGFYKKKQCRPSKG

binding site of AT using chemically modified, naturally occurring mutant and recombinant ATs (110–113). These basic residues participate in ionic interactions with the negatively charged groups of heparin, as observed in the crystal structure of an AT-pentasaccharide complex (93). Residues Arg 129, Lys 114, and Arg 47 are critical for the heparin-induced conformational change of AT, contributing to the resulting high affinity of interaction (113).

Consensus sequences in GAG binding proteins

The X-ray crystal structures of many GAG-binding proteins has helped to determine the existence of amino acid consensus sequences for GAG binding with common features such as the arrangement of basic amino acids. For example, the heparin-binding sequence WQPPRARI and the sequence WSPW have been identified as the GAG binding motifs at the C-terminal region of fibronectin and thrombospondin, respectively (114,115). Cardin and Weintraub (116) analysed the structures of 21 heparin-binding proteins and proposed that typical heparin-binding sites have the sequence XBBXBX or XBBBXXBX, where B is a lysine or arginine (with a very rare occurrence of His) and X is a hydropathic residue. The 'X' in the consensus sequences was defined as a hydropathic residue based on the frequency of occurrence of residues at specific positions from known heparin binding proteins. The residues Asn, Ser, Ala, Gly, Ile, Leu and Tyr were more common at positions 'X'. Residues such as Cys, Glu, Asp, Met, Phe and Trp exhibited a very low occurrence at positions 'X' in either the α-helical or β-sheet domains of heparin binding proteins. Table 5 lists the heparin-binding proteins that contain the Cardin–Weintraub consensus sequence.

Depending on the secondary structure of the protein, very few residues in these consensus sequences may actually participate in GAG binding. Glycosaminoglycan-binding sites are often found along one exposed face of a protein and sometimes wrap around multiple faces in the case of β-sheets.

The basic amino acids of the sequence XBBBXXBX, when they belong to an α-helix, are usually displayed on one side forming an amphiphatic helical arrangement (Figure 6A). Therefore, in order to interact with a linear GAG chain, it would be predicted that the

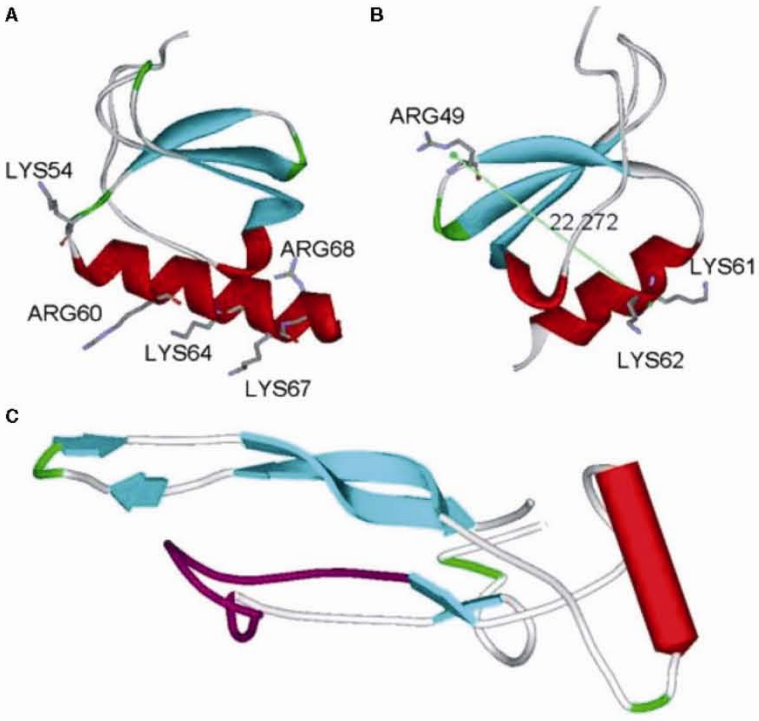


Figure 6: Types of GAG consensus sequences present in proteins. (A) Linear XBBBXXBX motif with basic arginine and lysine residues (blue) oriented on one surface of a helix (green, residues 53–72), based on the structure of interleukin-8 (PDB code 3IL8). (B) Linear motif with basic arginine and lysine residues (blue) spaced at a 20 Å linear distance, located on opposite surfaces (green, residues 48–50 and 60–62), as observed in the structure of platelet factor-4 (PDB code HPF4). (C) Linearly contiguous GAG-binding domains with the consensus sequence TXXBXXTBXXXTBB (magenta), based on the structure of TGFβ-1 (PDB code 1KLC) (residues 23–41).

positively charged amino acid residues in α-helical proteins would have to line up along the same side of the protein segment. Comparative analysis of heparin binding sequences have shown that basic amino acids are generally located about 20 Å apart (Figure 6B) in an amphipathic helical structure, and the same spatial arrangement is preserved in a β-strand structure (117). For example, the sulphates that mimic HS in the Artemin crystal structure (118) were found to be separated by approximately 8–9 Å and arranged at the vertices of an approximate equilateral triangle in the prehelix (with a positively charged heparin consensus sequence XBBXBX) and amino-terminal regions.

In β-strands, the positively charged residues in a GAG-binding protein are located in a different way compared with α-helical structures. The basic amino acids in the sequence XBBXBX line up on one face of a β-strand, whereas the hydropathic residues points back into the protein core. An example of β-sheet heparin binding proteins is the cobra cardiotoxins, which contain nine discontinuous basic residues (-B-X_{2i-1}-B-, where X is any residue and B is basic residue) separated by an odd number of any other residue (119).

A third consensus sequence was similarly proposed in the heparin binding protein von Willebrand factor: XBBBXXBBBXXBBX, where 'B' represents a cationic residue (120). The consensus sequence TXXBXXTBXXXTBB, as shown in Figure 6C, was also observed in aFGF, bFGF and transforming growth factor β-1 (TGFβ-1), where T defines a turn, B a basic amino acid (arginine or lysine) and X a hydropathic residue. The spatial distance between each of the three turns present in the consensus of these crystal structures was 12–18 Å (105).

It should be noted that the spacing of clusters of basic residues can also provide structural clues about heparin-binding sites that may be important for interactions with GAGs and which can facilitate the design of peptides that bind heparin efficiently (105).

Glycosaminoglycans binding sites are often not conserved between proteins, as observed in the case of chemokines, which have high structural similarity but do not share common GAG binding regions (121). Platelet factor 4 and IL-8 are members of the α-chemokine family that have very similar monomeric three dimensional structures, with anti-parallel β-strands and an α-helix in the C-terminus. PF4 has a heparin/HS binding consensus sequence 'KKIIKK', where K is lysine and I is an isoleucine protruding from the α-helix. The GAG consensus sequence in the equivalent α-helical domain of IL-8 is 'KENWVQRVVEKFLKR', which is responsible for heparin/HS binding. The heparin/HS binding proteins of the β-chemokine subfamily [e.g. MIP-1α, regulated on activation, normal, T-cell expressed, and secreted (RANTES)] use a different structural motif, 'KRNR'. Members of both chemokine α and β families have additional residues and hence lack conservation in their GAG-binding regions, allowing specificity and selectivity of HS binding across chemokines.

Structural properties of GAG-binding proteins

The X-ray crystal structures of heparin–protein complexes have provided information on the structural features required for heparin binding, such as the protein fold, the periodicity of clusters of basic residues, the periodicity of sulphate clusters on the GAG chains and the sulphation level required for interactions with the binding site (105). Heparin binding sites can be formed by basic amino acids that are distant in sequence but are brought spatially close together in the final fold of the protein. The end-to-end lengths of these extended clusters are comparable to the minimum GAG chain lengths that are required for binding (typically 6–12 monosaccharide units, approximately 25–50 Å long). The binding of GAG fragments to chemokines has a strong length dependence but it is clearly not the only determinant of selectivity (122).

The periodicity of sulphate group clusters in an oligosaccharide chain can play a key role in determining the structure of the GAG binding site on the surface of either α-helical or β-sheet proteins. The regular periodicity of sulphate group clusters along one side of an oligosaccharide chain was consistent with the ability of heparin to induce an α-helical structure in polylysine peptides, allowing electrostatic interactions every three peptide turns between a HS cluster and a zeta-amino group of a polylysine peptide (123). A heparin octasaccharide was the minimal fragment size required for such interactions to occur with the polylysine peptide. A similar phenomenon has been detected for several lysine-rich regions in the Tau protein (124), wherein a heparin oligosaccharide wraps tightly around the outer surface of the (double) pleated sheets, inducing secondary structural changes and thereby neutralizing the inhibitory charge repulsions that would occur in a parallel stacking of the repeat regions formed by a polylysine stretch.

The relative proportion of *N*- and *O*-linked sulphate groups and *N*-linked acetyl groups in heparin/HS can effect their interaction with proteins. In the case of RANTES, *O*-sulphation appears to be more important than *N*-sulphation (122). Macrophage inflammatory protein-1 alpha, monocyte chemoattractant protein-1 (MCP-1) and IL-8 showed preference for both *N*- and *O*-sulphation. The binding of GAG fragments to chemokines requires both *N*- and *O*-sulphation (122). In addition, binding studies involving chemically modified heparins or HS preparations have shown that 2-*O*- and *N*-sulphate groups are important for interactions with bFGF (Figure 7) and do not require 6-*O*-sulphate group for binding. Glycosaminoglycan fragments requires 2-*O*-, 6-*O*- and *N*-sulphate groups for optimal interaction with the HIV-Tat protein (125).

Heparin–AT interactions: a case study of GAG–protein binding

The anti-coagulant activity of heparin arises primarily through its activation of the AT-mediated inhibition of blood coagulation factors such as thrombin and factor Xa, as depicted in Figure 8. The interaction of AT and its coagulation factors with heparin involves number of affinity states to terminate in a high affinity interaction. Heparin activates AT by two different mechanisms namely, conformational activation and bridging mechanism (126). First, the interaction between GAG and AT is mediated by a well-defined unique pentasaccharide sequence within heparin. This interaction generates a conformational change in the structure of AT, which enables additional interactions between AT and heparin, resulting in stronger binding. The conformational change also expels a protease reactive

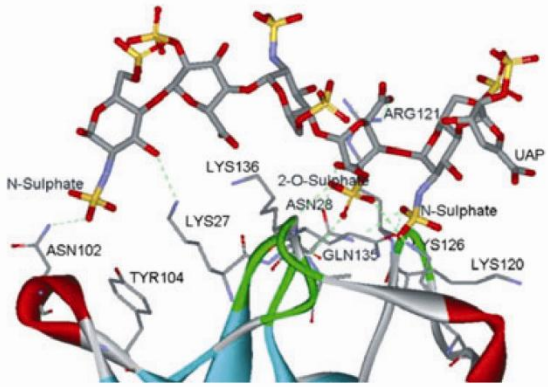


Figure 7: Distinct role of sulphate groups of heparin (shown in sticks) in the interactions with basic residues (shown as lines) of bFGF (shown in ribbons). The figure was prepared from PDB code 1BFC using DS Visualizer 2.0[c].

centre loop (RCL) in AT. This conformational activation mechanism promotes binding of RCL to the active site of factox Xa. After the complex formation, the AT-III interaction reverts to low-affinity binding followed by RCL cleavage, resulting in the release of heparin from the covalent AT-III–factor Xa complex. These conformational changes in RCL does not affect the binding of AT to thrombin caused by difference in active site in this proteinase. Full length heparin promotes the interaction of AT with thrombin by means of bridging mechanism of activation. During this mechanism, a positively charge surface in thrombin binds non-specifically to the extended heparin polysaccharide (Figure 8).

Several theories have been proposed in relation to the length dependence of the interaction of heparin with AT and serine proteases. Heparin chains of at least 16 saccharides in length are required to accelerate the reaction of AT with thrombin, even though only the pentasaccharide sequence is necessary to bind to AT (127). By contrast, heparin chains as small as the AT binding pentasaccharide are able to accelerate the inactivation of the other target coagulation enzymes, such as factor Xa.

GAG–FGF interactions

The extracellular domains of fibroblast growth factors aFGF (FGF-1) and bFGF (FGF-2) have been extensively studied to determine the thermodynamics and kinetics of their interactions with heparin. These growth factors exert their biological effects by binding to different, specific cell surface FGF receptors (FGFRs). In the high-resolution X-ray crystal structure of a 2:2:2 dimeric ternary complex of bFGF, FGFR-1 and a heparin decasaccharide, heparin makes numerous contacts with both bFGF and FGFR-1, stabilizing the FGF-FGFR interaction (128). Heparin also makes contacts with the FGFR-1 of the adjacent FGF-FGFR complex, thus seemingly promoting FGFR dimerization (Figure 9). The 6-*O*-sulphate groups of heparin play a major role in promoting these interactions (129).

The crystal structure of a 2:2:1 complex of aFGF, FGFR-2, and a heparin decasaccharide has also been determined at a resolution of 2.8 Å (130). The complex is assembled around a central asymmetric heparin molecule linking two aFGF ligands into a dimer that bridges the interaction between the two receptor chains (Figure 10). The heparin fragment makes contact with both aFGF molecules but only with one receptor chain. It is clear those different members of the FGF family and their respective receptors may interact differently

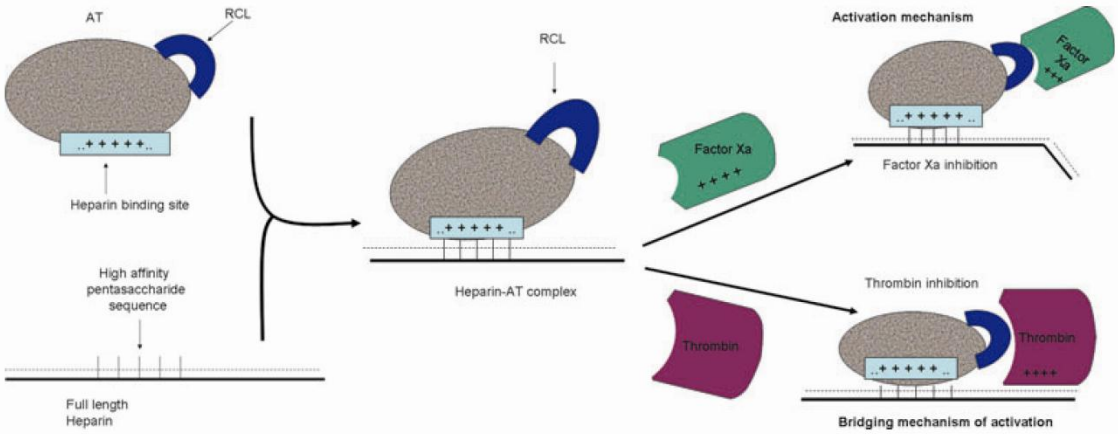


Figure 8: Schematic representation of the mechanism of AT inhibition of factor Xa and thrombin in the presence of full length heparin. Circulating AT forms a complex with endothelial heparin following an interaction of the high affinity pentasaccharide sequence in full length heparin with the heparin binding site. This leads to the exposure of the reactive site loop (RCL), which recognizes factor Xa and is known to provide a conformational activation mechanism. Thrombin inhibition occurs due to the interaction of thrombin and AT with full length heparin through a bridging mechanism of activation. Some negative charges available at the full length heparin chain binds non-specifically to the exosite (positively charged region) of thrombin. In the figure, + + + denotes a positively charged region and – – – denotes a negatively charged polysaccharide.

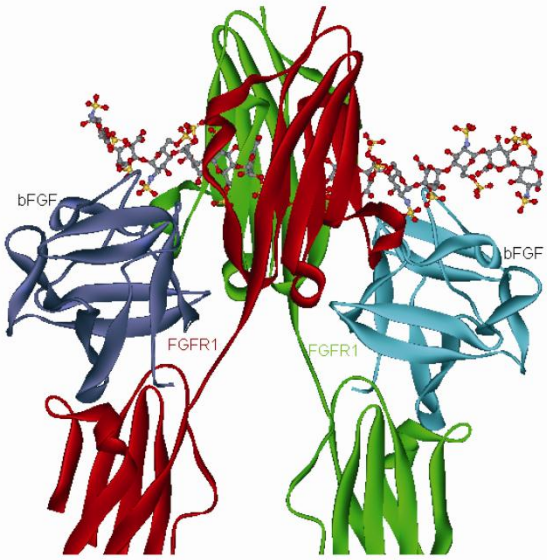


Figure 9: Schematic representation of the bFGF–FGFR1 complex. Heparin requires both 2-*O*, 6-*O*- and *N*-sulphate groups to promote the binding of bFGF to soluble FGFR-1. The binding of heparin/HS to bFGF, without 6-*O*-sulphate groups is not sufficient to induce the interaction of bFGF with FGFR. The heparin fragment (shown in balls and sticks) makes contact with both bFGF molecules (beta strands shown in purple and blue) and the FGFR1 receptor chain (immunoglobulin domains shown in red and green). The figure was prepared using PDB code 1FQ9 with DS Visualizer 2.0[c].

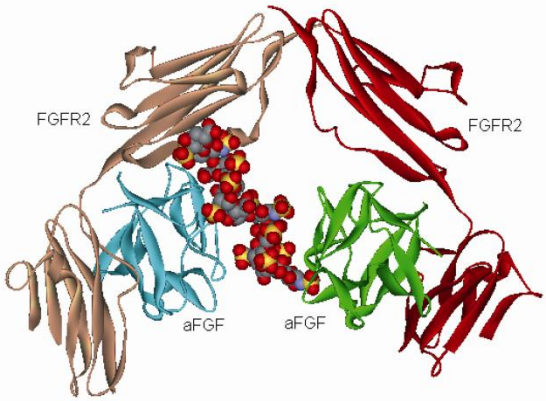


Figure 10: Ribbon diagram of the aFGF-FGFR2-heparin complex (PDB code 1E00). The heparin fragment (shown in CPK) makes contacts with both aFGF molecules (beta strands shown in green and blue) but only with one FGFR2 receptor chain (immunoglobulin domains shown in red and brown).

with heparin/HS as a result of the heterogeneity in the structure of HSPGs and FGFRs on cell surfaces in different tissues. It has been reported that aFGF may recognize several conformations of the iduronic residues of a heparin hexasaccharide. It is believed that the hexasaccharide undergoes local $^{1}C_{4}$–$^{2}S_{0}$ equilibrium conformational changes as a result of ionic interactions with flexible Arg and Lys side chains present in the protein (131).

Role of pH in GAG Binding

Certain heparin/HS–protein interactions are regulated by pH. Alteration of the pH can have profound effects on the ability of some proteins to bind heparin or HS. This is the case of the synthetic beta-amyloid peptide (Aβ) (132), selenoprotein P (133), the granulocyte macrophage colony stimulating factor (GM-CSF) (134), the mouse mast cell protease 7 (135), the stromal cell-derived factor-1 (SDF-1) (136) and the platelet endothelial cell adhesion molecule 1 (PECAM-1) (137,138). The modifying effect of pH arises particularly when the GAG binding site in a protein contains histidines, because the side chains of these amino acids have a pKa of approximately 6. Hence, if the pH falls closer to 6 an increasingly larger proportion of histidines will become protonated and hence positively charged, thus favouring the formation of electrostatic interactions with the negatively charged sulphate groups of GAGs (Figure 11).

A further example is that of mouse mast cell protease 6 (MCP-6). Molecular modelling of MCP-6 identified four conserved, pH dependent and surface exposed histidine residues (His 35, His 106, His 108 and His 238) that mediate the interaction of the protein with heparin in a pH dependent fashion (139). The electropositive nature of the surface of the protease, as shown in Figure 11, is due to the presence of protonated histidines that can make favourable interactions with GAGs, as compared with the accessible surface in the presence of deprotonated histidines. Histidine proline-rich glycoprotein (HPRG) is another example wherein binding to heparin is minimal at neutral pH but increases rapidly to a maximum at pH 6.5 (140). At an intermediate pH, both the protonation of histidines and the binding of zinc promote the interaction of HPRG with heparin. It is probable that there is a pH range where all histidines will be protonated, whereas most, if not all, of the glutamic and aspartic acid residues will still be negatively charged. This is likely to be the most favourable situation for heparin binding.

Effect of Metal Ions on GAG Binding

Sulphated GAG chains also bind strongly to divalent metal ions present in proteins or in solution (87). Molecular modelling and NMR studies have indicated the binding preferences of heparin for Ca^{2+} in solution (141,142) and is similar to the co-ordination observed in heparin-metal-protein complexes such as annexin proteins (92,143). The carboxylate groups of the iduronate residue and the *N*-sulphate and 6-*O*-sulphate of GlcNS6S are essential for Ca^{2+} as compared with 2-*O*-sulphate of IdoA2S. Combined modelling and NMR studies have indicated that the heparin-Ca^{2+} binding site has preference for the $^{1}C_{4}$ iduronate form.

Divalent cation binding may be expected to influence the specificity and affinity of GAG–protein interactions. The binding of heparin/HS to proteins is enhanced in the presence of divalent cations such as

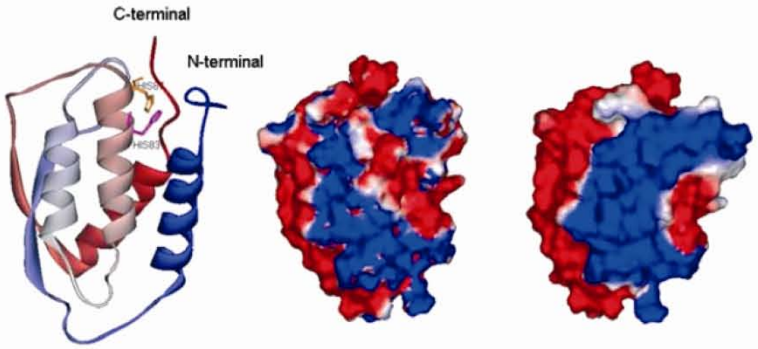


Figure 11: Crystal structure of recombinant human GM-CSF (PDB code 1CSG). Histidine residues (His^{83} and His^{87}) in helix C of hGM-CSF appear to act as a pH-dependent molecular switch to control the interaction with GAGs (202). The electrostatic potential surface is shown for both neutral (deprotonated His residues, positive charge contributed by Lys/Arg residues) and acidic pH (protonated His residues).

zinc, as is the case of endostatin (144). Crystallographic studies of the complexes of heparin-derived oligosaccharides with human annexin A2 suggest that this protein exhibits significant Ca^{2+}-dependent heparin-binding properties (Figure 12) at pH 7.4, either as a monomeric protein or as a component of an A2t heterotetramer (143). In the complex of a heparin oligosaccharide with annexin V the calcium cation does not interact directly with the heparin fragment but it induces the conformation of protein loops necessary for binding (92). Glycosaminoglycans also bind to prion proteins (PrP) at pH values above the pKa of histidine and in a metal ion-dependent fashion (145). Prion protein–GAG complexes are stabilized by Cu^{2+}or $Zn^{2+}$and PrP–GAG interactions are mediated largely by protonated and Cu(II)-bound His side-chains present at the N-terminal domain of PrP. Divalent cations were not found to be a prerequisite for the interaction of GAGs with lipoproteins but were found to stabilize their resulting complexes. It has been observed that Mn^{2+} is better than Mg^{2+} or Ca^{2+}at promoting stronger binding between the acidic groups of heparin and the phospholipid portion of low density lipoproteins (146).

It is known that Zn^{2+} binds selectively to heparin rather than to other GAGs (147), which suggests that binding of divalent cations to GAG chains is not always a simple electrostatic interaction between the negatively charged groups on the carbohydrate and the positively charged metal ion. NMR studies have revealed that iduronic acid is the main binding site in heparin for divalent cations. Spectral data also revealed that Zn^{2+} binding controls the ring conformation of iduronate in heparin and HS, with the $^1C_4$ ring conformation being stabilized over the $^2S_0$ conformation (148,149).

Molecular Modelling of GAGS and GAG–Protein Interactions

In view of the limited structural knowledge available on GAG–protein interactions and the phenomenal structural diversity of heparin and HS, molecular modelling approaches have assisted the understanding of GAG binding affinity and specificity. Glycosaminoglycans are challenging from a molecular modelling perspective

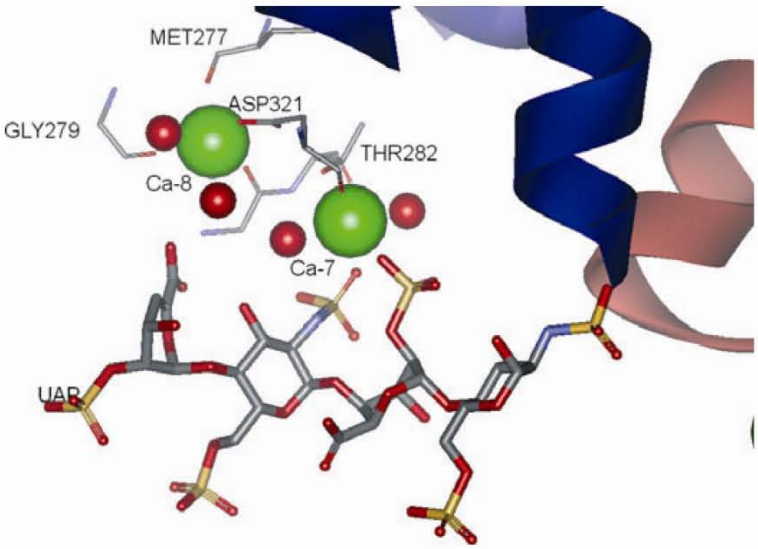


Figure 12: Calcium co-ordination at the heparin binding site in the crystal structure of annexin A2 (PDB code 2HYU). The Ca^{2+} ions are shown as green spheres, the water molecules as red spheres and the heparin tetrasaccharides as sticks.

because of their high negative charge density and their conformational flexibility. Protein side chains also have a high degree of conformational flexibility. Hence, if all possible conformations of the sulphate and hydroxyl groups on a GAG oligosaccharide and all rotamers of charged side chains in a protein are to be taken into account, an accurate prediction of GAG–protein binding becomes an extremely challenging task.

Several molecular modelling techniques have been described in the literature for the successful prediction of sulphated GAG binding sites on the surface of proteins and for the prediction of their relative affinities. These methods include energy mapping of ligand probes on the surface of proteins, molecular docking and scoring, and molecular dynamics (MD) simulations (150). The prediction of binding sites for GAGs can also be validated upon survey of bound sulphates originating from the crystallization buffer found in the crystal structures of proteins (137,151).

Prediction of GAG binding sites on protein surfaces using GRID

The prediction of the location of GAG binding sites on the surface of proteins has been attempted by searching for the most positively charged patches of amino acids. The GRID algorithm (152), using atom probes to represent polar or charged groups on saccharide molecules, has been used successfully to map the most energetically favourable positions where sulphate groups may bind to the surface of proteins. Such studies have been performed with a number of proteins such as aFGF, bFGF, AT and IL-8 (153). The mapping of sulphate interaction energies can be first computed using GRID and then followed by ligand–protein docking to predict the most favourable anchoring position for a charged sulphate group on the surface of proteins. Different binding modes were proposed in this way for the interactions of HS with the chemokines RANTES, MIP-1α, and chemokine domain of fractalkine, showing that the different types of interactions that may arise on the surface of proteins are determined by their three-dimensional structures (151). However, this study did not consider an analysis of the optimum GAG sequence required for binding (i.e. effect of *N* and *O*-sulphation).

Molecular docking

The purpose of docking GAG fragments to the surface of a protein is to identify the likely position of its heparin-binding site(s), predict the binding mode of GAG fragments and obtain an estimate of the free energy of binding (and dissociation constant). Most docking studies reported so far for heparin-binding proteins have focussed on predicting the amino acids that make up the heparin binding sites, but there have been few reports on the calculation of binding affinities (137,154).

Simulated annealing and genetic algorithms have been extensively used to dock GAGs to their putative proteins or receptors (137,144,154–158). It is clear that the prediction of binding energies of heparin and related sulphated GAGs to their biological targets requires a large number of docking evaluations in order to achieve energy convergence and sufficient conformational sampling.

Several ligand–protein docking studies have been reported for the prediction of heparin-binding sites on AT, aFGF and bFGF, which have been contrasted with the X-ray crystal structures of the complexes of these proteins with heparin oligosaccharide fragments (153). After correctly predicting the location of heparin binding sites in these proteins, docking simulations were also used to predict which are the key residues that make up the heparin binding site of IL-8 (153). Molecular modelling studies have also been carried out to predict the binding of a heparin hexasaccharide to the multicomponent complex between bFGF and FGFR1, with findings consistent with experimental data on the binding mechanism of bFGF to its receptor, the dimerization of the receptor, and site-specific mutagenesis and biochemical cross-linking data (159). Molecular docking has also been used to predict that long heparin fragments such as a dodecasaccharide or a tetradecasaccharide are required for binding to the dimer of chemokine SDF-1α (160). In another study, different protein models for the dimer of MIP-1α were built on the basis of the crystal structures of PF4 and IL-8 (67). Docking simulations using heparin penta- and endecasaccharides predicted the interaction of these GAGs with the S-domains (requiring fragments 12–14 saccharides long) and the electropositive surface on the opposite face of the MIP-1α dimer (67).

A study of the interaction between a heparin pentasaccharide and AT has been carried out, despite the difficulty posed by the known conformational change that occurs in the protein upon ligand binding. Homology modelling of the protein structure and manual docking of the pentasaccharide were used to determine the basic amino acids involved in the recognition of the sulphate and carboxylate groups of the oligosaccharide. These predictions were then confirmed by automated docking simulations (161). The crystal structure of the complex between AT and the pentasaccharide revealed the existence of contacts between heparin and arginine and lysine residues on three different helices in the protein (93). The crystal structures of ternary complexes of AT, thrombin and heparin and AT and factor Xa and heparin, provided further insights into the large conformational changes that occur in AT upon activation.

Docking simulations have been used to predict the binding mode of a heparin oligosaccharide onto the surface of endostatin (144), as well as to determine the binding mode of a hexasaccharide to aFGF (162). In the case of aFGF, most of the low energy docking poses of a hexasaccharide were oriented towards Lys127 and Lys142 on the surface of the protein.

Docking methods have also been used to screen a combinatorial virtual library of heparin/HS hexasaccharides against the crystal structure of AT, identifying high specificity heparin/HS sequences (163). A good correlation was observed between the GOLD docking scores and the experimental binding affinity. Different methods have been used to dock heparin fragments onto activated protein C (APC), supported by experimental data (164). Short heparin oligosaccharides were determined to bind to various loops in APC, impairing the interaction of APC with factor Va during APC-catalysed cleavage.

Current docking methods aimed at predicting high affinity GAG sequences generally fail to take into account any conformational changes that may occur in the protein receptor. In addition, GAG

oligosaccharides have many rotatable bonds (large degrees of freedom), posing a significant challenge for the search of the correct binding mode, particularly for molecules larger than a pentasaccharide. An additional problem arises due to the presence of distant, discontinuous heparin binding sites. Because most docking methods perform coarse docking, two model oligosaccharide fragments are needed for improved accuracy: one in which all IdoA2S residues are in the $^{1}C_{4}$ ring conformation and another one in which they adopt the $^{2}S_{0}$ ring conformation. Some of these limitations are being overcome by recent docking and scoring methods, such as those implemented in Glide (Schrödinger LLC) and AutoDock 4.0, which take into account partial flexibility of the protein and allow full treatment of ligand flexibility. Monte Carlo Multiple Minima (MCMM) calculations have also been used to try to overcome the problem of ligand flexibility (165) in the search for binding modes of cyclitols (GAG-like sulphated molecules) in complex with aFGF (166).

Recent docking simulations of GAG fragments to the homology model of PECAM-1 considered the sugars in different $^{1}C_{4}$, $^{4}C_{1}$ and $^{2}S_{0}$ conformations (137,138). Glycosaminoglycan binding sites were initially predicted on the basis of a survey of sulphate groups in known crystal structures of proteins followed by docking of heparin fragments of various size and ring conformations (Figure 13). AutoDock scores gave a good correlation with experimental data suggesting the existence of high and low affinity GAG binding sites in PECAM-1. Docking simulations also predicted the effect of pH on binding of GAGs due to the presence of a key protonated histidine residue in Ig-domains 2 and 3. When these calculations were repeated at neutral pH, the free energy of binding increased to approximately 3 kcal/mole because of the loss of ionic interactions. The free energy of binding was determined to decrease with the increase in the size of the heparin fragments, with the optimum size being a pentasaccharide for an interaction with the 'closed' conformation of the receptor (137,138).

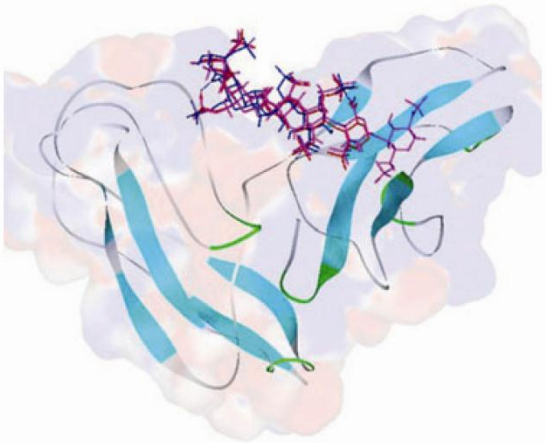


Figure 13: Heparin fragments of varying length docked onto extracellular domains 2 and 3 of PECAM-1. The protein domains are represented as ribbons and the heparin fragments are shown in sticks.

Energy scoring of GAG–protein interactions

The accurate computational prediction of the affinity of binding of GAG–protein complexes is still in its infancy. This is mostly due to the poorly defined contribution of water (solvation/desolvation) to the binding interaction and current limitations in the force fields (such as in the parameterization of charges and bonded terms and the neglect of polarization effects) and scoring functions used to represent GAG structure, dynamics and interactions.

Specific scoring functions have been developed for ranking the binding modes of non-GAG carbohydrates to proteins (167,168) which can take into account CH···π interactions, hydrogen bonds and electrostatic interactions. These functions can also be used for the sugar backbone present in sulphated GAGs but have not been optimized to deal with anionic substituents, such as sulphates and carboxylic acid groups. The GScore and EScore functions implemented in Glide (Schrödinger) have been used successfully for the screening of interactions of GAGs and their mimetics with FGFs, with good correlation with experimental binding affinities (169,170). Both scoring functions performed the same for the prediction of binding modes and the orientation of sulphate groups; however, Escore was better at predicting binding affinities.

The biomolecular ligand energy evaluation protocol (BLEEP) method has been used successfully to identify low-energy binding modes of heparin fragments (171). Various conformations of heparin were generated and the structure of human bFGF was kept rigid, all in the presence of explicit water molecules. The method correctly assigned the lowest energy to the binding mode observed in the crystal structure, indicating that its potential of mean force score (PMFscore) function is able to measure correctly the interaction energies of GAGs.

The linear interaction energy (LIE) method (172) and MCMM conformational search have been used to predict the binding affinities of tetracyclitols to aFGF and bFGF (166), but met with little success. This is due to the large magnitude of the electrostatic contributions to the free energy, resulting in unreasonable LIE coefficients.

Molecular dynamics simulations

Molecular dynamics simulations have been reported for complexes of oligosaccharides with proteins such as galectin-1 (152) and endo-1,4-b-xylanase II (XynII) (173), but very few MD simulations have been performed for sulphated GAGs such as heparin and HS. A 5 ns MD simulation has been reported for the complex of a heparin disaccharide with IL-8, where the binding affinities of the Ala mutants of basic residues histidines, lysines and arginines were predicted compared to wild type IL-8, and the structural stability of the monomeric and dimeric forms of IL-8 was also determined (68).

Molecular dynamics simulations have also been performed for the complex of a heparin pentasaccharide with AT in order to characterize the energetic contribution of important amino acids required for the interaction with GAG fragments and the ability of GAG fragments to induce the observed conformational change in AT (174). These simulations revealed that there is no specific conformational requirement for IdoA2S, as either the skew-boat or chair conformation is appropriate for binding with AT with a similar enthalpy of interaction.

A number of MD simulations have been carried out for heparin fragments in aqueous solution. Simulations of a system comprising a heparin decasaccharide in water with explicit sodium ions determined the conformation of heparin in solution under physiological conditions, which was found to be in agreement with NMR data (175). These simulations investigated the conformational changes in iduronic acid and the conformational flexibility of the glycosidic linkage. Larger variability in the conformation of heparin with respect to NMR-determined structures was observed, although this may have been due to the choice of partial atomic charges. Molecular dynamics simulations suggested that chair forms predominate monosaccharide level of IdoA2S (78) and the skew-boat may contribute approximately 40–60% of the total IdoA2S conformational preference in the entire polysaccharide chain depending on the heparin sequence (78). More recently, MD simulations of IdoA2S containing oligosaccharides has been reported to calculate the forces responsible for conformational preference of IdoA2S residue in aqueous solution and the results indicated that stabilization due to intramolecular hydrogen bonds around IdoA2S is highly correlated with the expected conformational equilibrium for this residue in solution (i.e. $^{2}S_{0}$ conformation) (176).

Other simulations have investigated the binding of divalent metal ions such as Ca^{2+} and the effect of the sulphation pattern of heparin in aqueous solution (141,142). These simulations revealed that IdoA2S residues adopt the $^{1}C_{4}$ conformation when co-ordinating the metal. The co-ordination shell of Ca^{2+} was made up of the *N*-sulphate of GlcNS6S and the carboxylate of IdoA2S, whereas the sulphate at position 2 of IdoA2S is not essential for binding to the metal.

The force constants used in molecular mechanics force fields to characterize equilibrium bond lengths and angles, as well as partial charges and VDW interaction parameters can significantly affect the accuracy of simulations of ligand–protein interactions. A variety of force fields have been designed for the simulation of carbohydrates, including optimized potential for liquid simulations (177), GROningen MOlecular Simulation (GROMOS package) (178), carbohydrate solution force field (179), CHARMM (180), CHARMM CHEAT95 (181), Glycam/AMBER (182), MM2 (183) and MM3 (184), PEF95SAC (185) and PIM (set of carbohydrate parameters) (186). These force fields do not always contain parameters for sulphated carbohydrates such as GAGs, but various approaches can be followed to develop specific parameters for GAGs using the MM2 (187,188), CHARMm and AMBER force fields (189). Parameters (i.e. non-bonded) for sulphates not available from the work of Huige and Altona (189) can be approximated from those for phosphates available from AMBER or CHARMm or Glycam06 (190) or from Lan Jin's thesis (191).

Therapeutic Potential of GAG Molecules and GAG Mimetics

Carbohydrates such as GAGs derive their biological activity through binding to their protein receptors. These carbohydrate–protein interactions may be mimicked by designing small molecule drugs with appropriate binding affinity and selectivity. However, a significant problem is that the binding affinities of many carbohydrate–protein interactions are in the milli- to micromolar range, whereas small molecule drugs tend to require nanomolar binding affinities. Consequently, synthetic compounds that have been specifically designed to mimic the structure and interactions of carbohydrates, such as GAG mimetics, need to bind to their receptors with higher affinity than naturally occurring GAG oligosaccharides. Potential strategies based on heparin/HS–protein interactions have recently been described to assist GAG-based drug discovery (192). Glycosaminoglycan-based drugs can act in several ways by activating (agonists) or inactivating (antagonists) protein-based receptors, competing with endogenous GAGs and/or inhibit GAG biosynthesis.

The molecular diversity of heparin/HS interactions has been exploited for the development and clinical progression of GAG mimetics (193). Discrete GAG sequences can bind specifically and make unique interactions with a large number of proteins, such as chemokines (194), growth factors (195), proteases [e.g. AT (93)] and adhesion molecules (64). Nevertheless, the design of GAG mimetics requires an understanding of the mechanism and specificity of a given GAG–protein interaction. One mechanism of interaction of cell surface GAGs involves binding to their receptor through the formation of a complex involving the GAG, ligand and receptor molecules. In this case, the GAG and receptor binding sites are spatially distinct and uncompetitive for binding, resulting in the formation of ternary complexes, as in the case of IFN-γ, PF4, IL-3, G-CSF and GM-CSF (196). The other mechanism of interaction of soluble GAGs involves the binding competition with ligands, such as the chemokine IL-8 (197) and the helical cytokine IL-5, to their high affinity receptor (196). Cell surface GAGs help present chemokines to their GPCRs (G-Protein Coupled Receptor) by increasing the local concentration of protein (122).

Very few GAG fragments have been developed for therapeutic use (Figure 14), mostly because the synthesis of such fragments is difficult. The synthetic challenges associated with the complex structures of these oligosaccharides arise from the low availability of L-idose and L-iduronic acid from commercial or natural sources and the lack of efficient synthetic routes to make sufficient amounts of these monosaccharides. Other difficulties involve the development of a suitable protecting-group strategy to allow the implementation of a high degree of functionalization of heparin/HS fragments and the stereo-selective and efficient formation of interglycosidic bonds in the carbohydrate backbone (198).

The most recognized pharmaceutical application of GAGs is in anticoagulation. Many pharmaceutical companies, such as Organon and Sanofi-Aventis, have been working on the development of commercial GAG-based drugs that can bind AT and thereby cause anticoagulation efficiently without the need for frequent administration compared to full length heparin. For example, the synthetic pentasaccharide Arixtra® (Fondaparinux or SR90107/Org31540) (199) binds to AT and has better efficacy at low doses (with a half life of 17 h).

The crystal structure of the complex of Arixtra® with AT in the absence of coagulation factors has been determined (200). The crystal structure of this synthetic pentasaccharide complexed to thrombospondin-1 (TSPN-1) has also been resolved (201). The structural requirements for the binding of heparin to AT, as shown in

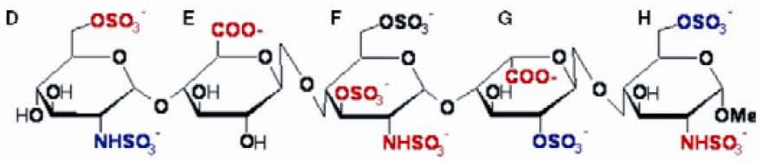


Figure 14: Structural requirement for heparin binding to AT based on the structure-activity relationships of Fondaparinux. The groups highlighted in red are absolutely essential for the activation of AT, whereas the groups in blue only help to increase the biological activity. The 3-*O*-sulphate group at position F is responsible for strong binding to AT-III. The 3-*O*-sulphate group at position F is also specific for binding to AT. The *N*-sulphated groups in Fondaparinux are replaced by *O*-methyl groups to form the more potent pentasaccharide Idraparinux.

Figure 14, were determined on the basis of the crystal structure and the structure-activity relationships for a series of pentasaccharides, using various combinations of sulphate and carboxylate groups (127,202). The 3-*O*-sulphate group at position F of Fondaparinux was shown to exhibit the strongest binding to AT by interacting with positively charged amino acids, whereas the lack of the 3-*O*-sulphate group at position F results in a decreased binding affinity to AT of nearly 20,000-fold.

Several other clinical candidates have been developed, such as Idraparinux (SANORG 34006), which also selectively inhibits coagulation factor Xa and binds to AT (203). Idraparinux is currently in phase III clinical trials for the treatment of venous thromboembolism. The synthesis of Idraparinux is much easier than that of Fondaparinux or heparin. It has a higher affinity (K_d of 1 nM) and better efficacy than Fondaparinux (K_d of 25 nM). The hydroxyl groups in Idraparinux are methylated and the *N*-sulphate groups in Fondaparinux are replaced by *O*-sulphates in Idraparinux. Idraparinux has an increased half life (120 h) in the bloodstream. The higher activity of Idraparinux appears to be due to the presence of its methyl ethers and of IdoA in the favourable $^2S_0$ conformation (202). The crystal structure of the pentasaccharide analogue of idraparinux complexed with activated AT has been reported (93), providing convincing evidence that the IdoA adopts the $^2S_0$ conformation.

Heparin binds to both AT and thrombin simultaneously to form a ternary complex, and it also binds to and inhibits factor Xa. Significantly longer heparin oligosaccharides are required to inhibit thrombin activity compared to the specific pentasaccharide that is required to bind to AT and inhibit factor Xa. The synthetic hexadecasaccharide SR123781 has tailor-made factor Xa and thrombin inhibitory activities combined with more selectivity in their mode of action. The likely molecular interactions of this hexadecasaccharide have been determined from the X-ray crystal structures of ternary complexes of AT/thrombin/heparin (204). This oligosaccharide consists of an AT-binding domain (ABD) (S12–S16) at the reducing end of the non-sulphated linker, a non-sulphated linker region (S6–S11), and a thrombin-binding domain (TBD) (S1–S5) at the non-reducing end of the linker (Figure 15). It also contains methoxy groups MeO and 2-*O*-sulpho substituted glucose units in the AT binding domain instead of the *N*-sulpho substituted glucosamine (S12–S16) that occurs in the natural heparin pentasaccharide. The highly sulphated glucose units allow non-specific binding to thrombin, which is dependent primarily on the overall charge density of the GAG fragment rather than on the precise sequence of the variously substituted sugar residues. The non-sulphated linker region

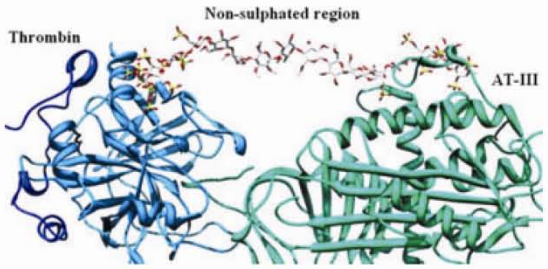


Figure 15: Structure of SR123781A and ribbon representation of the ternary complex. Thrombin is blue and anti-thrombin is green and the 16-mer heparin is in stick form.

(S6–S11) does not interact with any protein residues, but rather it facilitates the formation of the ternary complex, giving rise to increased anti-thrombin activity but with minimal interaction with PF4. It was reasoned that an oligosaccharide in which charged ABD and TBD are separated by a non-sulphated neutral linker could deliver an AT-mediated factor Xa inhibitor and block the active site of clot-bound thrombin, hence not giving undesired interactions with PF4.

PI-88 (Progen, Toowong, Queensland, Australia) (Figure 16) has progressed to clinical trials to treat inflammatory diseases, thrombosis, virus infections and cancer (205). PI-88 acts as a substrate analogue to inhibit heparanase to prevent HS degradation. This enzyme is the therapeutic target in disease states such as tumour cell invasion, metastasis, and angiogenesis. PI-88 is a phosphomannopentose sulphate (6-*O*-PO_3H_2-α-D-Man-(1 → 3)-α-D-Man-(1 → 3)-α-D-Man-(1 → 3)-α-D-Man-(1 → 2)-D-Man), wherein the chain length, sugar composition and glycosidic linkages α(1 → 3) and α(1 → 2) play important roles in its anti-coagulation activity compared with the anti-coagulant activity of sulphated glucose-containing oligosaccharides with β(1 → 4) and β(1 → 3) linkages (206). A number of sulphated pseudo sugar molecules such as cyclitols have been identified as selective inhibitors of certain protein–heparin interactions (207). In general, so called GAG mimetics consisting of different sugar backbones (such as mannose, triose, Xyl or inositols) with different degrees of sulphation and coupled by linkers of variable chain length, flexibility, orientation, and hydrophobicity can help to probe differences in the heparin/HS binding specificity and selectivity of proteins.

The chemical synthesis of carbohydrate-based oligosaccharides is plagued with problems of poor yield, multitude of products and long

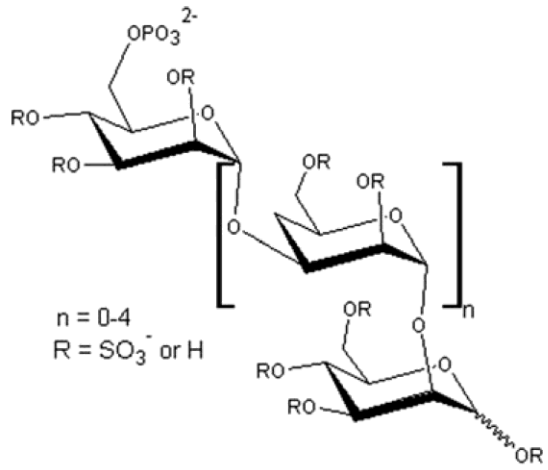


Figure 16: Structure of phosphomannopentose sulphate (PI-88).

reaction times. Microwave-based synthesis of variably functionalized, per-sulphated organic molecules has been reported to result in high yields and purity to facilitate the rapid screening of these molecules as GAG mimetics (208). An alternative approach (i.e. peptide synthesis), in which homo-oligomers of tyrosine instead of carbohydrates are used as the backbone, has recently been reported to overcome the limitation of oligosaccharide synthesis (209). These oligo (tyrosine sulphate) molecules bind more strongly to AT than a heparin-derived hexasaccharide, with the binding strength being dependent on chain length (209).

Anti-coagulants comprise direct and indirect inhibitors of enzymes involved in the coagulation pathways, primarily thrombin and factor Xa (210). Direct inhibitors (including thrombin inhibitors, e.g. Lepirudin, Argatroban, Bivalirudin and factor Xa inhibitors, e.g. Rivaroxaban) (211) interact with either the active site or an exosite of the pro-coagulant enzyme, blocking its proteinase activity. Indirect inhibitors enhance the proteinase inhibitory activity of the natural anticoagulants, AT and heparin co-factor II (e.g. LMWH). Significant efforts have been made to design GAG mimetics as inhbitors of direct and indirect mechanisms (210,211). A carboxylic acid-based polymer, poly-acrylic acid (PAA), demonstrated a surprisingly high acceleration in thrombin inhibition through a 'bridging' mechanism of activation (212). Further exploration of this scaffold led to the design of 4-hydroxycinnamic acid-based dehydrogenation polymer DHP (dehydropolymers) oligomers with dual inhibitory action against thrombin and factor Xa (213). Although high activation is achieved with these carboxylate-only molecules, SAR (structure activity relationship) studies show the need to retain critical sulphate groups and a more rigid backbone when designing GAG mimetics (or activators) in order to retain interactions in the ABD.

Molecular modelling has been used to design a small non-sugar based AT activator, epicatechin sulphate (ECS), from a pharmacophore deduced from the DEF portion of the natural heparin pentasaccharide DEFGH (214). Epicatechin sulphate was rationally designed using the hydropathic interaction technique (HINT) (215,216) to target the binding and activation of AT for enhanced inhibition of factor Xa in a pH- and salt-dependent manner. However, +/−catechin sulphate, a chiral stereoisomer of ECS, did not bind in the pentasaccharide binding site of AT, resulting in only weak activation (217). In fact, catechin sulphate was found to bind in the extended heparin binding site, which is adjacent to the binding domain for the reference trisaccharide DEF (217). Similarly, HINT analysis was performed for the PAA–AT complex, indicating that residues Arg 13, Arg 47, Lys 125, Arg 129, Arg 132, Lys 133 and Lys 136 are required for the interaction with PAA (212), as opposed to Lys 114, which is a key residue for the interaction with the critical 3-sulphonate group of residue F in the heparin pentasaccharide (113).

Several molecules capable of blocking heparin binding to growth factors such as VEGF (Vascular Endothelial Growth Factor) and bFGF have been identified (218). Most of these molecules have a linear extended structure containing heparin mimetic functional groups, such as carbohydrates, sulfonates, carboxylates and hydroxyl groups (218). This is important as negative charges on functional groups in GAG molecules, such as *N*-sulphates and *O*-sulphates, have been found to play an important role in the interactions with growth factors. Multiple sulphated peptides have been also reported to bind to a set of heparin-binding peptides and $VEGF_{165}$ (219).

The entry of viruses into cells can be effectively blocked by either the removal of HS chains from the cell surface through enzymatic treatment or the presence of soluble forms of HS or HS-like molecules which competitively bind to the virus. Lignin sulphate, a sulphated form of lignin, was identified as an inhibitor of HSV-1 entry into cells (220). Recently, Surfen (*bis*-2-methyl-4-amino-quinolyl-6-carbamide) has been identified as small molecule antagonist of HS, as cell attachment and infection by HSV was diminished in the presence of this molecule (221). Furthermore, Surfen is also known to inhibit FGF signalling by blocking the formation of the FGF/HSPG complex (221).

A variety of different approaches such as solution-phase and solid-phase chemistry have been used for the polymer-supported synthesis of GAG and non-GAG derivatives in order to develop a large variety of oligosaccharides (198,207). Heparin microarray technology has provided the tools for rapid screening of specific GAG sequences (natural or synthetic oligosaccharides) that interact with proteins such as chemokines (222) and signalling proteins involved in inflammation, viral infection and immune system regulation (223–225). The introduction of non-anionic structural motifs into heparin/HS (226,227) may also provide a route for the development of novel, potent drug-like GAG mimetic molecules that can treat various diseases.

Summary

The use of biochemical, structural and molecular modelling methods to study the structure and function of GAGs has allowed significant progress in the description of the conformational properties of these complex molecules, the elucidation of the structural determinants of their interactions with proteins and the development of GAG mimetic molecules of therapeutic value.

The number of X-ray determinations of crystal structures of diverse GAG–protein complexes has increased in recent years. This has made possible to obtain a broader characterization of the properties of heparin binding regions in proteins, the preferred conformations and sulphation patterns that heparin has when it interact with proteins, the role of metals in mediating these interactions and the modifying effect of pH on the affinity of binding. The successful design and development of GAG-mimetic drug molecules depends on understanding how these factors shape the relationship between GAG structure and the nature of the interactions of GAGs with different proteins.

Molecular modelling techniques have enabled substantial progress to be made on the prediction of heparin binding sites in proteins, the prediction of biologically active GAG conformations, the dissection of the different intermolecular forces that dictate GAG–protein binding affinity and the rationalization of the specificity and selectivity of GAG interactions with proteins. These methods, in combination with new synthetic approaches and the use of fast screening tools, are enabling the successful development of new potent drug molecules.

Outlook

A number of challenges lie ahead in the investigation of the interactions of GAGs with proteins from a structural and functional point of view. For example, a number of proteins of immunological importance (such as cytokines and chemokines) have the ability to oligomerize and so do their receptors (228–231). Heparin is known to interact with many of these proteins but the precise mechanism of interaction is not known. Structural biology determinations and molecular modelling approaches will be required to determine whether heparin and receptor binding sites occupy different positions on the surfaces of these proteins and to elucidate the possible role of heparin in mediating the oligomerization of these immunological molecules and/or their interactions with their receptors.

The realistic description of the conformational flexibility of proteins and GAGs and its effect on heparin binding, and the incorporation of the effect of the aqueous solvent in mediating heparin–protein interactions are two important areas where more molecular modelling studies are clearly needed. From a methodological point of view, one of the main challenges facing molecular modelling and computational chemistry is the accurate representation of the polarization effects in GAG molecules because of their high charge densities, which are likely to affect the computational prediction of their affinity of binding to proteins.

Acknowledgments

NG gratefully acknowledges the award of an Endeavour International Postgraduate Research Scholarship from the Australian Government. The authors acknowledge Dr Rupesh Khunt at Vienna University of Technology for his help with the molecular drawings.

References

1. Rademacher T.W., Parekh R.B., Dwek R.A. (1988) Glycobiolngy. Annu Rev Biochem;57:785–838.
2. Jackson R.L., Busch S.J., Cardin A.D. (1991) Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes. Physiol Rev;71:481–539.
3. Casu B., Lindahl U. (2001) Structure and biological interactions of heparin and heparan sulfate. Adv Carbohydr Chem Biochem;57:159–206.
4. Conrad H.E. (1998) Heparin-Binding Proteins. San Diego, CA: Academic Press.
5. Champe P.C., Harvey R.A. (1994) Biochemistry (Lippincott's Illustrated Reviews). Philadelphia, PA: Lippincott-Raven.
6. Sasisekharan R., Venkataraman G. (2000) Heparin and heparan sulfate: biosynthesis, structure and function. Curr Opin Chem Biol;4:626–631.
7. Capila I., Linhardt R.J. (2002) Heparin–protein interactions. Angew Chem Int Ed Engl;41:390–412.
8. Sharon N. (1986) IUPAC–IUB Joint Commission on Biochemical Nomenclature (JCBN). Nomenclature of glycoproteins, glycopeptides and peptidoglycans. Recommendations 1985. Eur J Biochem;159:1–6.
9. Lawrence R., Lu H., Rosenberg R.D., Esko J.D., Zhang L. (2008) Disaccharide structure code for the easy representation of constituent oligosaccharides from glycosaminoglycans. Nat Methods;5:291–292.
10. Bohne-Lang A., Lang E., Förster T., von der Lieth C.-W. (2001) LINUCS: Linear Notation for Unique description of Carbohydrate Sequences. Carbohydr Res;336:1–11.
11. Banin E., Neuberger Y., Altshuler Y., Halevi A., Inbar O., Dotan N. *et al.* (2002) A novel linear code nomenclature for complex carbohydrates. Trends Glycosci Glycotechnol;14:127–137.
12. Itano N. (2008) Simple primary structure, complex turnover regulation and multiple roles of hyaluronan. J Biochem;144:131–137.
13. Toole B.P. (1997) Hyaluronan in morphogenesis. J Intern Med;242:35–40.
14. Toole B.P., Wight T.N., Tammi M.I. (2002) Hyaluronan–cell interactions in cancer and vascular disease. J Biol Chem;277:4593–4596.
15. Chen W.Y., Abatangelo G. (1999) Functions of hyaluronan in wound repair. Wound Repair Regen;7:79–89.
16. Laurent T.C., Laurent U.B., Fraser J.R. (1996) The structure and function of hyaluronan: an overview. Immunol Cell Biol;74:A1–A7.
17. Stern R. (2008) Association between cancer and ``acid mucopolysaccharides'': an old concept comes of age, finally. Semin Cancer Biol;18:238–243.
18. Aruffo A., Stamenkovic I., Melnick M., Underhill C.B., Seed B. (1990) CD44 is the principal cell surface receptor for hyaluronate. Cell;61:1303–1313.
19. Heldin P., Karousou E., Bernert B., Porsch H., Nishitsuka K., Skandalis S.S. (2008) Importance of hyaluronan-CD44 interactions in inflammation and tumorigenesis. Connect Tissue Res;49:215–218.
20. Turley E.A., Noble P.W., Bourguignon L.Y.W. (2002) Signaling properties of hyaluronan receptors. J Biol Chem;277:4589–4592.

21. Lehninger A.L., Nelson D.L., Cox M.M. (2004) Principles of Biochemistry 4e. New York: W.H. Freeman & Company;p. 238–271.
22. Linhardt R.J., Gunay N.S. (1999) Production and chemical processing of low molecular weight heparins. Semin Thromb Hemost;25:5–16.
23. Wardrop D., Keeling D. (2008) The story of the discovery of heparin and warfarin. Br J Haematol;141:757–763.
24. Marcum J.A. (2000) The origin of the dispute over the discovery of heparin. J Hist Med Allied Sci;55:37–66.
25. Iozzo R.V., San Antonio J.D. (2001) Heparan sulfate proteoglycans: heavy hitters in the angiogenesis arena. J Clin Invest;108:349–355.
26. Holt C.E., Dickson B.J. (2005) Sugar codes for axons? Neuron;46:169–172.
27. Liu D., Shriver Z., Qi Y., Venkataraman G., Sasisekharan R. (2002) Dynamic regulation of tumor growth and metastasis by heparan sulfate glycosaminoglycans. Semin Thromb Hemost;28:67–78.
28. Tímár J., Lapis K., Dudás J., Sebestyén A., Kopper L., Kovalszky I. (2002) Proteoglycans and tumor progression: Janus-faced molecules with contradictory functions in cancer. Semin Cancer Biol;12:173–186.
29. Sanderson R.D. (2001) Heparan sulfate proteoglycans in invasion and metastasis. Semin Cell Dev Biol;12:89–98.
30. Casu B., Guerrini M., Torri G. (2004) Structural and conformational aspects of the anticoagulant and anti-thrombotic activity of heparin and dermatan sulfate. Curr Pharm Des;10:939–949.
31. Fareed J., Hoppensteadt D.A., Bick R.L. (2000) An update on heparins at the beginning of the new millennium. Semin Thromb Hemost;26:5–18.
32. Sasisekharan R., Shriver Z., Venkataraman G., Narayanasami U. (2002) Roles of heparan–sulphate glycosaminoglycans in cancer. Nat Rev Cancer;2:521–528.
33. Yip G.W., Smollich M., Gotte M. (2006) Therapeutic value of glycosaminoglycans in cancer. Mol Cancer Ther;5:2139–2148.
34. Bandtlow C.E., Zimmermann D.R. (2000) Proteoglycans in the developing brain: new conceptual insights for old proteins. Physiol Rev;80:1267–1290.
35. Kisilevsky R., Ancsin J.B., Szarek W.A., Petanceska S. (2007) Heparan sulfate as a therapeutic target in amyloidogenesis: prospects and possible complications. Amyloid;14:21–32.
36. Snow A.D., Willmer J., Kisilevsky R. (1987) Sulfated glycosaminoglycans: a common constituent of all amyloids? Lab Invest;56:120–123.
37. Ley K., Laudanna C., Cybulsky M.I., Nourshargh S. (2007) Getting to the site of inflammation: the leukocyte adhesion cascade updated. Nat Rev Immunol;7:678–689.
38. Fry E.E., Lea S.M., Jackson T., Newman J.W.I., Ellard F.M., Blakemore W.E. *et al.* (1999) The structure and function of a foot-and-mouth disease virus–oligosaccharide receptor complex. EMBO J;18:543–554.
39. Rostand K.S., Esko J.D. (1997) Microbial adherence to and invasion through proteoglycans. Infect Immun;65:1–8.
40. Rider C.C. (1997) The potential for heparin and its derivatives in the therapy and prevention of HIV-1 infection. Glycoconj J;14:639–642.
41. Rider C.C., Coombe D.R., Harrop H.A., Hounsell E.F., Bauer C., Feeney J. *et al.* (1994) Anti-HIV-1 activity of chemically modified heparins: correlation between binding to the V3 loop of gp120 and inhibition of cellular HIV-1 infection in vitro. Biochemistry;33:6974–6980.
42. Spear P.G., Shieh M.T., Herold B.C., WuDunn D., Koshy T.I. (1992) Heparan sulfate glycosaminoglycans as primary cell surface receptors for herpes simplex virus. Adv Exp Med Biol;313:341–353.
43. Shieh M.T., WuDunn D., Montgomery R.I., Esko J.D., Spear P.G. (1992) Cell surface receptors for herpes simplex virus are heparan sulfate proteoglycans. J Cell Biol;116:1273–1281.
44. Spear P.G., Eisenberg R.J., Cohen G.H. (2000) Three classes of cell surface receptors for Alphaherpesvirus entry. Virology;275:1–8.
45. Patricia G.S. (2004) Herpes simplex virus: receptors and ligands for cell entry. Cell Microbiol;6:401–410.
46. Shukla D., Spear P.G. (2001) Herpesviruses and heparan sulfate: an intimate relationship in aid of viral entry. J Clin Invest;108:503–510.
47. Kokenyesi R., Bernfield M. (1994) Core protein structure and sequence determine the site and presence of heparan sulfate and chondroitin sulfate on syndecan-1. J Biol Chem;269:12304–12309.
48. Varki A. (1999) The Essentials of Glycobiology. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
49. Bernfield M., Gotte M., Park P.W., Reizes O., Fitzgerald M.L., Lincecum J. *et al.* (1999) Functions of cell surface heparan sulfate proteoglycans. Annu Rev Biochem;68:729–777.
50. Iozzo R.V., Cohen I.R., Grassel S., Murdoch A.D. (1994) The biology of perlecan: the multifaceted heparan sulphate proteoglycan of basement membranes and pericellular matrices. Biochem J;302:625–639.
51. Nader H.B., Chavante S.F., dos-Santos E.A., Oliveira F.W., de-Paiva J.F., Jerônimo S.M.B. *et al.* (1999) Heparan sulfates and heparins: similar compounds performing the same functions in vertebrates and invertebrates? Braz J Med Biol Res;32:529–538.
52. Warda M., Mao W., Toida T., Linhardt R.J. (2003) Turkey intestine as a commercial source of heparin? Comparative structural studies of intestinal avian and mammalian glycosaminoglycans. Comp Biochem Physiol B Biochem Mol Biol;134:189–197.
53. Ototani N., Kikuchi M., Yosizawa Z. (1981) Comparative studies on the structures of highly active and relatively inactive forms of whale heparin. J Biochem;90:241–246.
54. Warda M., Gouda E.M., Toida T., Chi L., Linhardt R.J. (2003) Isolation and characterization of raw heparin from dromedary intestine: evaluation of a new source of pharmaceutical heparin. Comp Biochem Physiol C Toxicol Pharmacol;136:357–365.
55. Bland C.E., Ginsburg H., Silbert J.E., Metcalfe D.D. (1982) Mouse heparin proteoglycan. Synthesis by mast cell-fibroblast monolayers during lymphocyte-dependent mast cell proliferation. J Biol Chem;257:8661–8666.
56. Linhardt R.J., Ampofo S.A., Fareed J., Hoppensteadt D., Folkman J., Mulliken J.B. (1992) Isolation and characterization of human heparin. Biochemistry;31:12441–12445.

57. Hovingh P., Linker A. (1982) An unusual heparan sulfate isolated from lobsters (*Homarus americanus*). J Biol Chem;257:9840–9844.
58. Dietrich C.P., Paiva J.F., Castro R.A.B., Chavante S.F., Jeske W., Fareed J. *et al.* (1999) Structural features and anticoagulant activities of a novel natural low molecular weight heparin from the shrimp *Penaeus brasiliensis*. Biochim Biophys Acta;1428:273–283.
59. Hovingh P., Linker A. (1993) Glycosaminoglycans in *Anodonta californiensis*, a freshwater mussel. Biol Bull;185:263–276.
60. Pejler G., Danielsson A., Bjork I., Lindahl U., Nader H.B., Dietrich C.P. (1987) Structure and antithrombin-binding properties of heparin isolated from the clams *Anomalocardia brasiliana* and *Tivela mactroides*. J Biol Chem;262:11413–11421.
61. Medeiros G.F., Mendes A., Castro R.A.B., Baú E.C., Nader H.B., Dietrich C.P. (2000) Distribution of sulfated glycosaminoglycans in the animal kingdom: widespread occurrence of heparin-like compounds in invertebrates. Biochim Biophys Acta;1475:287–294.
62. Lindahl U., Kjellen L. (1991) Heparin or heparan sulfate – what is the difference? Thromb Haemost;66:44–48.
63. Lindahl U., Kjellen L. (1991) Heparin or heparan sulfate – what is the difference? Thromb Haemost;66:44–48.
64. Lyon M., Gallagher J.T. (1998) Bio-specific sequences and domains in heparan sulphate and the regulation of cell growth and adhesion. Matrix Biol;17:485–493.
65. Spillmann D., Witt D., Lindahl U. (1998) Defining the interleukin-8-binding domain of heparan sulfate. J Biol Chem;273:15487–15493.
66. Stringer S.E., Gallagher J.T. (1997) Specific binding of the chemokine platelet factor 4 to heparan sulfate. J Biol Chem;272:20508–20514.
67. Stringer S.E., Forster M.J., Mulloy B., Bishop C.R., Graham G.J., Gallagher J.T. (2002) Characterization of the binding site on heparan sulfate for macrophage inflammatory protein 1alpha. Blood;100:1543.
68. Krieger E., Geretti E., Brandner B., Goger B., Wells T.N., Kungl A.J. (2004) A structural and dynamic model for the interaction of interleukin-8 and glycosaminoglycans: support from isothermal fluorescence titrations. Proteins;54:768–775.
69. Lortat-Jacob H., Turnbull J.E., Grimaud J.A. (1995) Molecular organization of the interferon gamma-binding domain in heparan sulphate. Biochem J;310:497–505.
70. Lohse D.L., Linhardt R.J. (1992) Purification and characterization of heparin lyases from *Flavobacterium heparinum*. J Biol Chem;267:24347–24355.
71. Linhardt R.J., Turnbull J.E., Wang H.M., Loganathan D., Gallagher J.T. (1990) Examination of the substrate specificity of heparin and heparan sulfate lyases. Biochemistry;29:2611–2617.
72. Desai U.R., Wang H.M., Linhardt R.J. (1993) Specificity studies on the heparin lyases from *Flavobacterium heparinum*. Biochemistry;32:8140–8145.
73. Wu Z.L., Lech M. (2005) Characterizing the non-reducing end structure of heparan sulfate. J Biol Chem;280:33749–33755.
74. Pikas D.S., Li J.-p., Vlodavsky I., Lindahl U. (1998) Substrate specificity of heparanases from human hepatoma and platelets. J Biol Chem;273:18770–18777.
75. Bishop J.R., Schuksz M., Esko J.D. (2007) Heparan sulphate proteoglycans fine-tune mammalian physiology. Nature;446:1030–1037.
76. Lindahl U., Kusche-Gullberg M., Kjellen L. (1998) Regulated diversity of heparan sulfate. J Biol Chem;273:24979–24982.
77. Ferro D.R., Provasoli A., Ragazzi M., Casu B., Torri G., Bossennec V. *et al.* (1990) Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences. Carbohydr Res;195:157–167.
78. Ferro D.R., Provasoli A., Ragazzi M., Torri G., Casu B., Gatti G. *et al.* (1986) Evidence for conformational equilibrium of the sulfated L-iduronate residue in heparin and in synthetic heparin mono-and oligo-saccharides: NMR and force-field studies. J Am Chem Soc;108:6773–6778.
79. Mulloy B., Forster M.J., Jones C., Davies D.B. (1993) NMR and molecular–modelling studies of the solution conformation of heparin. Biochem J;293:849–858.
80. Desai U.R., Wang H.M., Kelly T.R., Linhardt R.J. (1993) Structure elucidation of a novel acidic tetrasaccharide and hexasaccharide derived from a chemically modified heparin. Carbohydr Res;241:249–259.
81. Yates E.A., Santini F., Guerrini M., Naggi A., Torri G., Casu B. (1996) 1H and 13C NMR spectral assignments of the major sequences of twelve systematically modified heparin derivatives. Carbohydr Res;294:15–27.
82. Ragazzi M., Ferro D.R., Provasoli A., Pumilia P., Cassinari A., Torri G. *et al.* (1993) Conformation of the unsaturated uronic acid residues of glycosamtnoglycan disaccharides. J Carbohydr Chem;12:523–535.
83. van Boeckel C.A., van Aelst S.F., Wagenaars G.N., Mellema J.R., Paulsen H., Peters T. *et al.* (1987) Conformational analysis of synthetic heparin-like oligosaccharides containing {alpha}-L-idopyranosyluronic acid. Rec! Trav Chim Pays Bas;106:19–29.
84. Mikhailov D., Mayo K.H., Vlahov I.R., Toida T., Pervin A., Linhardt R.J. (1996) NMR solution confurmation of heparin-derived tetrasaccharide. Biochem J;318:93–102.
85. Mulloy B., Forster M.J., Jones C., Drake A.F., Johnson E.A., Davies D.B. (1994) The effect of variation of substitution on the solution conformation of heparin: a spectroscopic and molecular modelling study. Carbohydr Res;255:1–26.
86. DiGabriele A.D., Lax I., Chen D.I., Svahn C.M., Jaye M., Schlessinger J. *et al.* (1998) Structure of a heparin-linked biologically active dimer of fibroblast growth factor. Nature;393:812–817.
87. Hricovíni M., Nieto P.N., Torri G. (2003) NMR of Sulfated Oligo and Polysaccharides. Weinhem, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.
88. Remko M., Swart M., Bickelhaupt F.M. (2007) Conformational behavior of basic monomeric building units of glycosaminoglycans: isolated systems and solvent effect. J Phys Chem B;111:2313–2321.
89. Remko M., Hricovíni M. (2007) Theoretical study of structure and properties of hexuronic acid and D-glucosamine structural units of glycosaminoglycans. Struct Chem;18:537–547.
90. Mulloy B., Forster M.J. (2000) Conformation and dynamics of heparin and heparan sulfate. Glycobiology;10:1147–1156.
91. Faham S., Hileman R.E., Fromm J.R., Linhardt R.J., Rees D.C. (1996) Heparin structure and interactions with basic fibroblast growth factor. Science;271:1116.

92. Capila I., Hernáiz M.J., Mo Y.D., Mealy T.R., Campos B., Dedman J.R. *et al.* (2001) Annexin V–heparin oligosaccharide complex suggests heparan sulfate-mediated assembly on cell surfaces. Structure,9:57–64.
93. Jin L., Abrahams J.P., Skinner R., Petitou M., Pike R.N., Carrell R.W. (1997) The anticoagulant activation of antithrombin by heparin. Proc Natl Acad Sci USA;94:14683–14688.
94. Ragazzi M., Ferro D.R., Perly B., Sinay P., Petitou M., Choay J. (1990) Conformation of the pentasaccharide corresponding to the binding site of heparin for antithrombin III. Carbohydr Res;195:169–185.
95. Ragazzi M., Ferro D.R., Provasoli A. (1986) A force–field study of the conformational characteristics of the iduraonate ring. J Comput Chem;7:105–112.
96. Hricovini M., Guerrini M., Bisio A. (1999) Structure of heparin-derived tetrasaccharide complexed to the plasma protein antithrombin derived from NOEs, J-couplings and chemical shifts. Eur J Biochem;261:789–801.
97. Hricovini M., Guerrini M., Bisio A., Torri G., Petitou M., Casu B. (2001) Conformation of heparin pentasaccharide bound to antithrombin III. Biochem J;359:265–272.
98. Hricovini M., Guerrini M., Bisio A., Torri G., Naggi A., Casu B. (2002) Active conformations of glycosaminoglycans. NMR determination of the conformation of heparin sequences complexed with antithrombin and fibroblast growth factors in solution. Semin Thromb Hemost;28:325–334.
99. Guerrini M., Guglieri S., Beccati D., Torri G., Viskov C., Mourier P. (2006) Conformational transitions induced in heparin octasaccharides by binding with antithrombin III. Biochem J;399:191–198.
100. Sasisekharan R., Raman R., Prabhakar V. (2006) Glycomics approach to structure–function relationships of glycosaminoglycans. Annu Rev Biomed Eng;8:181–231.
101. Malik A., Ahmad S. (2007) Sequence and structural features of carbohydrate binding in proteins and assessment of predictability using a neural network. BMC Struct Biol;7:1.
102. Shionyu-Mitsuyama C., Shirai T., Ishida H., Yamane T. (2003) An empirical approach for structure-based prediction of carbohydrate-binding sites on proteins. Protein Eng Des Sel;16:467–478.
103. Taroni C., Jones S., Thornton J.M. (2000) Analysis and prediction of carbohydrate binding sites. Protein Eng Des Sel;13:89–98.
104. Fromm J.R., Hileman R.E., Caldwell E.E.O., Weiler J.M., Linhardt R.J. (1997) Pattern and spacing of basic amino acids in heparin binding sites. Arch Biochem Biophys;343:92–100.
105. Hileman R.E., Fromm J.R., Weiler J.M., Linhardt R.J. (1998) Glycosaminoglycan–protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. BioEssays;20:156–167.
106. Caldwell E.E., Nadkarni V.D., Fromm J.R., Linhardt R.J., 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.
107. Thompson L.D., Pantoliano M.W., Springer B.A. (1994) Energetic characterization of the basic fibroblast growth factor–heparin interaction: identification of the heparin binding domain. Biochemistry;33:3831–3840.
108. Bae J., Desai U.R., Pervin A., Caldwell E.E., Weiler J.M., Linhardt R.J. (1994) Interaction of heparin with synthetic antithrombin III peptide analogues. Biochem J;301:121–129.
109. Jairajpuri M.A., Lu A., Desai U., Olson S.T., Bjork I., Bock S.C. (2003) Antithrombin III phenylalanines 122 and 121 contribute to its high affinity for heparin and its conformational activation. J Biol Chem;278:15941–15950.
110. Arocas V., Bock S.C., Olson S.T., Bjork I. (1999) The role of Arg46 and Arg47 of antithrombin in heparin binding. Biochemistry;38:10196–10204.
111. Desai U., Swanson R., Bock S.C., Bjork I., Olson S.T. (2000) Role of arginine 129 in heparin binding and activation of antithrombin. J Biol Chem;275:18976–18984.
112. Schedin-Weiss S., Desai U.R., Bock S.C., Gettins P.G., Olson S.T., Bjork I. (2002) Importance of lysine 125 for heparin binding and activation of antithrombin. Biochemistry;41:4779–4788.
113. Arocas V., Bock S.C., Raja S., Olson S.T., Bjork I. (2001) Lysine 114 of antithrombin is of crucial importance for the affinity and kinetics of heparin pentasaccharide binding. J Biol Chem;276:43809–43817.
114. Guo N.H., Krutzsch H.C., NÄgre E., Vogel T., Blake D.A., Roberts D.D. (1992) Heparin- and sulfatide-binding peptides from the type I repeats of human thrombospondin promote melanoma cell adhesion. Proc Natl Acad Sci USA;89:3040–3044.
115. Woods A., McCarthy J.B., Furcht L.T., Couchman J.R. (1993) A synthetic peptide from the COOH-terminal heparin-binding domain of fibronectin promotes focal adhesion formation. Mol Biol Cell;4:605–613.
116. Cardin A.D., Weintraub H.J. (1989) Molecular modeling of protein–glycosaminoglycan interactions. Arterioscler Thromb Vasc Biol;9:21–32.
117. Margalit H., Fischer N., 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.
118. Silvian L., Jin P., Carmillo P., Boriack-Sjodin P.A., Pelletier C., Rushe M. *et al.* (2006) Artemin crystal structure reveals insights into heparan sulfate binding. Biochemistry;45:6801–6812.
119. Vyas A.A., Pan J.J., Patel H.V., Vyas K.A., Chiang C.M., Sheu Y.C. *et al.* (2005) Analysis of binding of cobra cardiotoxins to heparin reveals a new beta-sheet heparin-binding structural motif. J Biol Chem;280:9567–9577.
120. Sobel M., Soler D.F., Kermode J.C., Harris R.B. (1992) Localization and characterization of a heparin binding domain peptide of human von Willebrand factor. J Biol Chem;267:8857–8862.
121. Johnson Z., Power C.A., Weiss C., Rintelen F., Ji H., Ruckle T. *et al.* (2004) Chemokine inhibition – why, when, where, which and how. Biochem Soc Trans;32:366–377.
122. Kuschert G.S., Coulin F., Power C.A., Proudfoot A.E., Hubbard R.E., Hoogewerf A.J. *et al.* (1999) Glycosaminoglycans interact selectively with chemokines and modulate receptor binding and cellular responses. Biochemistry;38:12959–12968.
123. Mulloy B., Crane D.T., Drake A.F., Davies D.B. (1996) The interaction between heparin and polylysine: a circular dichroism and molecular modelling study. Braz J Med Biol Res;29:721–729.

124. Sibille N., Sillen A., Leroy A., Wieruszeski J.M., Mulloy B., Landrieu I. *et al.* (2006) Structural impact of heparin binding to full-length Tau as studied by NMR spectroscopy. Biochemistry;45:12560–12572.
125. Rusnati M., Coltrini D., Oreste P., Zoppetti G., Albini A., Noonan D. *et al.* (1997) Interaction of HIV-1 Tat protein with heparin. Role of the backbone structure, sulfation and size. J Biol Chem;272:11313–11320.
126. Olson S.T., Chuang Y.-J. (2002) Heparin activates antithrombin anticoagulant function by generating new interaction sites (exosites) for blood clotting proteinases. Trends Cardiovasc Med;12:331–338.
127. Petitou M., Duchaussoy P., Driguez P.A., Jaurand G., Herault J.P., Lormeau J.C. *et al.* (1998) First synthetic carbohydrates with the full antocoagulant properties of heparin. Angew Chem Int Ed Engl;37:3009–3014.
128. Schlessinger J., Plotnikov A.N., Ibrahimi O.A., Eliseenkova A.V., Yeh B.K., Yayon A. *et al.* (2000) Crystal structure of a ternary FGF–FGFR–heparin complex reveals a dual role for heparin in FGFR binding and dimerization. Mol Cell;6:743–750.
129. Rusnati M., Coltrini D., Caccia P., Dell'Era P., Zoppetti G., Oreste P. *et al.* (1994) Distinct role of 2-*O*-, *N*-, and 6-*O*-sulfate groups of heparin in the formation of the ternary complex with basic fibroblast growth factor and soluble FGF receptor-1. Biochem Biophys Res Commun;203:450–458.
130. Pellegrini L., Burke D.F., von Delft F., Mulloy B., Blundell T.L. (2000) Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. Nature;407:1029–1034.
131. Canales A., Angulo J., Ojeda R., Bruix M., Fayos R., Lozano R. *et al.* (2005) Conformational flexibility of a synthetic glycosylaminoglycan bound to a fibroblast growth factor. FGF-1 recognizes both the (1) C (4) and (2) S (0) conformations of a bioactive heparin-like hexasaccharide. J Am Chem Soc;127:5778–5779.
132. Fraser P.E., Nguyen J.T., Surewicz W.K., Kirschner D.A. (1991) pH-dependent structural transitions of Alzheimer amyloid peptides. Biophys J;60:1190–1201.
133. Arteel G.E., Franken S., Kappler J., Sies H. (2000) Binding of Selenoprotein P to heparin: characterization with surface plasmon resonance. Biol Chem;381:265–268.
134. Wettreich A., Sebollela A., Carvalho M.A., Azevedo S.P., Borojevic R., Ferreira S.T. *et al.* (1999) Acidic pH modulates the interaction between human granulocyte-macrophage colony-stimulating factor and glycosaminoglycans. J Biol Chem; 274:31468–31475.
135. Matsumoto R. (1995) Packaging of proteases and proteoglycans in the granules of mast cells and other hematopoietic cells. J Biol Chem;270:19524–19531.
136. Veldkamp C.T., Peterson F.C., Pelzek A.J., 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.
137. Gandhi N.S., Coombe D.R., Mancera R.L. (2008) Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 1. Molecular modeling studies. Biochemistry;47:4851–4862.
138. Coombe D.R., Stevenson S.M., Kinnear B.F., Gandhi N.S., Mancera R.L., Osmond R.I.W. *et al.* (2008) Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 2. Biochemical analyses. Biochemistry;47:4863–4875.
139. Hallgren J., Backstrom S., Estrada S., Thuveson M., Pejler G. (2004) Histidines are critical for heparin-dependent activation of mast cell tryptase 1. J Immunol;173:1868–1875.
140. Borza D.B., Morgan W.T. (1998) Histidine–proline-rich glycoprotein as a plasma pH sensor. Modulation of its interaction with glycosaminoglycans by pH and metals. J Biol Chem;273:5493–5499.
141. Chevalier F., Lucas R., Angulo J., Martin-Lomas M., Nieto P.M. (2004) The heparin–Ca (2+) interaction: the influence of the O-sulfation pattern on binding. Carbohydr Res;339:975–983.
142. Chevalier F., Angulo J., Lucas R., Nieto P.M., Martin M. (2002) The heparin–Ca^{2-} interaction: structure of the Ca^{2+} binding site. Eur J Org Chem;2367:2376.
143. Shao C., Zhang F., Kemp M.M., Linhardt R.J., Waisman D.M., Head J.F. *et al.* (2006) Crystallographic analysis of calcium-dependent heparin binding to Annexin A2. J Biol Chem;281:31689–31695.
144. Ricard-Blum S., Feraud O., Lortat-Jacob H., Rencurosi A., Fukai N., Dkhissi F. *et al.* (2004) Characterization of endostatin binding to heparin and heparan sulfate by surface plasmon resonance and molecular modeling: role of divalent cations. J Biol Chem;279:2927.
145. Gonza'lez-Iglesias R., Pajares M.A., Ocal C., Oesch B., Gasset M. (2002) Prion protein interaction with glycosaminoglycan occurs with the formation of oligomeric complexes stabilized by Cu (II) bridges. J Mol Biol;319:527–540.
146. Srinivasan S.R., Radhakrishnamurthy B., Berenson G.S. (1975) Studies on the interaction of heparin with serum lipoproteins in the presence of Ca^{2+}, Mg^{2+}, and Mn^{2+}. Arch Biochem Biophys;170:334–340.
147. Parrish R.F., Fair W.R. (1981) Selective binding of zinc ions to heparin rather than to other glycosaminoglycans. Biochem J;193:407–410.
148. Whitfield D.M., Choay J., Sarkar B. (1992) Heavy metal binding to heparin disaccharides. I. Iduronic acid is the main binding site. Biopolymers;32:585–596.
149. Whitfield D.M., Sarkar B. (1992) Heavy metal binding to heparin disaccharides. II. First evidence for zinc chelation. Biopolymers;32:597–619.
150. Imberty A., Lortat-Jacob H., Pérez S. (2007) Structural view of glycosaminoglycan–protein interactions. Carbohydr Res;342:430–439.
151. Lortat-Jacob H., Grosdidier A., Imberty A. (2002) Structural diversity of heparan sulfate binding domains in chemokines. Proc Natl Acad Sci USA;99:1229–1234.
152. Goodford P.J. (1985) A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J Med Chem;28:849–857.
153. Bitomsky W., Wade R.C. (1999) Docking of glycosaminoglycans to heparin-binding proteins: validation for aFGF, bFGF, and antithrombin and application to IL-8. J Am Chem Soc;121:3004–3013.
154. Mulloy B., Forster M.J. (2008) Application of drug discovery software to the identification of heparin-binding sites on protein surfaces: a computational survey of the 4-helix cytokines. Mol Simul;34:481–489.

155. Fallahi A., Kroll B., Warner L.R., Oxford R.J., Irwin K.M., Mercer L.M. *et al.* (2005) Structural model of the amino propeptide of collagen XI {alpha}1 chain with similarity to the LNS domains. Protein Sci;14:1526–1537.
156. da.Rocha P.S.S., Tácio V., Amorim F., Raul C.E., Geraldo P.P. (2008) Studies of molecular docking between fibroblast growth factor and heparin using generalized simulated annealing. Int J Quantum Chem;108:2608–2614.
157. Forster M., Mulloy B. (2006) Computational approaches to the identification of heparin-binding sites on the surfaces of proteins. Biochem Soc Trans;34:431–434.
158. Kern A., Schmidt K., Leder C., Muller O.J., Wobus C.E., Bettinger K. *et al.* (2003) Identification of a heparin-binding motif on adeno-associated virus type 2 capsids. J Virol;77:11072–11081.
159. Lam K., Rao V.S., Qasba P.K. (1998) Molecular modeling studies on binding of bFGF to heparin and its receptor FGFR1. J Biomol Struct Dyn;15:1009–1027.
160. Sadir R., Baleux F., Grosdidier A., Imberty A., Lortat-Jacob H. (2001) Characterization of the stromal cell-derived factor-1a–heparin complex. J Biol Chem;276:8288–8296.
161. Grootenhuis P.D.J., Van Boeckel C.A.A. (1991) Constructing a molecular model of the interaction between antithrombin III and a potent heparin analog. J Am Chem Soc;113:2743–2747.
162. Canales A., Lozano R., Lopez-Mendez B., Angulo J., Ojeda R., Nieto P.M. *et al.* (2006) Solution NMR structure of a human FGF-1 monomer, activated by a hexasaccharide heparin-analogue. FEBS J;273:4716.
163. Raghuraman A., Mosier P.D., Desai U.R. (2006) Finding a needle in a haystack: development of a combinatorial virtual screening approach for identifying high specificity heparin/heparan sulfate sequence (s). J Med Chem;49:3553–3562.
164. Nicolaes G.A.F., rensen K.W., Friedrich U., Tans G., Rosing J., Autin L. *et al.* (2004) Altered inactivation pathway of factor Va by activated protein C in the presence of heparin. Eur J Biochem;271:2724–2736.
165. Keserü G.M., Kolossváry I. (1999) Molecular Mechanics and Conformational Analysis in Drug Design. Oxford, UK: Blackwell Publishing.
166. Cochran S., Li C.P., Bytheway I. (2005) An experimental and molecular–modeling study of the binding of linked sulfated tetracyclitols to FGF-1 and FGF-2. Chembiochem;6:1882–1890.
167. Kerzmann A., Neumann D., Kohlbacher O. (2006) SLICK-scoring and energy functions for protein–carbohydrate interactions. J Chem Inf Model;46:1635–1642.
168. Laederach A., Reilly P.J. (2003) Specific empirical free energy function for automated docking of carbohydrates to proteins. J Comput Chem;24:1748–1757.
169. Liu L., Bytheway I., Karoli T., Fairweather J.K., Cochran S., Li C. *et al.* (2008) Design, synthesis, FGF-1 binding, and molecular modeling studies of conformationally flexible heparin mimetic disaccharides. Bioorg Med Chem Lett;18:344–349.
170. Bytheway I., Cochran S. (2004) Validation of molecular docking calculations involving FGF-1 and FGF-2. J Med Chem;47:1683–1693.
171. Mitchell J.B.O., Laskowski R.A., Alex A., Thornton J.M. (1999) BLEEP – potential of mean force describing protein–ligand interactions: I. Generating potential. J Comput Chem;20:1165–1176.
172. Aqvist J., Marelius J. (2001) The linear interaction energy method for predicting ligand binding free energies. Comb Chem High Throughput Screen;4:613–626.
173. Laitinen T., Rouvinen J., Peräkylä M. (2003) MM–PBSA free energy analysis of endo-1, 4-xylanase II (XynII)–substrate complexes: binding of the reactive sugar in a skew boat and chair conformation. Org Biomol Chem;1:3535–3540.
174. Verli H., Guimarães J.A. (2005) Insights into the induced fit mechanism in antithrombin–heparin interaction using molecular dynamics simulations. J Mol Graph Model;24:203–212.
175. Verli H., Guimarães J.A. (2004) Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution. Carbohydr Res;339:281–290.
176. Pol-Fachin L., Verli H. (2008) Depiction of the forces participating in the 2-*O*-sulfu-[alpha]-l-iduronic acid conformational preference in heparin sequences in aqueous solutions. Carbohydr Res;343:1435–1445.
177. Kony D., Damm W., Stoll S., Gunsteren W.F.V. (2002) An improved OPLS–AA force field for carbohydrates. J Comput Chem;23:1416–1429.
178. Lins R.D., Hünenberger P.H. (2005) A new GROMOS force field for hexopyranose-based carbohydrates. J Comput Chem;26:1400–1412.
179. Kuttel M., Brady J.W., Naidoo K.J. (2002) Carbohydrate solution simulations: producing a force field with experimentally consistent primary alcohol rotational frequencies and populations. J Comput Chem;23:1236–1243.
180. Brooks B.R., Bruccoleri R.E., Olafson B.D., States D.J., Swaminathan S., Karplus M. (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem;4:187–217.
181. Grootenhuis P.D.J., Haasnoot C.A.G. (1993) A CHARMm based force field for carbohydrates using the CHEAT approach: carbohydrate hydroxyl groups represented by extended atoms. Mol Simul;10:75–95.
182. Woods R.J., Dwek R.A., Edge C.J., Fraser-Reid B. (1995) Molecular mechanical and molecular dynamic simulations of glycoproteins and oligosaccharides. 1. GLYCAM_93 parameter development. J Phys Chem;99:3832–3846.
183. Allinger N.L. (1977) Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. J Am Chem Soc;99:8127–8134.
184. Allinger N.L., Yuh Y.H., Lii J.H. (1989) Molecular mechanics. The MM3 force field for hydrocarbons. 1. J Am Chem Soc;111:8551–8566.
185. Fabricius J., Engelsen S.B., Rasmussen K. (1997) The Consistent Force Field. 5. PEF95SAC: optimized potential energy function for alcohols and carbohydrates. J Carbohydr Chem;16:751–772.
186. Imberty A., Bettler E., Karababa M., Mazeau K., Petrova P., Pérez S. (1999) Building Sugars: The Sweet Part of Structural Biology. Hyderabad: Perspectives in Structural Biology Indian Academy of Sciences and Universities Press;p. 392–409.
187. Ferro D.R., Pumilia P., Cassinari A., Ragazzi M. (1995) Treatment of ionic species in force–field calculations: sulfate and carboxylate groups in carbohydrates. Int J Biol Macromol;17:131–136.

188. Ferro D.R., Provasoli A., Ragazzi M. (1987) An improved force field for conformational analysis of sulfated polysaccharides. J Comput Chem;8:351–367.
189. Huige C.J.M., Altona C. (1995) Force field parameters for sulfates and sulfamates based on Ab initio calculations: extensions of AMBER and CHARMm fields. J Comput Chem; 16:56–79
190. Kirschner K.N., Yongye A.B., Tschampel S.M., Gonzalez-Outeirino J., Daniels C.R., Foley B.L. *et al.* (2008) GLYCAM06: a generalizable biomolecular force field. Carbohydrates. J Comput Chem;29:622–655.
191. Jin L. (2007) PhD Thesis: University of Edinburgh.
192. Lindahl U. (2007) Heparan sulfate–protein interactions – a concept for drug design? Thromb Haemost;98:109–115.
193. Fugedi P. (2003) The potential of the molecular diversity of heparin and heparan sulfate for drug development. Mini Rev Med Chem;3:659–667.
194. Johnson Z., Proudfoot A.E., Handel T.M. (2005) Interaction of chemokines and glycosaminoglycans: a new twist in the regulation of chemokine function with opportunities for therapeutic intervention. Cytokine Growth Factor Rev;16:625–636.
195. Spillmann D., Lindahl U. (1994) Glycosaminoglycan–protein interactions: a question of specificity. Curr Opin Struct Biol;4:677–682.
196. Coombe D.R. (2008) Biological implications of glycosaminoglycan interactions with haemopoietic cytokines. Immunol Cell Biol;86:598–607.
197. Webb L.M., Ehrengruber M.U., Clark-Lewis I., Baggiolini M., Rot A. (1993) Binding to heparan sulfate or heparin enhances neutrophil responses to interleukin 8. Proc Natl Acad Sci USA;90:7158–7162.
198. Codée J.D.C., Overkleeft H.S., van der Marel G.A., van Boeckel C.A.A. (2004) The synthesis of well-defined heparin and heparan sulfate fragments. Drug Discov Today Technol;1:317–326.
199. Choay J., Petitou M., Lormeau J.C., Sinay P., Casu B., Gatti G. (1983) Structure activity relationship in heparin: a synthetic pentasaccharide with high affinity for antithrombin III and eliciting high anti-factor Xa activity. Biochem Biophys Res Commun;116:492–499.
200. Johnson D.J.D., Li W., Adams T.E., Huntington J.A. (2006) Antithrombin–S195A factor Xa–heparin structure reveals the allosteric mechanism of antithrombin activation. EMBO J;25:2029–2037.
201. Tan K., Duquette M., Liu J.-h., Zhang R., Joachimiak A., Wang J.-h. *et al.* (2006) The structures of the thrombospondin-1 N-terminal domain and its complex with a synthetic pentameric heparin. Structure;14:33–42.
202. Petitou M., van Boeckel C.A.A. (2004) A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next. Angew Chem Int Ed Engl;43:3118–3133.
203. Herbert J.M., Herault J.P., Bernat A., van Amsterdam R.G.M., Lormeau J.C., Petitou M. *et al.* (1998) Biochemical and pharmacological properties of SANORG 34006, a potent and long-acting synthetic pentasaccharide. Blood;91:4197.
204. Li W., Johnson D.J., Esmon C.T., Huntington J.A. (2004) Structure of the antithrombin-thrombin–heparin ternary complex reveals the antithrombotic mechanism of heparin. Nat Struct Mol Biol;11:857–862.
205. Ferro V., Don R. (2003) The development of the novel angiogenesis inhibitor PI-88 as an anticancer drug. Australas Biotechnol;13:38–39.
206. Wall D., Douglas S., Ferro V., Cowden W., Parish C. (2001) Characterisation of the anticoagulant properties of a range of structurally diverse sulfated oligosaccharides. Thromb Res;103:325–335.
207. Freeman C., Liu L., Banwell M.G., Brown K.J., Bezos A., Ferro V. *et al.* (2005) Use of sulfated linked cyclitols as heparan sulfate mimetics to probe the heparin/heparan sulfate binding specificity of proteins. J Biol Chem;280:8842–8849.
208. Raghuraman A., Riaz M., Hindle M., Desai U.R. (2007) Rapid and efficient microwave-assisted synthesis of highly sulfated organic scaffolds. Tetrahedron Lett;48:6754–6758.
209. Yamaguchi M., Ohmori T., Sakata Y., Ueki M. (2008) Oligo(tyrosine sulfate)s as heparin pentasaccharide mimic: evaluation by surface noncovalent affinity mass spectrometry. Bioorg Med Chem;16:3342–3351.
210. Desai U.R. (2004) New antithrombin-based anticoagulants. Med Res Rev;24:151–181.
211. Alban S. (2008) Pharmacological strategies for inhibition of thrombin activity. Curr Pharm Des;14:1152–1175.
212. Monien B.H., Cheang K.I., Desai U.R. (2005) Mechanism of Poly(acrylic acid) acceleration of antithrombin inhibition of thrombin: implications for the design of novel heparin mimics. J Med Chem;48:5360–5368.
213. Monien B.H., Henry B.L., Raghuraman A., Hindle M., Desai U.R. (2006) Novel chemo-enzymatic oligomers of cinnamic acids as direct and indirect inhibitors of coagulation proteinases. Bioorg Med Chem;14:7988–7998.
214. Gunnarsson G.T., Desai U.R. (2002) Designing small, nonsugar activators of antithrombin using hydropathic interaction analyses. J Med Chem;45:1233–1243.
215. Kellogg G.E., Joshi G.S., Abraham D.J. (1992) New tools for modeling and understanding hydrophobicity and hydrophobic interactions. Med Chem Res;1:444–453.
216. Kellogg G.E., Semus S.F., Abraham D.J. (1991) HINT: a new method of empirical hydrophobic field calculation for CoMFA. J Comput Aided Mol Des;5:545–552.
217. Gunnarsson G.T., Desai U.R. (2002) Interaction of designed sulfated flavanoids with antithrombin: lessons on the design of organic activators. J Med Chem;45:4460–4470.
218. Zhang J., Rivers G., Zhu Y., Jacobson A., Peyers J., Grundstrom G. *et al.* (2001) Identification of inhibitors of heparin growth factor interactions from combinatorial libraries of four-component condensation reactions. Bioorg Med Chem;9:825–836.
219. Kim S.H., Kiick K.L. (2007) Heparin-mimetic sulfated peptides with modulated affinities for heparin-binding peptides and growth factors. Peptides;28:2125–2136.
220. Raghuraman A., Tiwari V., Thakkar J.N., Gunnarsson G.T., Shukla D., Hindle M. *et al.* (2005) Structural characterization of a serendipitously discovered bioactive macromolecule, lignin sulfate. Biomacromolecules;6:2822–2832.
221. Schuksz M., Fuster M.M., Brown J.R., Crawford B.E., Ditto D.P., Lawrence R. *et al.* (2008) Surfen, a small molecule antagonist of heparan sulfate. Proc Natl Acad Sci USA;105:13075–13080.

222. de Paz J.L., Moseman E.A., Noti C., Polito L., von Andrian U.H., Seeberger P.H. (2007) Profiling heparin–chemokine interactions using synthetic tools. ACS Chem Biol;2:735–744.
223. de Paz J.L., Noti C., Seeberger P.H. (2006) Microarrays of synthetic heparin oligosaccharides. J Am Chem Soc;128:2766–2767.
224. Paz J.L., Seeberger P.H. (2008) Deciphering the glycosaminoglycan code with the help of microarrays. Mol BioSys;4:707–711.
225. Seeberger P.H., Werz D.B. (2007) Synthesis and medical applications of oligosaccharides. Nature;446:1046.
226. Fernández C., Hattan C.M., Kerns R.J. (2006) Semi-synthetic heparin derivatives: chemical modifications of heparin beyond chain length, sulfate substitution pattern and *N*-sulfo/*N*-acetyl groups. Carbohydr Res;341:1253–1265.
227. Huang L., Kerns R.J. (2006) Diversity-oriented chemical modification of heparin: identification of charge-reduced *N*-acyl heparin derivatives having increased selectivity for heparin-binding proteins. Bioorg Med Chem;14:2300–2313.
228. Boulanger M.J., Garcia K.C., Garcia K.C. (2004) Shared Cytokine Signaling Receptors: Structural Insights From the Gp130 System. San Diego, CA, USA: Academic Press;p. 107–146.
229. Walter M.R., Garcia K.C. (2004) Structural Analysis of IL-10 and Type I Interferon Family Members and THEIR Complexes with Receptor. San Diego, CA, USA: Academic Press;p. 171–223.
230. Lau E.K., Allen S., Hsu A.R., Handel T.M., Garcia K.C. (2004) Chemokine–Receptor Interactions: GPCRs, Glycosaminoglycans and Viral Chemokine Binding Proteins. San Diego, CA, USA: Academic Press;p. 351–391.
231. Oppenheim J.J., Feldmann M., Durum S.K. (2001) Cytokine Reference: A Compendium of Cytokines and Other Mediators of Host Defense. San Diego, CA; London: Academic
232. Ceroni A., Dell A., Haslam S. (2007) The GlycanBuilder: a fast, intuitive and flexible software tool for building and displaying glycan structures. Source Code Biol Med;2:3.

Notes

[3]http://www.glycosciences.de/tools/linucs/
[4]http://www.cermav.cnrs.fr/cgi-bin/gag/gag.cgi.
[5]Accelrys, Inc., San Diego, CA, USA.

3 Heparin/Heparan Sulphate-based Drugs

ELSEVIER

Heparin/heparan sulphate-based drugs

Neha S. Gandhi[1] and Ricardo L. Mancera[1,2,]

[1] Curtin Health Innovation Research Institute, Western Australian Biomedical Research Institute, School of Biomedical Sciences, Curtin University, GPO Box U1987, Perth, WA 6845, Australia
[2] School of Pharmacy, Curtin University, GPO Box U1987, Perth, WA 6845, Australia

Glycosaminoglycans (GAGs) are an untapped source of novel chemical entities and, therefore, offer exciting new opportunities for the development of novel drug molecules because of their unique physical and biological properties. Advances in the functional understanding of GAG–protein interactions are enabling the development of GAG mimetics for use as anti-angiogenic, anti-metastatic, anti-inflammatory, anticoagulant and anti-thrombotic agents. Many anti-thrombotic molecules, such as Fondaparinux and Idraparinux, have been successful in clinical trials, and a new generation of heparin mimetic oligosaccharides and small molecules are currently in different stages of clinical development. In particular, the recent increased activity in the development of new mimetics by altering the composition of sulphated GAGs is very encouraging. This article reviews structurally defined heparin-mimetic oligosaccharides and small molecules currently in development or clinical trials.

Sulphated glycosaminoglycans (GAGs) are glycans found inside the cell and in the extracellular matrix, which act by binding selectively to a variety of proteins and pathogens and are crucially relevant to many disease processes, such as inflammation [1–3], neurodegeneration [4], angiogenesis [5], cardiovascular disorders [6], cancer [7] and infectious diseases [8–10]. Heparin and heparan sulphate (HS) are GAGs consisting of 1–4 linked uronic acid and glucosamine and encompassing varying degrees of sulphation, and they are involved in many of these activities [11]. Heparin is a minor form of the ubiquitous HS, and the anticoagulant activity of pharmaceutical heparin is mainly accounted by fractions containing a pentasaccharide sequence with binding affinity for anti-thrombin (AT) (see below).

The wide range and intricacy of glycan-mediated cellular interactions have turned glycans into novel targets for future drug development [12–14], with drugs already being developed for the treatment of metabolic disorders, cancer and infection. In recent years, there has been a renaissance in the development of carbohydrate-based therapeutics that involve the inhibition of carbohydrate–lectin interactions and carbohydrate-based anticoagulant and AT agents [15]. The pharmacological and therapeutic value of heparin/HS and their mimetics is now recognized because of their ability to bind and cause immobilization and/or activation of a variety of proteins, such as growth factors, chemokines and metalloproteinases [16,17]. Potential strategies based on heparin/HS–protein interactions have recently been described to assist GAG-based drug discovery [18]. GAG-based drugs can act in several ways by activating (agonists) or inactivating (antagonists) protein-based receptors, competing with endogenous GAGs and/or inhibiting GAG biosynthesis. The molecular diversity of heparin/HS interactions has been exploited for the development and clinical progression of GAG mimetics [19].

The anticoagulant market has been very active recently because of the development of new compounds, including indirect factor Xa (FXa) inhibitors (such as Fondaparinux and Idraparinux and its new biotinylated form), direct inhibitors of FXa (such as Rivaroxaban and Apixaban) and direct inhibitors of thrombin (such as Dabigatran) [20,21]. The mechanism of action of these anticoagulants has been reviewed extensively [22,23]. The discovery of the mechanism of binding of heparin to AT and FXa has focused interest on the development of small, structurally defined heparin mimetics with AT activity but with reduced side-effects [24]. This

Corresponding author:. Mancera, R.L. (r.mancera@curtin.edu.au)

1359-6446/06/$ - see front matter © 2010 Elsevier Ltd. All rights reserved. doi:10.1016/j.drudis.2010.10.009

article reviews the principles under which recently developed GAG (heparin/HS)-based inhibitors act and gives a description of the different classes of inhibitors and their development as drugs.

Heparin/HS mimetics as anticoagulants

Anticoagulants based on heparin/HS are drugs of choice in the therapy and prophylaxis of thromboembolic diseases [25]. Structural and functional studies have shown that a unique pentasaccharide (sometimes referred to as AGA*IA or DEFGH), GlcNAc/NS6S → GlcA → GlcNS3S6S → IdoA2S → GlcNS6S (where Glc is glucosamine, IdoA is iduronic acid and GlcA is glucoronic acid, which are either sulphated or acetylated), comprises the AT-binding domain and is responsible for the anticoagulant activity of heparin. The 3-*O*-sulphate group at position F is responsible for strong and specific interactions with AT [26]. The pentasaccharide unit only inhibits the activity of FXa mediated by AT; however, a much larger oligosaccharide is required for the AT-mediated inhibition of thrombin [27]. The structural requirements for the binding of heparin (Fig. 1a) to AT, as shown in Fig. 1, were determined by analysing the crystal structure and by determining the structure–activity relationships of a series of pentasaccharides using various combinations of sulphate and carboxylate groups [28,29]. This approach helped to establish that charged groups, as depicted in Fig. 1, are absolutely essential for the activation of AT (highlighted in the blue boxes) and required to increase the biological activity (in the red boxes). Moreover, hydrophobic interactions between the heparin pentasaccharide and AT also contribute to increasing the binding affinity [30]. Several review

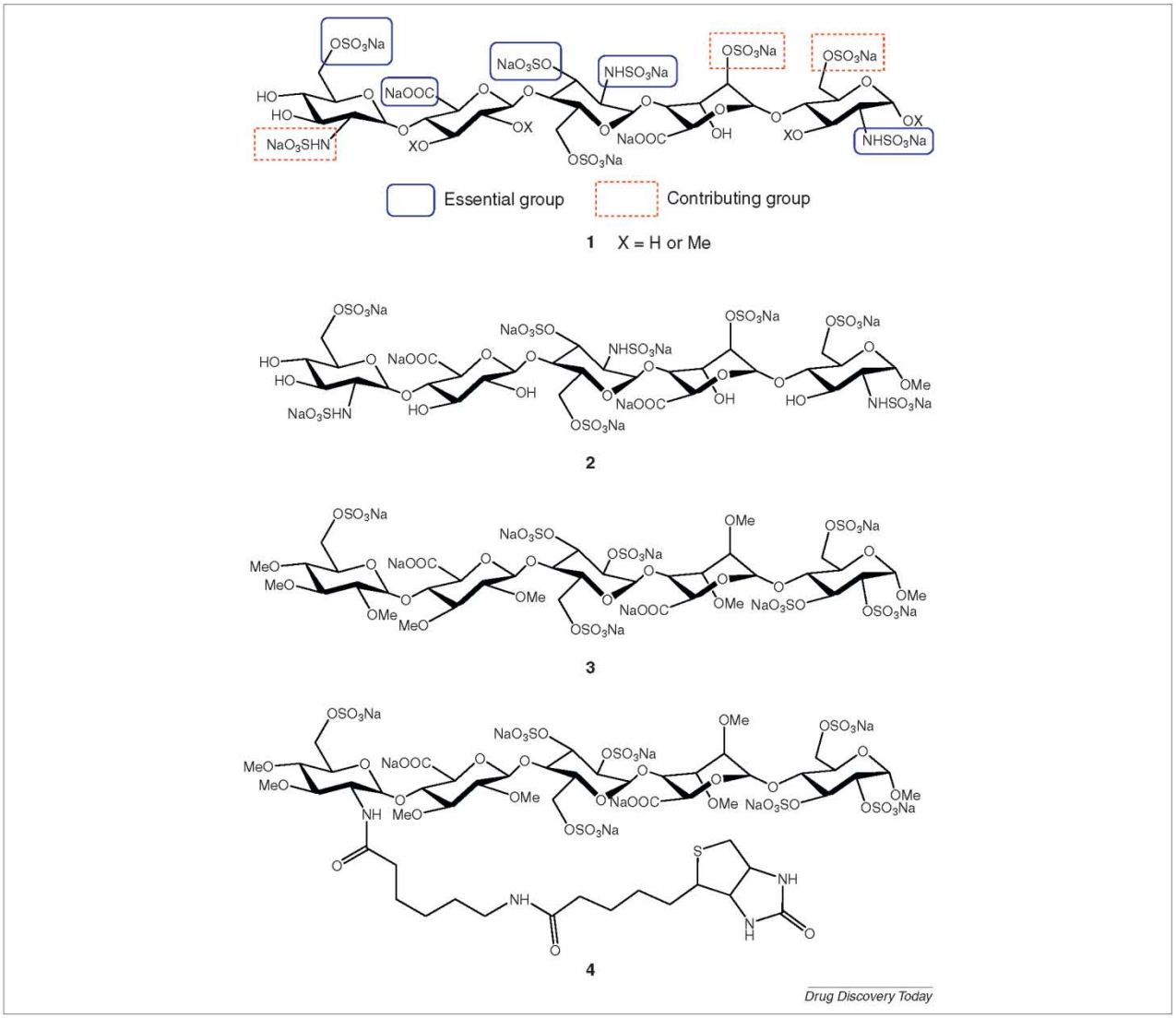


FIGURE 1

Chemical structures of heparin pentasaccharide derivatives. (**a**) The AT-binding pentasaccharide motif. Highlighted functional groups are essential for AT activation. (**b**) Structure of Fondaparinux. (**c**) Structure of Idraparinux. (**d**) Structure of Idrabiotaparinux. Natural heparin and synthetic pentasaccharides differ in their substitution pattern (synthetic pentasaccharides contains –OMe substitutions).

articles have been published describing the structure–activity relationship and mechanism of action of heparin mimetic anticoagulants [24,28,29,31,32].

Research on heparin mimetic anticoagulants has gained momentum since the successful clinical development programs of the 1990s. GlaxoSmithKline registered Fondaparinux (Fig. 1b) (SR90107, Org31540) as a new anti-thrombotic drug under the name Arixtra® after being granted approval from the US FDA and the European Committee for Proprietary Medical Products [33]. SR123781 is a short-acting hexadecasaccharide analogue of heparin with *N*-sulphate groups replaced by *O*-sulphates and alkylated hydroxyl groups in the AT-binding domain. It has tailor-made FXa- and thrombin-inhibitory activities combined with more selectivity in its mode of action. Sanofi-Aventis discontinued the development of SR123781, however, after the success of heparin mimetic AVE 5026 [34].

Idraparinux (SanOrg34006, SR34006)

Idraparinux (Fig. 1c) is a synthetic pentasaccharide analogue of Fondaparinux, in which the hydroxyl groups are methylated and the *N*-sulphate groups are replaced by *O*-sulphates [35]. Idraparinux (K_d of 1 nM) interacts more strongly with AT than Fondaparinux (K_d of 50 nM) through non-ionic interactions [36] and also exhibits superior anti-Xa activity (1600 versus 700 U μg^{-1}) [35,37,38]. Idraparinux can be synthesized more easily than Fondaparinux because of the presence of a 'pseudo'-alternating sequence that can be easily prepared from glucose [39]. The crystal structures of the complexes of a pentasaccharide analogue of Idraparinux (modified by the addition of a sulphate at the H unit) [40] with a dimer consisting of activated AT and latent AT [41], and with an intermediate state [42], have been reported. Analysis of these structures explains the lower affinity of heparin for AT on the basis of induced conformational changes in AT, such as the expulsion of the hinge region and the closure of β-sheet A to the normal five-stranded form. This then leads to the activation of AT and the allosteric inhibition of coagulation factors IXa, Xa and thrombin.

Idraparinux has been evaluated in clinical trials for the treatment and secondary prevention of venous thromboembolism (VTE) and the prevention of thromboembolic events associated with atrial fibrillation (AF) [43]. The pharmacokinetics, pharmacodynamics and tolerability of Idraparinux were evaluated in several phase I studies [44,45]. Idraparinux has a long half-life (80–120 h) in the bloodstream, thus enabling once-weekly administration [46]. The phase II dose-ranging PERSIST study established that subcutaneous once-weekly administration of 2.5 mg of Idraparinux, with an increased elimination half-life of approximately 600 h, was as effective as warfarin and demonstrated dose-dependent increases in major bleeding for the secondary prevention of deep vein thrombosis (DVT) [47–49]. Idraparinux was also evaluated for the long-term treatment of patients with DVT and pulmonary embolism (PE) using a subcutaneous once-weekly dose of 2.5 mg in the three van Gogh trials [50,51]. Unfortunately, the results of these trials indicated that the rate of recurrent VTE was considerably higher with Idraparinux than with conventional therapy. In addition, the elimination half-life of 60 days led to a prolonged anticoagulant effect after completion of the therapy [52]. These findings also explained the complication of increased bleeding during a 12-month treatment period compared to a six-month treatment period in patients randomly treated with Idraparinux in the DVT and PE study, as carried out in the van Gogh Extension trial [53,54]. However, repeated doses of reversal agents, such as rFVIIa, are required in bleeding patients during surgery to neutralize the anticoagulant effect of Idraparinux [55].

The van Gogh and Amadeus phase III clinical trials of Idraparinux established that its pharmacokinetics were best described by a three-compartment model in patients with DVT, PE or AF at risk of thromboembolic events [56]. The terminal half-life was measured to be 66.3 days, and the half-life during steady state was determined to be 35 weeks. Idraparinux clearance was notably related to subject weight, creatinine clearance, sex and age. The phase III AMADEUS non-inferiority trial enrolled patients with AF at risk for thromboembolism to compare the efficacy and safety of Idraparinux to therapy with vitamin K antagonists [57]. Idraparinux proved as effective as vitamin K antagonists; however, the trial was cut short because of an excess of bleeding complications in Idraparinux-treated patients and a few documented ischemic events. Idraparinux remains in the late clinical development pipeline of Sanofi-Aventis because of its advantageous (compared to oral anticoagulants and Fondaparinux) once-a-week dosing regimen.

Idrabiotaparinux (SSR126517) (Fig. 1d) is a novel synthetic anticoagulant linked to biotin at position 2 of the non-reducing end of glucose in Idraparinux [58]. Linkage of biotin at this position in the pentasaccharide prevents interaction of the pentasaccharide with AT or FXa *in vitro* [59]. The optimal length of the spacer was found to be a 6C-length arm. Administration of Idrabiotaparinux to rats by either the intravenous or the subcutaneous route resulted in a similar pharmacokinetic profile to that of Idraparinux. Further animal studies into Idrabiotaparinux showed that the injection of avidin triggered the immediate elimination of the molecule from the bloodstream, resulting in the complete neutralization of the anti-thrombotic activity of Idrabiotaparinux. Sanofi-Aventis has halted the development of Idrabiotaparinux in AF in phase III trials, because of its lack of potential benefit over oral anticoagulants, such as vitamin K antagonists, which are currently being evaluated in clinical trials [60].

AVE5026

AVE5026 (Sanofi-Aventis) is in clinical development for the prevention of VTE [61]. This molecule is a complex mixture of oligomeric ultra-low-molecular-weight heparin (LMWH) fragments (molecular weights 2000–3000 Da) with a polydispersity index of approximately 1.0. It is prepared by partial and controlled chemioselective depolymerization of porcine unfractionated heparin (UFH). AVE5026 primarily targets FXa and has only a minimal effect on thrombin. It exhibits a higher ratio of FXa to anti-Factor IIa activity (>30:1). In addition, it shows dose-dependent anti-thrombotic activity in a rat microvascular thrombosis disease model, suggesting that this agent might provide the optimal treatment for cancer-associated thrombosis [62]. When given subcutaneously, the half-life of AVE5026 is 16–20 h, enabling once-daily administration. AVE5026 is excreted renally and, like Fondaparinux, its anticoagulant effects are not neutralized by protamine sulphate. An elective total knee replacement surgery study demonstrated a highly statistically significant dose-dependent response with AVE5026 for the prevention of VTE in patients

undergoing knee arthroplasty [63]. A 20 mg dose of AVE5026 was selected for further evaluation. An extensive phase III trial is currently comparing AVE5026 with the LMWH Enoxaparin for the prevention of VTE in patients undergoing hip, knee or abdominal surgery and in cancer patients receiving chemotherapy.

M118

Momenta Pharmaceuticals developed M118 (Fig. 2), a novel anticoagulant for the treatment of patients with acute coronary syndrome. It is currently being evaluated in a phase II clinical trial with patients undergoing percutaneous coronary intervention [64]. M118 is an optimized polysaccharide compound engineered from UFH using a specific enzymatic depolymerization process. It is designed to act at multiple points in the coagulation cascade by selectively binding to both AT and thrombin, two crucial factors involved in the formation of clots [65]. Preclinical and phase I studies have shown that M118 has the positive attributes of both UFH (reversibility, monitorability and broad inhibition of the coagulation cascade) and LMWH (adequate bioavailability and predictable pharmacokinetics that enable subcutaneous administration) and can thus be administered both intravenously and subcutaneously [66]. M118 exhibits clear dose-dependent inhibition of FXa and Factor IIa, with an anti-Xa:anti-IIa ratio that is constant over time [67,68]. M118 was found to be effective at preventing thrombosis in diseased arteries in a photochemical carotid artery injury model in ApoE$^{-/-}$ mice [69]. The reduced polydispersity (the ratio of weight averaged to number-averaged molecular weight) of M118 contributes to a more predictable pharmacokinetic profile. M118 lacks drug–drug interactions when co-administered with aspirin and clopidogrel or with glycoprotein IIb/IIIa inhibitors such as Eptifibatide [70].

EP42675 and EP217609

EP42675 is the first representative of a new class of synthetic parenteral anticoagulants with a dual mechanism of action combining the properties of an indirect FXa inhibitor and a direct thrombin inhibitor [71]. EP42675 is being trialled in patients with acute coronary syndrome undergoing percutaneous coronary intervention. The structure of EP42675 contains an AT-binding pentasaccharide (an indirect FXa inhibitor) coupled to a peptidomimetic α-NAPAP analogue (a direct inhibitor of the active site of both free and clot-bound thrombin). This dual mechanism imparts a unique pharmacological profile to EP42675: (i) inhibition of both fibrin-bound and fluid-phase thrombin owing to the presence of a direct thrombin-inhibiting moiety, (ii) inhibition of FXa in the presence of AT, (iii) a favourable pharmacokinetic profile that ensures prolonged anticoagulant coverage with improved control over its therapeutic window owing to the presence of the Fondaparinux pentasaccharide, and (iv) no cross-reaction with platelet factor 4 antibodies.

The use of anticoagulants can result in haemorrhagic adverse events and hence the availability of an antidote is highly desirable. An antidote can also be useful in case an anticoagulated patient needs to be operated on urgently. In the development of EP42675, therefore, a biotin entity was covalently linked to the spacer between the pentasaccharide portion and the direct thrombin inhibitor portion of the molecule to give EP217609 (Fig. 3), which enables it to be neutralized upon administration of avidin [72]. EP42675 has successfully completed phase I trials with 100 healthy subjects, where it was found to be well tolerated and showed predictable pharmacokinetic and pharmacodynamic profiles, with low intra- and inter-subject variabilities [73]. The half-lives of EP42675 and EP217609 in rats was determined to be approximately three hours [31]. In animals, the pharmacokinetic/pharmacodynamic profiles of EP217609 and EP42675 were found to be similar.

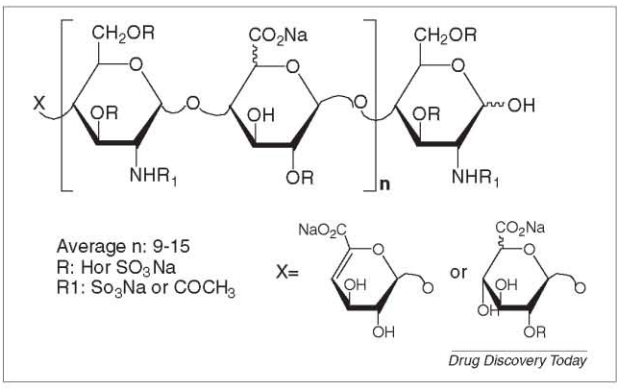


FIGURE 2

Chemical structure of M118, a LMWH characterized by a weight-averaged molecular mass between 5500 and 9000 Da and a polydispersity of approximately 1.0. M118 has the structural formula $C_{12m}H_{14m+1}O_{10m}N_mNa_mR_{3m-1}R1_m$, $C_{12m}H_{14m+2}O_{10m+1}N_mNa_mR_{3m}R1_m$, where n is equal to the average number of disaccharide repeats, $m = 1 + n$, R is H or SO_3Na and R_1 is SO_3Na or $COCH_3$.

Non-anticoagulant heparin/HS mimetics

Heparin is known to inhibit the synthesis, expression and/or function of adhesion molecules, cytokines, chemokines, proteases and viral proteins [74]. Consequently, attention has been focused recently on the non-anticoagulant properties of heparin, which are known to inhibit inflammation [63,75] and the metastatic spread of tumour cells [7,76].

PI-88 (Muparfostat)

Heparanase is an endoglycosidase enzyme that has vital roles in inflammation, tumour cell invasion, metastasis and angiogenesis [77,78]. Heparanase is the enzyme responsible for processing HS. Several sulphated sugar molecules such as cyclitols and glycol-split derivatives have been identified as selective inhibitors of heparanase–heparin interactions [79]. PI-88 (Progen Pharmaceuticals) (Fig. 4) is one such inhibitor. PI-88 inhibits heparanase and the cleavage of HS by binding competitively with HS, thereby preventing the release of growth factors, such as FGF-1, FGF-2 and VEGF, involved in angiogenesis [80]. PI-88 has progressed to clinical trials to treat inflammatory diseases, thrombosis, viral infections and cancer [81].

PI-88 is a phosphomannopentose sulphate (6-*O*-PO_3H_2-α-D-Man-(1 → 3)-α-D-Man-(1 → 3)-α-D-Man-(1 → 3)-α-D-Man-(1 → 2)-D-Man) (Fig. 4), wherein the chain length, sugar composition and glycosidic linkages α(1 → 3) and α(1 → 2) play important parts in its anticoagulation activity, compared with the anticoagulant activity of sulphated glucose-containing oligosaccharides with β(1 → 4) and β(1 → 3) linkages [82]. PI-88 is known to consistently prolong the activated partial thromboplastin time through the

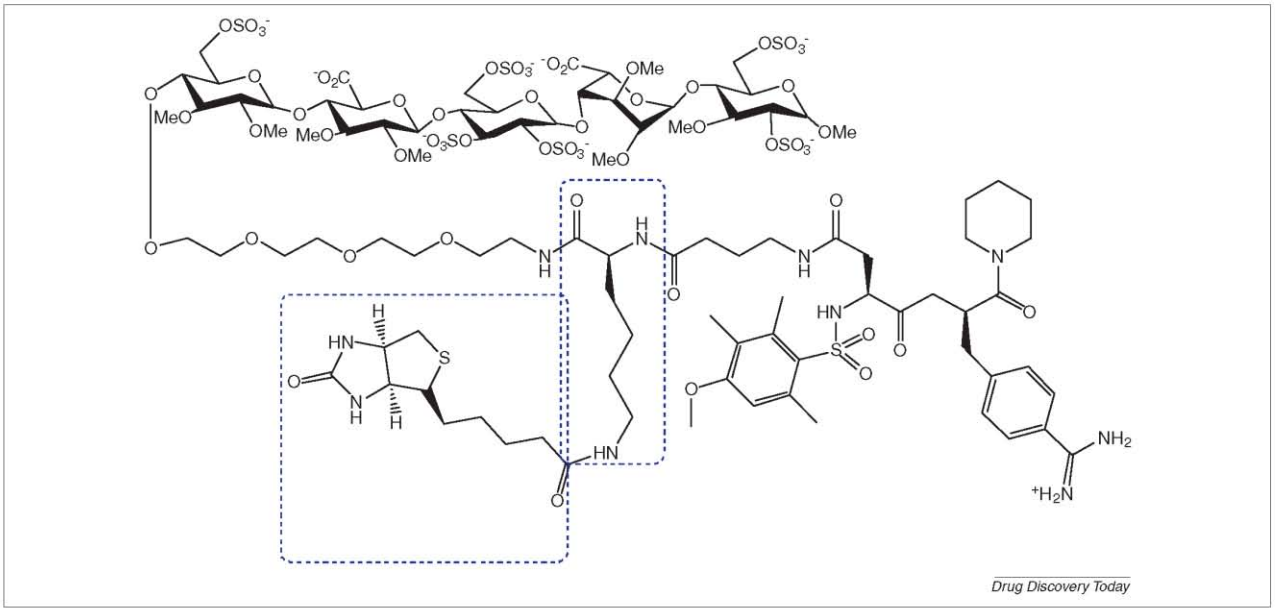


FIGURE 3

Chemical structures of EP217609 and EP42675. EP42675 and EP217609 are the first representatives of a new class of synthetic, parenteral anticoagulants with a dual mechanism of action combining the properties of an indirect FXa inhibitor and a direct thrombin inhibitor. The structure of EP42675 can be inferred by deleting the biotin and lysine moieties (shown in dotted rectangles) from the structure of EP217609.

activation of the endogenous heparin cofactor II. Apart from its anticipated anticoagulant effects, PI-88 was well tolerated in animal studies.

In the first phase I trial with patients with malignant disease, PI-88 was administered subcutaneously for four consecutive days either bimonthly or weekly [83]. Prolongation of the activated partial thromboplastin time was seen in only 2 of 14 patients. The recommended dose of PI-88 administered daily for four days every week was established to be 250 mg. Dose-limiting toxicity occasionally resulted in thrombocytopenia (at a dose of 2.28 mg/kg/day for 14 days) in patients with advanced malignancies, which seemed to be immunologically mediated through the development of anti-heparin platelet factor 4 complex antibodies [84]. The second phase I trial evaluated the safety, toxicity, pharmacological properties and biological activity of PI-88 with fixed weekly docetaxel (chemotherapy) in patients with advanced solid malignancies [85]. Sixteen patients received docetaxel at a 30 mg/m^2 dose on days 1, 8 and 15 of a 28 day cycle, with PI-88 injected subcutaneously for four days per week. Minor toxicity responses during the course of the therapy included fatigue, dysgeusia, thrombocytopenia, diarrhoea, nausea and emesis. Docetaxel and PI-88 did not alter the pharmacokinetics of each other. In another phase I trial, the recommended dose of PI-88 was reported to be 190 mg/m^2 alone and 1000 mg/m^2 in combination with dacarbazine every three weeks [86]. A phase I/II trial of daily PI-88 alone or with dacarbazine in patients with malignant melanoma was subsequently undertaken; however, the trial was stopped owing to cases of major febrile neutropenia [87]. A phase II trial of PI-88 in patients with advanced melanoma evaluated a fixed dose of 250 mg/day given by injection for four consecutive days followed by three drug-free days in a 28 day cycle [88]. Some patients developed serious bleeding events, with hemorrhagic cerebral metastases and arterial thrombosis. Nonetheless, in patients with advanced melanoma, PI-88 demonstrated noteworthy activity, but further investigations are needed of its use in combination with chemotherapy. A phase III trial investigating PI-88 as a post-resection treatment for hepatocellular carcinoma (liver cancer) was designed to establish the efficacy and safety of PI-88, but no results have been reported.

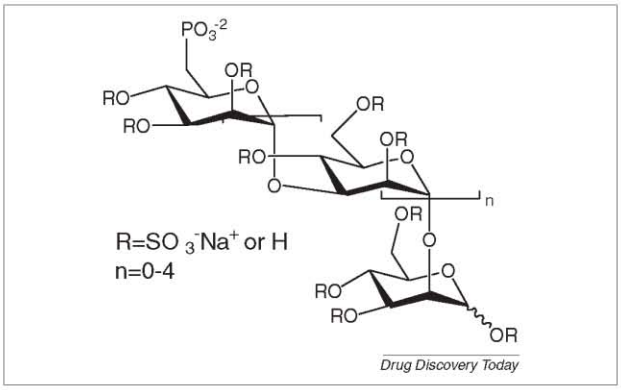


FIGURE 4

Chemical structure of PI-88. PI-88 is primarily composed of sulphated phosphomannopentaose and phosphomannotetraose oligosaccharide units. PI-88 is a potent anti-angiogenic, anti-tumour and anti-metastatic agent because of its inhibition of the heparan sulphate-degrading enzyme heparanase.

The PG500 series is a collection of newly designed compounds that are anomerically pure and fully sulphated and have single entity oligosaccharides attached to a lipophilic moiety, such as aglycone, at the reducing end of the molecule [89]. Compared with PI-88, some of these compounds are more potent inhibitors of

angiogenesis and metastasis and show strong anti-tumour activity in some aggressive tumour models [90]. These compounds are believed to interfere in processes such as tumour development, namely angiogenesis via inhibition of VEGF, FGF-1 and FGF-2, and metastasis via inhibition of heparanase [91]. PG545 was selected as the lead molecule based on its efficacy, pharmacokinetics, toxicology and ease of manufacture [92,93]. This compound has been in preclinical trials and administered subcutaneously once a week in mice for the treatment of cancer.

Tramiprosate (Alzhemed™)

HS/heparin have been widely reported to be associated with neuritic plaques in Alzheimer's disease (AD) [94]. HS has also been shown to promote the aggregation of amyloid β-peptide (Aβ) and have a pivotal role in plaque formation [95]. Several molecules have been proposed to be used to prevent HS-induced aggregation of Aβ: derivatives or fractions of heparin and other GAGs, sulphated compounds that act as HS mimetics (e.g. pentosan polysulphate and dextran sulphate) [96], small-molecule anionic sulphonates or sulphates [97], and amyloidophilic, sulphonated dyes, such as Congo Red and Thioflavin S.

Tramiprosate (also referred to as 3-amino-1-propanesulfonic acid, 3-aminopropylsulfonic acid, 3-APS, homotaurine or NC-531) is a GAG mimetic designed to interfere with the actions of Aβ early in the cascade of amyloidogenic events [98–100]. Structurally, Tramiprosate is a modification of the amino acid taurine (Fig. 5). It binds preferentially to soluble Aβ peptide and maintains Aβ in a random-coil/α-helical rich conformation and in non-fibrillar form, thereby inhibiting aggregation and hence plaque formation and deposition [101]. It can cross the blood–brain barrier effectively [102]. Recently, it has been reported that Tramiprosate also alters tau aggregation [103].

A phase II trial demonstrated that Tramiprosate reduces $A\beta_{42}$ in the cerebrospinal fluid of patients with mild to moderate AD [104]. The US phase III trial involved patients with mild to moderate AD, who were randomly assigned to receive placebo or 100 mg or 150 mg twice-daily doses of Tramiprosate. Although treatment was well tolerated, the study failed to demonstrate efficacy upon long-term clinical testing of cognitive improvement [105]. The European phase III trial has been discontinued. No further reports on the drug are available except columetric magnetic resonance imaging findings, which suggested less hippocampal shrinkage upon treatment with Tramiprosate [106]. Bellus Health (formerly Neurochem Inc.) has been promoting this medication as a nutraceutical, Vivimind™, which is being put forward as protecting against memory loss [107].

Bellus Health has recently initiated a phase I clinical trial of NRM8499, a prodrug of Tramiprosate for the treatment of AD. NRM8499 increases brain exposure to Tramiprosate, which might help improve the therapeutic effect on cognitive function and other clinical results in AD. This randomized, double-blind, placebo-controlled study is expected to investigate the safety, tolerability and pharmacokinetic profile of NRM8499 in a group of up to 84 young and elderly healthy subjects. Preclinical studies conducted in rodents showed that NRM8499 increased plasma and brain exposure to Tramiprosate by 1.5–3-fold.

Eprodisate sodium (NC-503, Kiacta® and Fibrillex™)

Eprodisate (1,3-propanedisulfonic acid disodium salt) is a low-molecular-weight, negatively charged sulphonated molecule (Fig. 6) that shares certain structural similarities with HS and is known to bind to serum amyloid protein A (SAA) [108]. Eprodisate binds to the GAG-binding site of SAA and competes with naturally occurring sulphated GAGs, thus targeting amyloid fibril polymerization and inhibiting amyloid deposition in tissues [97,109]. Eprodisate inhibits the development of amyloid deposits in *in vivo* mouse models of amyloid protein A (AA) amyloidosis [110]. In preclinical pharmacokinetic studies, Eprodisate has good bioavailability if administered orally; it is not metabolized, it does not bind to plasma proteins, and it is excreted primarily by the kidney, although pharmacokinetics analyses in its phase I trial revealed high inter-individual variability in its plasma concentrations [111]. Although Eprodisate is eliminated by the kidney, plasma concentrations were seen to increase as renal function decreased, resulting in a considerable increase in drug systemic exposure. An approximate terminal half-life of 10–20 h was derived from a multiple rising oral dose study. The efficacy and safety of Eprodisate was tested in a single phase II/III trial in AA amyloidosis patients [112]. Eprosidate was well tolerated, and its adverse events profile was comparable to placebo. Eprodisate can also be used with other types of amyloidosis. The results of a recent trial showed that it might slow the progression of AA amyloidosis-related renal disease [113], but no effect was seen on SAA levels, progression to end-stage renal disease or death, proteinuria and amyloid content of abdominal fat [114]. Despite having previously been granted orphan and fast-track status, the FDA and the EMEA both requested an additional confirmatory phase III trial before approval [114].

Studies using a preclinical rat model of diabetes and metabolic syndrome have confirmed that Eprodisate decreases glucose, cholesterol and triglycerides in the blood of obese diabetic Zucker rats compared with the control group, while preserving 40% more

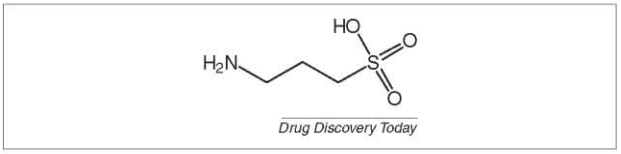


FIGURE 5

Chemical structure of Tramiprosate, an anti-amyloidogenic agent. In the context of Alzheimer's disease, this molecule acts by preventing and slowing the formation and the deposition of heparin/HS-induced amyloid fibrils in the brain and by binding to soluble beta-amyloid protein to reduce the amyloid-induced toxicity on neuronal and brain inflammatory cells.

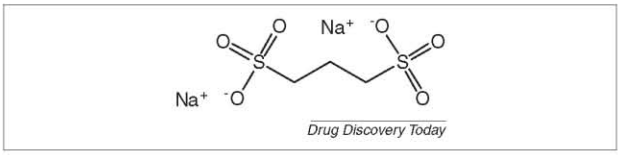


FIGURE 6

Chemical structure of Eprodisate. Eprodisate is a promising agent designed to prevent the worsening of renal function in patients with AA amyloidosis. It inhibits the polymerization of amyloid fibrils and the deposition of the fibrils in tissues by interfering with interactions between amyloidogenic proteins and heparin/HS.

pancreatic islet cells compared with the control group and showing some protective effect on renal function. However, Bellus Health has discontinued the development of Eprodisate in relation to diabetes after a phase IIa clinical proof-of-concept trial because the study did not meet its primary efficacy endpoint [115]. Instead, the company has pushed preclinical development of a prodrug of Eprodisate for the treatment of Type II diabetes and related metabolic syndromes.

Other novel GAG mimetics

Substrate-optimized glycans

Zacharon Pharmaceuticals has focused its research on small-molecule inhibitors of GAG biosynthesis for lysosomal storage disease. Mucopolysaccharidosis (MPS) is a form of lysosomal storage disease caused by a deficiency in the enzymes responsible for the degradation of GAGs, which makes lysosomes fill with partially degraded GAGs and resulting in serious systemic disease.

Zacharon uses substrate optimization therapy, which is a novel therapeutic approach for selectively modifying glycan structure without reducing the overall amount produced or altering normal glycan function [116]. This selective modification renders the glycan molecule more readily degradable, despite the presence of specific enzyme deficiencies. In MPS, this is accomplished by selectively and favourably modifying the glycan sulphation pattern. In MPS II, an inhibitor of the biosynthetic step involving the addition of 2-*O* sulphates to GAGs would produce GAGs with less 2-*O* sulphation and increased 6-*O* sulphation (Fig. 7). These GAGs would be easier to degrade for an MPS II (2-sulphatase-deficient) patient. To identify potential drug candidates, a library of 74,000 drug-like small molecules for inhibitors of HS biosynthesis were screened using cell-based assays. Of the 264 hit compounds identified in the primary screen, 30 were found to inhibit HS biosynthesis in cultured cells. Four of the compounds were found to reduce GAG accumulation in primary human fibroblasts obtained from MPS patients. ZP2345 was then chosen as the starting point for the development of a GAG mimetic based on substrate optimization therapy. ZP2345 is an HS inhibitor that reduces 2-*O* sulphation in a dose-dependent manner in a cultured human cell model of MPS-II [116]. Ongoing studies are focusing on analogue design, synthesis and testing to improve the potency and efficacy of these inhibitors.

Heparin mimetics in cancer

Exogenous heparin, LMWH and their mimetics have been shown to exert anti-metastatic and anti-angiogenic properties affecting cancer progression, such as inhibition of heparanase, blocking of P- and L-selectin-mediated cell adhesion, and inhibition of angiogenesis [117]. Heparanase has been targeted with several heparin-related inhibitors such as aza sugar derivatives, glycol-split derivatives and cyclitols that block the active site of the enzyme or the heparin/HS binding sites, or both [118]. Several LMWH derivatives have been described, among which are a deoxycholic acid conjugate and the fragmentation of a periodate-oxidized heparin, which have anti-angiogenic and anti-metastatic activities in different types of cancer models [119,120].

Endotis Pharma has created a platform for the development of anti-cancer 'small-glyco' drugs. These molecules are short, chemically synthesized oligosaccharides with potent affinity and selective inhibition of several growth factors and proteins (VEGF-A, FGF-2, PDGF-B, SDF-1α and heparanase) involved in tumour growth and dissemination [121]. A library of more than 100 synthetic oligosaccharides of different sizes containing various substitutions has been evaluated for their affinity for specific targets and their efficacy on cell proliferation and migration and *in vitro* endothelial tubule formation. The structure–activity relationship indicates that affinity and selectivity of these molecules for different targets can be fine-tuned through chemical substitutions. EP80061 is the lead compound in the series, and it induced a very potent anti-metastatic effect on a disseminated tumour model in C57Bl/6 mice [122]. However, the structure of these series has not yet been disclosed.

Momenta Pharmaceuticals presented preclinical data for M402, a HS mimetic containing a mixture of linear sugar chains and engineered to have potent anti-metastatic properties [123,124].

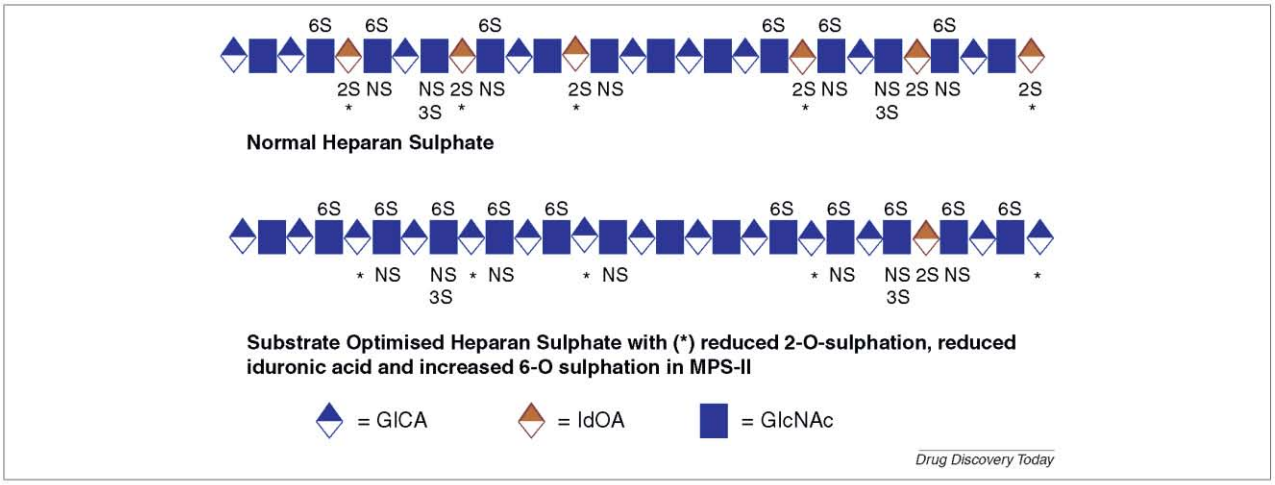


FIGURE 7

Substrate-optimized glycans (e.g. HS in MPS).Figure modified, with permission, from Ref. [159], Zacharon Pharmaceuticals.

Both *in vitro* and *in vivo*, M402 showed reduced anticoagulant activity and inhibited tumour metastasis through modulation of multiple factors, such as P-selectin, VEGF, FGF-2 and SDF-1α [124,125]. M402 as a monotherapy or in combination with chemotherapeutics showed statistically significant survival benefits in animal models with aggressive tumours [123,125]. In combination with gemcitabine, M402 produced much prolonged survival and reduced metastasis compared to groups treated with saline solution alone or gemcitabine alone in a murine pancreatic model [126]. Mice treated with M402 showed reduced epithelial-to-mesenchymal transition, a key step in the progression of tumour cells towards a more invasive phenotype.

Regenerating agents

Regenerating agents (RGTAs) are large biopolymers engineered to replace HS specifically bound to matrix proteins and growth factors destroyed after chronic tissue injury [127]. These polymers protect proteins bound to the extracellular matrix from proteolysis. RGTAs can interact with many heparin-binding growth factors, such as FGF-2 [128], transforming growth factor-β [129] and VEGF [130]. In addition to their heparin-binding-growth-factor-protecting and stabilizing properties, RGTAs have been found to inhibit human leukocyte elastase [131], plasmin [132,133] and heparanase [134]. RGTA-induced matrix therapy is a possible alternative to cell or gene therapy in regenerative medicine [135]. RGTA derivatives are potent activators of tissue repair in various *in vivo* wound-healing models: wound [136], bone defect [137,138], infarcted myocardium [139], colic ulceration [140] and periodontitis [141]. These RGTAs have also been shown to stimulate satellite cell growth and differentiation in primary cultures [142].

RGTAs are dextran derivatives with defined amounts of substituted carboxymethyl, benzylamide and sulfonate groups (Fig. 8). By varying the relative proportion of these substitutions, a library of heparin-mimetic biopolymers was produced. RGTAs with an increased level of sulphation and benzylamidation have shown anti-prion activity by blocking the conversion of prion protein PrP^C into the abnormal forms in scrapie-infected GT1 cells [143] and scrapie-infected and bovine-spongiform-encephalopathy-infected mice [144]. RGTA polymers are easier and less costly to produce, store and handle than growth factors. One such molecule is OTR4120 (alternatively called RGD120 or RG1192), derived from a glycosidic polymer of dextran T40 and functionalized with a 1.13 level of substitution of sulphate residues and a 0.46 level of substitution of carboxymethyl residues, with a maximal level of substitution of 3.0 [145], rendering this molecule structurally similar to heparin but having at least ten times less anticoagulant activity than heparin [146]. Nuclear magnetic resonance (NMR) analysis has shown that this polymeric compound is composed of a 15 sugar unit sequence statistically repeated along the molecule [147,148]. OTR4120 is known to enhance tissue repair in several animal models, including peripheral nerve injury in rats [149], burned skin in rats [150], chronic skin ulcers in mice [135] and cutaneous wound repair in rats [151]. Pharmacokinetics studies performed in a muscle crush model indicated that OTR4120 could replace degraded HS-GAG after tissue injury and bind to the heparin-binding sites present on many extracellular matrix proteins that have been freed from occupation by their endogenous GAGs [152]. In a recent clinical pilot study, an OTR4120 ophthalmic solution was found to improve the healing of severe corneal ulcers and dystrophy [153]. OTR3 is currently marketing CACIPLIQ20®, an active device based on RGTA for the treatment of chronic ulcers, diabetic foot ulcers, pressure ulcers, venous ulcers and burns.

Heptagonists

Use of heparin, LMWH, Fondaparinux and Idraparinux in cardiovascular surgeries often leads to a high incidence of bleeding complications. Protamine and LMW protamine are antidotes employed in heparin reversal; however, they can cause severe adverse reactions. PolyMedix has developed novel small synthetic salicylamide derivatives called heptagonists, which act as universal anticoagulant-reversing agents and are active against heparin, LMWH, Idraparinux and Fondaparinux [154–157]. One of the company's so-called 'heptagonists', PMX-60056 [157], can effectively neutralize the AT and anti-Xa activities of LMWH. PMX-50056 has been shown to completely reverse the anticoagulant effects of heparin and normalize blood clotting time in six human subjects in less than 10 min in a phase Ib clinical trial.

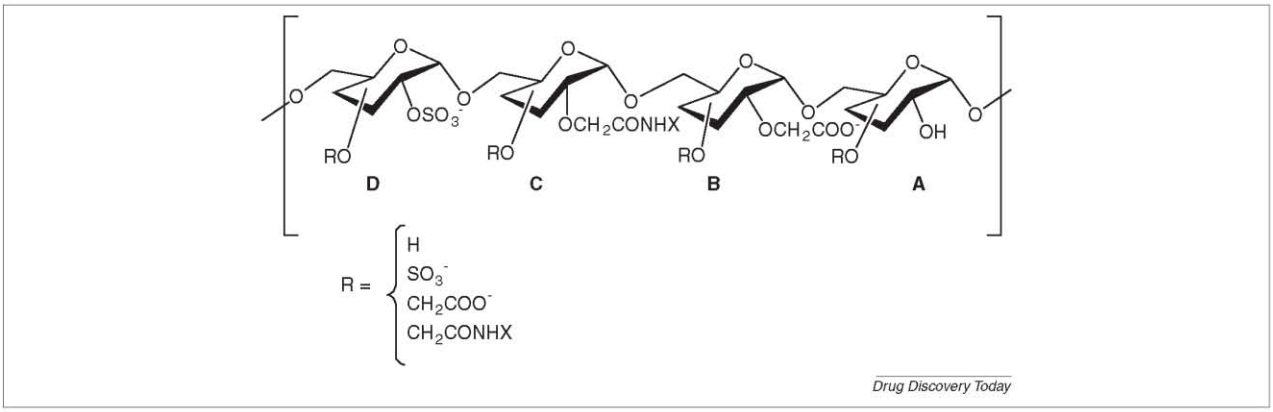


FIGURE 8

Schematic chemical structure of RGTA. Four differently substituted units, A (<1%), B (=32%), C (=0%) and D (=67%), can be present in OTR4120, as reported by titrimetry and ^{1}H NMR [147]. R represents the proportion of substituted group in the global C3 and C4 positions arranged to define the global dextran sulphate of each group.

Concluding remarks

Past research has highlighted the drawbacks of using native heparin oligosaccharides as drugs. Their anionic nature can result in large interactions with multiple, physiologically important proteins, leading to many side-effects. In addition to their lack of affinity, heparin oligosaccharides suffer from low tissue permeability, short serum half-life and poor stability. Consequently, the pharmacodynamic and pharmacokinetic properties of heparin make it inadequate for its direct therapeutic application. In addition, the multi-step synthesis of heparin/HS oligosaccharides poses challenges for medicinal chemists, both at the drug development and the production scale [158]. Furthermore, new therapeutic applications of sulphated GAGs now include the treatment of infectious diseases and inflammation and the control of cell growth in wound healing and cancer. These new applications require the elimination of the anticoagulant activity of heparin oligosaccharides and the engineering of appropriate pharmacokinetic properties and optimal oral bioavailability.

GAG mimetics are designed to overcome these shortcomings. Detailed insight into GAG–protein interactions has predominantly been provided by recent progress in NMR spectroscopy, X-ray crystallography and molecular modelling techniques. Identification of the bound conformation of heparin/HS to a protein enables the design of GAG mimetics and the identification of negligible and replaceable functional groups. As a consequence, the development of GAG mimetics that have improved absorption, distribution, metabolism and excretion properties can be accomplished.

The development of HS/heparin-based drugs is a fertile field of research that is providing enormous opportunities for the discovery of improved treatments for many diseases. This is evidenced by the existence of many newly established companies, such as Intellihep, Zacharon Pharmaceuticals, GlycoMimetics, Endotis Pharma, Polymedix, Progen, OTR3 and Momenta, to name a few, which are beginning to exploit the untapped potential of the structural diversity of heparin/HS in various therapeutic and biomedical applications.

Acknowledgements

N.S.G. is grateful for the award of an Endeavour International Postgraduate Research Scholarship. We apologize to all those authors whose work could not be cited owing to space limitations.

References

1 Lever, R. *et al.* (2001) Role of glycosaminoglycans in inflammation. *Inflammopharmacology* 9, 165–169
2 Parish, C.R. (2006) The role of heparan sulphate in inflammation. *Nat. Rev. Immunol.* 6, 633–643
3 Parish, C.R. (2005) Heparan sulfate and inflammation. *Nat. Immunol.* 6, 861–862
4 Díaz-Nido, J. *et al.* (2002) Glycosaminoglycans and β-amyloid, prion and tau peptides in neurodegenerative diseases. *Peptides* 23, 1323–1332
5 Iozzo, R.V. and San Antonio, J.D. (2001) Heparan sulfate proteoglycans: heavy hitters in the angiogenesis arena. *J. Clin. Invest.* 108, 349–355
6 Rosenberg, R.D. *et al.* (1997) Heparan sulfate proteoglycans of the cardiovascular system. Specific structures emerge but how is synthesis regulated? *J. Clin. Invest.* 99, 2062–2070
7 Yip, G.W. *et al.* (2006) Therapeutic value of glycosaminoglycans in cancer. *Mol. Cancer Ther.* 5, 2139–2148
8 Rostand, K.S. and Esko, J.D. (1997) Microbial adherence to and invasion through proteoglycans. *Infect. Immun.* 65, 1–8
9 Sawitzky, D. (1996) Protein–glycosaminoglycan interactions: infectiological aspects. *Med. Microbiol. Immunol. (Berl.)* 184, 155–161
10 Wadstrom, T. and Ljungh, A.S.A. (1999) Glycosaminoglycan-binding microbial proteins in tissue adhesion and invasion: key events in microbial pathogenicity. *J. Med. Microbiol.* 48, 223–233
11 Gandhi, N.S. and Mancera, R.L. (2008) The structure of glycosaminoglycans and their interactions with proteins. *Chem. Biol. Drug Des.* 72, 455–482
12 Fuster, M.M. and Esko, J.D. (2005) The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat. Rev. Cancer* 5, 526–542
13 Brown, J.R. *et al.* (2007) Glycan antagonists and inhibitors: a fount for drug discovery. *Crit. Rev. Biochem. Mol. Biol.* 42, 481–515
14 Shriver, Z. *et al.* (2004) Glycomics: a pathway to a class of new and improved therapeutics. *Nat. Rev. Drug Discov.* 3, 863–873
15 Osborn, H.M.I. *et al.* (2004) Carbohydrate-based therapeutics. *J. Pharm. Pharmacol.* 56, 691–702
16 Gesslbauer, B. and Kungl, A.J. (2006) Glycomic approaches toward drug development: therapeutically exploring the glycosaminoglycanome. *Curr. Opin. Mol. Ther.* 8, 521–528
17 Volpi, N. (2006) Therapeutic applications of glycosaminoglycans. *Curr. Med. Chem.* 13, 1799–1810
18 Lindahl, U. (2007) Heparan sulfate–protein interactions: a concept for drug design? *Thromb. Haemost.* 98, 109–115
19 Fugedi, P. (2003) The potential of the molecular diversity of heparin and heparan sulfate for drug development. *Mini Rev. Med. Chem.* 3, 659–667
20 Eikelboom, J.W. and Weitz, J.I. (2010) New anticoagulants. *Circulation* 121, 1523–1532
21 Samama, M.M. and Gerotziafas, G. (2010) Newer anticoagulants in 2009. *J. Thromb. Thrombolysis* 29, 92–104
22 Klement, P. and Rak, J. (2006) Emerging anticoagulants: mechanism of action and future potential. *Vnitr. Lek.* 52 (Suppl. 1), 119–122
23 Bauer, K.A. (2002) Selective inhibition of coagulation factors: advances in antithrombotic therapy. *Semin. Thromb. Hemost.* 28 (Suppl. 2), 15–24
24 De Kort, M. *et al.* (2005) Synthetic heparin derivatives as new anticoagulant drugs. *Drug Discov. Today* 10, 769–779
25 Alban, S. (2008) Natural and synthetic glycosaminoglycans. Molecular characteristics as the basis of distinct drug profiles. *Hamostaseologie* 28, 51–61
26 Atha, D.H. *et al.* (1985) Contribution of monosaccharide residues in heparin binding to antithrombin III. *Biochemistry* 24, 6723–6729
27 Oosta, G.M. *et al.* (1981) Multiple functional domains of the heparin molecule. *Proc. Natl. Acad. Sci. U. S. A.* 78, 829–833
28 Petitou, M. and v Boeckel, C.A.A. (2004) A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next? Angew *Chem. Int. Ed. Engl.* 43, 3118–3133
29 Van Boeckel, C.A.A. and Petitou, M. (1993) The unique antithrombin III binding domain of heparin: a lead to new synthetic antithrombotics. *Angew. Chem. Int. Ed. Engl.* 32, 1671–1690
30 Jairajpuri, M.A. *et al.* (2003) Antithrombin III phenylalanines 122 and 121 contribute to its high affinity for heparin and its conformational activation. *J. Biol. Chem.* 278, 15941–15950
31 Petitou, M. *et al.* (2009) From heparin to EP217609: the long way to a new pentasaccharide-based neutralisable anticoagulant with an unprecedented pharmacological profile. *Thromb. Haemost.* 102, 804–810
32 Avci, F.Y. *et al.* (2003) Synthetic oligosaccharides as heparin-mimetics displaying anticoagulant properties. *Curr. Pharm. Des.* 9, 2323–2335
33 Turpie, A.G.G. (2004) Fondaparinux: a Factor Xa inhibitor for antithrombotic therapy. *Expert Opin. Pharmacother.* 5, 1373–1384
34 Xu-song, Z. and Bing-ren, X. (2009) Discontinued drugs in 2008: cardiovascular drugs. *Expert Opin. Investig. Drugs* 18, 875–885
35 Herbert, J.M. *et al.* (1998) Biochemical and pharmacological properties of SANORG 34006, a potent and long-acting synthetic pentasaccharide. *Blood* 91, 4197–4205
36 Hjelm, R. and Schedin-Weiss, S. (2007) High affinity interaction between a synthetic, highly negatively charged pentasaccharide and alpha- or beta-antithrombin is predominantly due to nonionic interactions. *Biochemistry* 46, 3378–3384
37 Desai, U.R. *et al.* (1998) Mechanism of heparin activation of antithrombin: evidence for an induced-fit model of allosteric activation involving two interaction subsites. *Biochemistry* 37, 13033–13041

38 Desai, U.R. *et al.* (1998) Mechanism of heparin activation of antithrombin: role of individual residues of the pentasaccharide activating sequence in the recognition of native and activated states of antithrombin. *J. Biol. Chem.* 273, 7478–7487

39 Westerduin, P. *et al.* (1994) Feasible synthesis and biological properties of six 'non-glycosamino' glycan analogues of the antithrombin III binding heparin pentasaccharide. *Bioorg. Med. Chem.* 2, 1267–1280

40 Basten, J. *et al.* (1992) Biologically active heparin-like fragments with a "non-glycosamino"glycan structure. Part 2: a tetra-o-methylated pentasaccharide with high affinity for antithrombin III. *Bioorg. Med. Chem. Lett.* 2, 901–904

41 Jin, L. *et al.* (1997) The anticoagulant activation of antithrombin by heparin. *Proc. Natl. Acad. Sci. U. S. A.* 94, 14683–14688

42 McCoy, A.J. *et al.* (2003) Structure of β-antithrombin and the effect of glycosylation on antithrombin's heparin affinity and activity. *J. Mol. Biol.* 326, 823–833

43 Prandoni, P. *et al.* (2008) Idraparinux: review of its clinical efficacy and safety for prevention and treatment of thromboembolic disorders. *Expert Opin. Investig. Drugs* 17, 773–777

44 Faaij, R.A. *et al.* (1999) A phase I single rising dose study to investigate the safety, tolerance and pharmacokinetics of subcutaneous SANORG 34006 in healthy male and female elderly volunteers. *Thromb. Haemost.* 490–491 (Abstract 1547)

45 Faaij, R.A. *et al.* (1999) A phase I single rising dose study to investigate the safety, tolerance and pharmacokinetics of SANORG 34006 in healthy young male volunteers. *Thromb. Haemost.* 853–1853 (Abstract 2709)

46 Ma, Q. and Fareed, J. (2004) Idraparinux sodium. *IDrugs* 7, 1028–1034

47 (2002) PERSIST investigators. A novel long-acting synthetic factor Xa inhibitor (idraparinux sodium) to replace warfarin for secondary prevention in deep vein thrombosis. A phase II evaluation. *Blood* 100, 301

48 Buller, H.R. *et al.* (2004) A novel long-acting synthetic factor Xa inhibitor (SanOrg34006) to replace warfarin for secondary prevention in deep vein thrombosis. A phase II evaluation. *J. Thromb. Haemost.* 2, 47–53

49 Minar, E. and Investigators, T.P. (2004) A novel long-acting synthetic factor Xa inhibitor (SanOrg34006) to replace warfarin for secondary prevention in deep vein thrombosis. A phase II evaluation. *J. Thromb. Haemost.* 2, 540

50 Buller, H.R. *et al.* (2007) Extended prophylaxis of venous thromboembolism with idraparinux. *N. Engl. J. Med.* 357, 1105–1112

51 Buller, H.R. *et al.* (2007) Idraparinux versus standard therapy for venous thromboembolic disease. *N. Engl. J. Med.* 357, 1094–1104

52 Harenberg, J. *et al.* (2008) Anticoagulant effects of Idraparinux after termination of therapy for prevention of recurrent venous thromboembolism: observations from the van Gogh trials. *Eur. J. Clin. Pharmacol.* 64, 555–563

53 Harenberg, J. *et al.* (2008) Long elimination half-life of idraparinux may explain major bleeding and recurrent events of patients from the van Gogh trials. *J. Thromb. Haemost.* 6, 890–892

54 Harenberg, J. *et al.* (2009) The anticoagulant Idraparinux: is the extensive half life of 60 days the cause of bleeding complications. *Br. J. Clin. Pharmacol.* 68 (Suppl. 1), 21–121

55 Bijsterveld, N.R. *et al.* (2004) Recombinant factor VIIa reverses the anticoagulant effect of the long-acting pentasaccharide idraparinux in healthy volunteers. *Br. J. Haematol.* 124, 653–658

56 Veyrat-Follet, C. *et al.* (2009) The pharmacokinetics of idraparinux, a long-acting indirect factor Xa inhibitor: population pharmacokinetic analysis from Phase III clinical trials. *J. Thromb. Haemost.* 7, 559–565

57 (2008) The AMADEUS Investigators. Comparison of idraparinux with vitamin K antagonists for prevention of thromboembolism in patients with atrial fibrillation: a randomised, open-label, non-inferiority trial. *Lancet* 371, 315–321

58 Harenberg, J. (2009) Development of idraparinux and idrabiotaparinux for anticoagulant therapy. *Thromb. Haemost.* 102, 811–815

59 Savi, P. *et al.* (2008) Reversible biotinylated oligosaccharides: a new approach for a better management of anticoagulant therapy. *J. Thromb. Haemost.* 6, 1697–1706

60 Sobieraj-Teague, M. *et al.* (2009) New anticoagulants for atrial fibrillation. *Semin. Thromb. Hemost.* 35, 515–524

61 Viskov, C. *et al.* (2009) Description of the chemical and pharmacological characteristics of a new hemisynthetic ultra-low-molecular-weight heparin, AVE5026. *J. Thromb. Haemost.* 7, 1143–1151

62 Hoppensteadt, D. *et al.* (2008) AVE5026: a new hemisynthetic ultra low molecular weight heparin (ULMWH) with enriched anti-Xa activity and enhanced antithrombotic activity for management of cancer associated thrombosis. *J. Clin. Oncol.* 26 (15 Suppl.), 14653 (Meeting Abstracts)

63 Lassen, M.R. *et al.* (2009) AVE5026, a new hemisynthetic ultra-low-molecular-weight heparin for the prevention of venous thromboembolism in patients after total knee replacement surgery – TREK: a dose-ranging study. *J. Thromb. Haemost.* 7, 566–572

64 Melloni, C. *et al.* (2009) Design and rationale of the evaluation of M118 in percutaneous coronary intervention (EMINENCE) trial. *Am. Heart J.* 158, 726–733

65 Kishimoto, T.K. *et al.* (2009) M118-A rationally engineered low-molecular-weight heparin designed specifically for the treatment of acute coronary syndromes. *Thromb. Haemost.* 102, 900–906

66 Draganov, D. *et al.* (2009) Pharmacokinetics of M118, unfractionated heparin and enoxaparin sodium in normal and 5/6 nephrectomized uremic rats. *Toxicol. Lett.* 189 (Suppl. 1), S113–S1113

67 Volovyk, Z. *et al.* (2009) A rationally designed heparin, M118, has anticoagulant activity similar to unfractionated heparin and different from Lovenox in a cell-based model of thrombin generation. *J. Thromb. Thrombolysis* 28, 132–139

68 Fier, I. *et al.* (2007) A novel, rationally engineered heparin (M118) prevents thrombosis more effectively than unfractionated heparin in a canine model of deep arterial injury. *J. Am. Coll. Cardiol.* 49, 379A–380A

69 Chakrabarti, S. *et al.* (2009) M118, a novel low-molecular weight heparin with decreased polydispersity leads to enhanced anticoagulant activity and thrombotic occlusion in ApoE knockout mice. *J. Thromb. Thrombolysis* 28, 394–400

70 Fier, I.D. *et al.* (2009) Lack of pharmacokinetic and pharmacodynamic interactions between M118, a novel low-molecular-weight-heparin and Eptifibatide in healthy subjects. *J. Clin. Pharmacol.* 49, 73

71 Bal Dit Sollier, C. *et al.* (2009) Anticoagulant activities of EP42675 – synthetic direct inhibitor and indirect factor Xa inhibitor. In *Proceedings of the XXII Congress of the International Society of Thrombosis and Haemostasis*

72 De Kort, M. and Van Boeckel, C.A.A. (2010) *Antithrombotic Dual Inhibitors Comprising a Biotin Residue*. N.V. ORGANON (Oss, NL)

73 Bal Dit Sollier, C. *et al.* (2009) Pharmacokinetics and pharmacodynamics of EP42675 a new synthetic anticoagulant with a dual mechanism of action. In *Proceedings of the XII congress of the International Society of Thrombosis and Haemostasis*

74 Ludwig, R.J. (2009) Therapeutic use of heparin beyond anticoagulation. *Curr. Drug Discov. Technol.* 6, 281–289

75 Young, E. (2008) The anti-inflammatory effects of heparin and related compounds. *Thromb. Res.* 122, 743–752

76 Borsig, L. (2010) Antimetastatic activities of heparins and modified heparins. Experimental evidence. *Thromb. Res.* 125 (Suppl. 2), S66–S71

77 McKenzie, E.A. (2007) Heparanase: a target for drug discovery in cancer and inflammation. *Br. J. Pharmacol.* 151, 1–14

78 Vlodavsky, I. and Friedmann, Y. (2001) Molecular properties and involvement of heparanase in cancer metastasis and angiogenesis. *J. Clin. Invest.* 108, 341–347

79 Miao, H.-Q. *et al.* (2006) Development of heparanase inhibitors for anti-cancer therapy. *Curr. Med. Chem.* 13, 2101–2111

80 Kudchadkar, R. *et al.* (2008) PI-88: a novel inhibitor of angiogenesis. *Expert Opin. Investig. Drugs* 17, 1769–1776

81 Ferro, V. and Don, R. (2003) The development of the novel angiogenesis inhibitor PI-88 as an anticancer drug. *Australas. Biotechnol.* 13, 38–39

82 Wall, D. *et al.* (2001) Characterisation of the anticoagulant properties of a range of structurally diverse sulfated oligosaccharides. *Thromb. Res.* 103, 325–335

83 Basche, M. *et al.* (2006) A phase I biological and pharmacologic study of the heparanase inhibitor PI-88 in patients with advanced solid tumors. *Clin. Cancer Res.* 12, 5471–5480

84 Rosenthal, M.A. *et al.* (2002) Treatment with the novel anti-angiogenic agent PI-88 is associated with immune-mediated thrombocytopenia. *Ann. Oncol.* 13, 770–776

85 Chow, L.Q. *et al.* (2008) A phase I pharmacological and biological study of PI-88 and docetaxel in patients with advanced malignancies. *Cancer Chemother. Pharmacol.* 63, 65–74

86 Millward, M. *et al.* (2007) Final results of a phase I study of daily PI-88 as a single agent and in combination with dacarbazine (D) in patients with metastatic melanoma. *J. Clin. Oncol.* 25 (Suppl. 18), 8532 (Meeting Abstracts)

87 Khasraw, M. *et al.* (2009) Multicentre phase I/II study of PI-88, a heparanase inhibitor in combination with docetaxel in patients with metastatic castrate-resistant prostate cancer. *Ann. Oncol.* 21, 1302–1307

88 Lewis, K.D. *et al.* (2008) A phase II study of the heparanase inhibitor PI-88 in patients with advanced melanoma. *Invest. New Drugs* 26, 89–94

89 Karoli, T. *et al.* (2005) Synthesis, biological activity, and preliminary pharmacokinetic evaluation of analogues of a phosphosulfomannan angiogenesis inhibitor (PI-88). *J. Med. Chem.* 48, 8229–8236

90 Ferro, V. *et al.* (2007) PI-88 and novel heparan sulfate mimetics inhibit angiogenesis. *Semin. Thromb. Hemost.* 33, 557–568

91 Dredge, K. *et al.* (2009) The PG500 series: novel heparan sulfate mimetics as potent angiogenesis and heparanase inhibitors for cancer therapy. *Invest. New Drugs* 28, 276–283

92 Bytheway, I. *et al.* (2009) The dual angiogenesis/heparanase inhibitor PG545, but not the tyrosine kinase inhibitor sorafenib, inhibits spontaneous metastasis in models of breast and lung cancer. In *Proceedings of the AACR-NCI-EORTC International Conference: Molecular Targets and Cancer Therapeutics, Mol. Cancer Ther. (Meeting Abstract Supplement)*

93 Hammond, E. *et al.* (2009) The dual angiogenesis/heparanase inhibitor PG545 inhibits solid tumor progression in models of breast, prostate and liver cancer: a comparative assessment of once versus twice weekly administration schedules. In *Proceedings of the AACR-NCI-EORTC International Conference: Molecular Targets and Cancer Therapeutics, Mol. Cancer Ther. (Meeting Abstract Supplement)*

94 van Horssen, J. *et al.* (2003) Heparan sulphate proteoglycans in Alzheimer's disease and amyloid-related disorders. *Lancet Neurol.* 2, 482–492

95 McLaurin, J.A. *et al.* (1999) Interactions of Alzheimer amyloid-β peptides with glycosaminoglycans. *Eur. J. Biochem.* 266, 1101–1110

96 Leveugle, B. *et al.* (1994) Binding of heparan sulfate glycosaminoglycan to [beta]-amyloid peptide: inhibition by potentially therapeutic polysulfated compounds. *Neuroreport* 5, 1389–1392

97 Kisilevsky, R. *et al.* (1995) Arresting amyloidosis in vivo using small-molecule anionic sulphonates or sulphates: implications for Alzheimer's disease. *Nat. Med.* 1, 143–148

98 Tremblay, P. *et al.* (2005) Functional GAG mimetics as an approach for the treatment of amyloid diseases. *Alzheimers Dement.* 1 (Suppl. 1), S2–S12

99 Aisen, P.S. *et al.* (2007) Alzhemed: a potential treatment for Alzheimers disease. *Curr. Alzheimer Res.* 4, 473–478

100 Geerts, H. (2004) NC-531 (Neurochem). *Curr. Opin. Investig. Drugs* 5, 95–100

101 Wright, T.M. (2006) Tramiprosate. *Drugs Today (Barc)* 42, 291–298

102 Gervais, F. *et al.* (2007) Targeting soluble Aβ peptide with Tramiprosate for the treatment of brain amyloidosis. *Neurobiol. Aging* 28, 537–547

103 Santa-Maria, I. *et al.* (2007) Tramiprosate, a drug of potential interest for the treatment of Alzheimer's disease, promotes an abnormal aggregation of tau. *Mol. Neurodegener.* 2, 17

104 Aisen, P.S. *et al.* (2006) A Phase II study targeting amyloid-β with 3APS in mild-to-moderate Alzheimer disease. *Neurology* 67, 1757–1763

105 Rafii, M.S. and Aisen, P. (2009) Recent developments in Alzheimer's disease therapeutics. *BMC Med.* 7, 7

106 Saumier, D. *et al.* (2009) Lessons learned in the use of volumetric MRI in therapeutic trials in Alzheimer's disease: the Alzhemed™ (Tramiprosate) experience. *J. Nutr. Health Aging* 13, 370–372

107 Neugroschl, J. and Sano, M. (2010) Current treatment and recent clinical research in Alzheimer's disease. *Mt. Sinai J. Med.* 77, 3–16

108 Revill, P. *et al.* (2006) Eprodisate sodium. *Drugs Future* 31, 576–578

109 Ancsin, J.B. and Kisilevsky, R. (1999) The heparin/heparan sulfate-binding site on apo-serum Amyloid A. *J. Biol. Chem.* 274, 7172–7181

110 Gervais, F. *et al.* (2003) Proteoglycans and amyloidogenic proteins in peripheral amyloidosis. *Curr. Med. Chem. Immunol. Endocr. Metab. Agents* 3, 361–370

111 Kisilevsky, R. (2000) The relation of proteoglycans, serum amyloid P and Apo E to amyloidosis current status, 2000. *Amyloid* 7, 23–25

112 Clinicaltrials.gov. (2002) A phase II/III study of the safety and efficacy of NC-503 in patients suffering from secondary (AA) amyloidosis (NCT00035334).

113 Dember, L.M. *et al.* (2007) Eprodisate for the treatment of renal disease in AA amyloidosis. *N. Engl. J. Med.* 356, 2349–2360

114 Manenti, L. *et al.* (2008) Eprodisate in amyloid A amyloidosis: a novel therapeutic approach? *Expert Opin. Pharmacother.* 9, 2175–2180

115 Bellini, R. (2010) *BELLUS Health Ends NC-503 Diabetes Development Program following Results*. Bellus Health Inc.

116 Brown, J. *et al.* (2010) Small molecule inhibitors of glycosaminoglycan biosynthesis as substrate optimization therapy for the mucopolysaccharidoses. In *Proceedings of the Lysosomal Disease Network WORLD Symposium, vol. 99* pp. S12–S112

117 Casu, B. *et al.* (2010) Heparin-derived heparan sulfate mimics to modulate heparan sulfate–protein interaction in inflammation and cancer. *Matrix Biol.* 29, 442–452

118 Vlodavsky, I. *et al.* (2007) Heparanase: structure, biological functions, and inhibition by heparin-derived mimetics of heparan sulfate. *Curr. Pharm. Des.* 13, 2057–2073

119 Mousa, S.A. *et al.* (2006) Anti-metastatic effect of a non-anticoagulant low-molecular-weight heparin versus the standard low-molecular-weight heparin, enoxaparin. *Thromb. Haemost.* 96, 816–821

120 Lee, D.Y. *et al.* (2009) Antiangiogenic activity of orally absorbable heparin derivative in different types of cancer cells. *Pharm. Res.* 26, 2667–2676

121 Cabannes, E. *et al.* (2009) Heparan sulfate mimetics as anticancer small-glyco drugs. In *Proceedings of the 67th Harden Conference*

122 Serina, G. *et al.* (2010) Antitumor activity of EP80061, a small-glyco drug in preclinical studies. In *Proceedings of the AACR meeting*

123 Zhou, H. *et al.* (2010) *M402 – A Novel Heparan Sulfate Proteoglycan Mimetic Targeting Tumor–Host Interactions*. American Association for Cancer Research (AACR)

124 Zhou, H. *et al.* (2009) M-ONC 402-a non anticoagulant low molecular weight heparin inhibits tumor metastasis. In *Proceedings of the 100th Annual Meeting of American Association for Cancer Research (AACR)*

125 Chu, C. *et al.* (2009) M-ONC 402, A novel non-anticoagulant heparin, inhibits P-Selectin function and metastatic seeding of tumor cells in mice. In *Proceedings of the100th Annual Meeting of American Association for Cancer Research (AACR)*

126 Lolkema, M.P. *et al.* (2010) *M402, A Novel Heparan Sulfate Mimetic, Synergizes with Gemcitabine to Improve Survival and Reduce Metastasis and Epithelial-to-mesenchymal Transition (EMT) in a Genetically Engineered Mouse Model for Pancreatic Cancer*. American Association for Cancer Research (AACR)

127 Barritault, D. and Caruelle, J.P. (2006) Regenerating agents (RGTAs): a new therapeutic approach. *Ann. Pharm. Fr.* 64, 135–144

128 Tardieu, M. *et al.* (1992) Derivatized dextrans mimic heparin as stabilizers, potentiators, and protectors of acidic or basic FGF. *J. Cell. Physiol.* 150, 194–203

129 Meddahi, A. *et al.* (1996) Heparin-like polymers derived from dextran enhance colonic anastomosis resistance to leakage. *J. Biomed. Mater. Res.* 31, 293–297

130 Rouet, V. *et al.* (2005) A synthetic glycosaminoglycan mimetic minds vascular endothelial growth factor and modulates angiogenesis. *J. Biol. Chem.* 280, 32792–32800

131 Meddahi, A. *et al.* (1996) FGF protection and inhibition of human neutrophil elastase by carboxymethyl benzylamide sulfonate dextran derivatives. *Int. J. Biol. Macromol.* 18, 141–145

132 Meddahi, A. *et al.* (1995) Inhibition by dextran derivatives of FGF-2 plasmin-mediated degradation. *Biochimie* 77, 703–706

133 Ledoux, D. *et al.* (2000) Human plasmin enzymatic activity is inhibited by chemically modified dextrans. *J. Biol. Chem.* 275, 29383–29390

134 Rouet, V. *et al.* (2006) Heparin-like synthetic polymers, named RGTAs, mimic biological effects of heparin in vitro. *J. Biomed. Mater. Res. A* 78, 792–797

135 Barbier-Chassefière, V. *et al.* (2009) Matrix therapy in regenerative medicine, a new approach to chronic wound healing. *J. Biomed. Mater. Res. A* 90, 641–647

136 Meddahi, A. *et al.* (1994) New approaches to tissue regeneration and repair. *Pathol. Res. Pract.* 190, 923–928

137 Albo, D. *et al.* (1996) Modulation of cranial bone healing with a heparin-like dextran derivative. *J. Craniofac. Surg.* 7, 19–22

138 Blanquaert, F. *et al.* (1995) Heparan-like molecules induce the repair of skull defects. *Bone* 17, 499–506

139 Yamauchi, H. *et al.* (2000) New agents for the treatment of infarcted myocardium. *FASEB J.* 14, 2133–2134

140 Meddahi, A. *et al.* (2002) Heparin-like polymer improved healing of gastric and colic ulceration. *J. Biomed. Mater. Res.* 60, 497–501

141 Escartin, Q. *et al.* (2003) A new approach to treat tissue destruction in periodontitis with chemically modified dextran polymers. *FASEB J.* 17, 644–651

142 Papy-Garcia, D. *et al.* (2002) Glycosaminoglycan mimetics (RGTA) modulate adult skeletal muscle satellite cell proliferation in vitro. *J. Biomed. Mater. Res.* 62, 46–55

143 Schonberger, O. *et al.* (2003) Novel heparan mimetics potently inhibit the scrapie prion protein and its endocytosis. *Biochem. Biophys. Res. Commun.* 312, 473–479

144 Adjou, K.T. *et al.* (2003) A novel generation of heparan sulfate mimetics for the treatment of prion diseases. *J. Gen. Virol.* 84, 2595–2603

145 Barbosa, I. *et al.* (2005) A synthetic glycosaminoglycan mimetic (RGTA) modifies natural glycosaminoglycan species during myogenesis. *J. Cell Sci.* 118, 253–264

146 Aamiri, A. *et al.* (1995) Effect of a substituted dextran on reinnervation during regeneration of adult rat skeletal muscle. *C. R. Acad. Sci. III. Sci. Vie* 318, 1037–1044

147 Papy-Garcia, D. *et al.* (2005) Nondegradative sulfation of polysaccharides. Synthesis and structure characterization of biologically active heparan sulfate mimetics. *Macromolecules* 38, 4647–4654

148 Martelly, I. *et al.* (2010) Glycosaminoglycan mimetics trigger IP3-dependent intracellular calcium release in myoblasts. *Matrix Biol.* 29, 317–329

149 Zuijdendorp, H.M. *et al.* (2008) Significant reduction in neural adhesions after administration of the regenerating agent OTR4120, a synthetic glycosaminoglycan mimetic, after peripheral nerve injury in rats. *J. Neurosurg.* 109, 967–973

150 Garcia-Filipe, S. *et al.* (2007) RGTA OTR4120, a heparan sulfate mimetic, is a possible long-term active agent to heal burned skin. *J. Biomed. Mater. Res. A* 80, 75–84

151 Tong, M. *et al.* (2009) Stimulated neovascularization, inflammation resolution and collagen maturation in healing rat cutaneous wounds by a heparan sulfate glycosaminoglycan mimetic, OTR4120. *Wound Repair Regen.* 17, 840–852

152 Meddahi, A. *et al.* (2002) Pharmacological studies of $RGTA_{11}$, a heparan sulfate mimetic polymer, efficient on muscle regeneration. *J. Biomed. Mater. Res.* 62, 525–531
153 Chebbi, C.K. *et al.* (2008) Pilot study of a new matrix therapy agent (RGTA OTR4120®) in treatment-resistant corneal ulcers and corneal dystrophy. *J. Fr. Opthalmol.* 31, 465–471
154 Jeske, W. *et al.* (2007) *In Vitro Characterization of the Neutralization of Unfractionated Heparin and Low Molecular Weight Heparin by Novel Salicylamide Derivatives.* American Society of Hematology
155 Kuziej, J. *et al.* (2009) Neutralization of hemorrhagic and antithrombotic activities of heparins by a novel salicylamide derivative. *FASEB J.* 23, 566–569 (Meeting Abstracts 1)
156 Fareed, J. *et al.* (2008) Neutralization of the anticoagulant and anti-Xa effects of fondaparinux and idraparinux by a novel synthetic antagonist. Pharmacologic implications. *FASEB J.* 22, 1117–1118 (Meeting Abstracts 1)
157 Jeske, W. *et al.* (2009) *Novel Antagonists for Low Molecular Weight Heparin and Heparin-like Drugs.* American Society of Hematology
158 Codée, J.D.C. *et al.* (2004) The synthesis of well-defined heparin and heparan sulfate fragments. *Drug Discov Today: Technol.* 1, 317–326
159 Brown, J. *et al.* (2010) Small molecule inhibitors of glycosaminoglycan biosynthesisas substrate optimization therapy for the mucopolysaccharidoses. *Mol. Genet. Metab.* 99, S12–S112

ELSEVIER

Erratum to: Heparin/heparan sulphate-based drugs

Neha S. Gandhi[1] and Ricardo L. Mancera[1,2,*]

[1] Curtin Health Innovation Research Institute, Western Australian Biomedical Research Institute, School of Biomedical Sciences, Curtin University, GPO Box U1987, Perth, WA 6845, Australia
[2] School of Pharmacy, Curtin University, GPO Box U1987, Perth, WA 6845, Australia

We have discovered a small mistake in the chemical structure of a drug molecule shown in Fig. 3 of our recently published paper. The correct structure for Fig. 3 is given below.

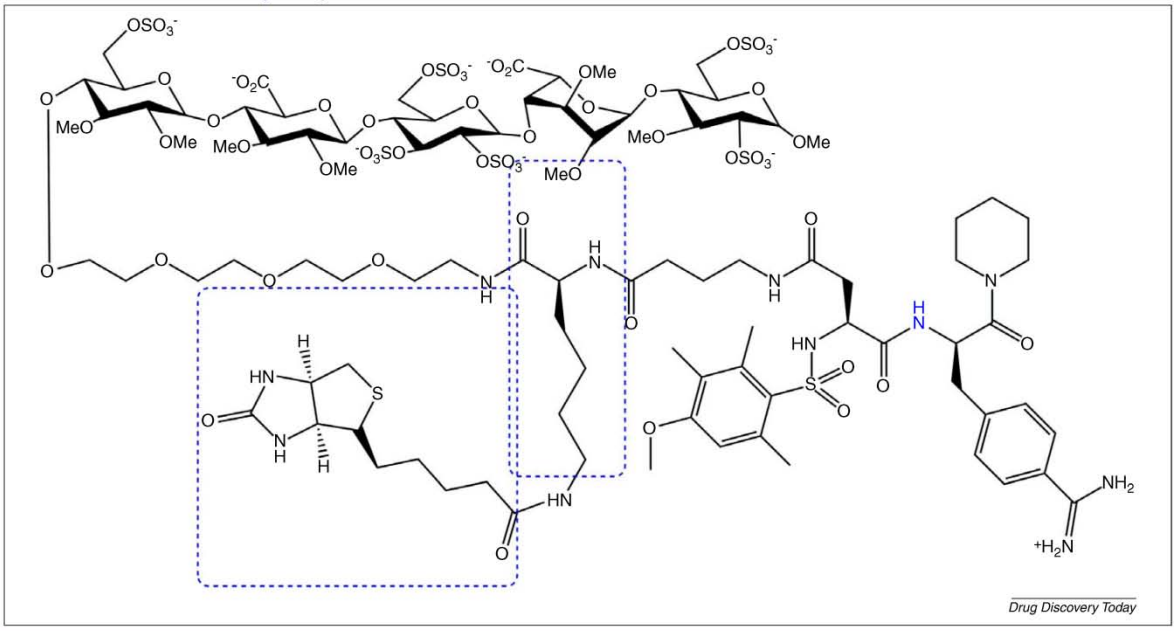


FIG. 3

Chemical structures of EP217609 and EP42675. EP42675 and EP217609 are the first representatives of a new class of synthetic, parenteral anticoagulants with a dual mechanism of action combining the properties of an indirect FXa inhibitor and a direct thrombin inhibitor. The structure of EP42675 can be inferred by deleting the biotin and lysine moieties (shown in dotted rectangles) from the structure of EP217609.

DOI of original article: 10.1016/j.drudis.2010.10.009
Corresponding author:. Mancera, R.L. (R.Mancera@curtin.edu.au)

1359-6446/06/$ - see front matter © 2011 Elsevier Ltd. All rights reserved. doi:10.1016/j.drudis.2011.02.009

4 Molecular Dynamics Simulations of CXCL-8 and its Interactions with a Receptor Peptide, Heparin Fragments, and Sulphated linked Cyclitols

JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

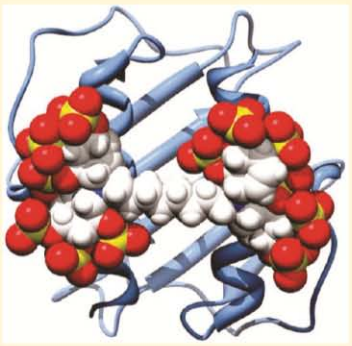
Molecular Dynamics Simulations of CXCL-8 and Its Interactions with a Receptor Peptide, Heparin Fragments, and Sulfated Linked Cyclitols

Neha S. Gandhi[†,‡] and Ricardo L. Mancera[*,†,‡,§]

[†]Curtin Health Innovation Research Institute, Western Australian Biomedical Research Institute, [‡]School of Biomedical Sciences, and [§]School of Pharmacy, Curtin University, GPO Box U1987, Perth WA 6845, Australia.

Supporting Information

ABSTRACT: CXCL-8 (Interleukin 8) is a CXC chemokine with a central role in the human immune response. We have undertaken extensive in silico analyses to elucidate the interactions of CXCL-8 with its various binding partners, which are crucial for its biological function. Sequence and structure analyses showed that residues in the thirdq β-sheet and basic residues in the heparin binding site are highly variable, while residues in the second β-sheet are highly conserved. Molecular dynamics simulations in aqueous solution of dimeric CXCL-8 have been performed with starting geometries from both X-ray and NMR structures showed shearing movements between the two antiparallel C-terminal helices. Dynamic conservation analyses of these simulations agreed with experimental data indicating that structural differences between the two structures at quaternary level arise from changes in the secondary structure of the N-terminal loop, the 3_{10}-helix, the 30s, 40s, and 50s loops and the third β-sheet, resulting in a different interhelical separation. Nevertheless, the observation of these different states indicates that CXCL-8 has the potential to undergo conformational changes, and it seems likely that this feature is relevant to the mode of binding of glycosaminoglycan (GAG) mimetics such as cyclitols. Simulations of the receptor peptide fragment−CXCL-8 complex identified several specific interactions of the receptor peptide with CXCL-8 that could be exploited in the structure-based design of competitive peptides and nonpeptidic molecules targeting CXCL-8 for combating inflammatory diseases. Simulations of the CXCL-8 dimer complexed with a 24-mer heparin fragment and of the CXCL-8−receptor peptide complex revealed that Arg60, Lys64, and Arg68 in the dimer bind to cyclitols in a horseshoe pattern, defining a region which is spatially distinct from the receptor binding site. There appears to be an optimum number of sulfates and an optimum length of alkyl spacers required for the interaction of cyclitol inhibitors with the dimeric form of CXCL-8. Calculation of the binding affinities of cyclitol inhibitors reflected satisfactorily the ranking of experimentally determined inhibitory potencies. The findings of these molecular modeling studies will help in the search for inhibitors which can modulate various CXCL-8 biological activities and serve as an excellent model system to study CXC-inhibitor interactions.



INTRODUCTION

CXCL-8 (Interleukin-8/IL-8) is a pro-inflammatory chemokine produced by various types of cells upon inflammatory stimuli.[1] It is secreted by endothelial and epithelial cells, fibroblasts, neutrophils, T lymphocytes, hepatocytes, keratinocytes, and peripheral blood monocytes.[2] CXCL-8 has a major role in neutrophil recruitment to and activation at the inflammation site where it enhances the adherence of circulating neutrophils to endothelial cells involving proteoglycans and subendothelial matrix proteins.[3] A series of cell-physiological responses required for neutrophil migration and its target function phagocytosis are also induced, like calcium mobilization, actin polymerization, degranulation, and respiratory burst.[2,4−6] CXCL-8 can then establish a concentration gradient to direct leukocytes to migrate across the endothelium and through the extracellular matrix into the tissue.[7] Because neutrophils are principal effector cells in acute ischemia-reperfusion injury, CXCL-8 has gained a major focus in transplantation research.[8,9] CXCL-8 has also been reported as a pro-inflammatory mediator in dermatologic diseases such as psoriasis.[10] In the context of inflammatory diseases of the lung, the concentration of CXCL-8 in bronchoalveolar lavage fluid from patients with idiopathic pulmonary fibrosis, bacterial pneumonia, and the acute respiratory distress syndrome is usually elevated and is associated with increased mortality.[11−14] CXCL-8 is thus an important mediator in the innate immune system response by acting as a potent chemo-attractant for neutrophils.[1] CXCL-8 is also known to promote angiogenesis and tumor metastasis.[15,16]

Interleukin-8 (IL-8) is the older name for CXCL-8, which received its name[17] based on the so-called CXC-chemokines (wherein the two N-terminal cysteines of chemokines are separated by one amino acid, represented in this name with an "X").[18] The CXCL-8 protein is encoded by the SCYB8 (small inducible cytokine subfamily B member 8) gene. CXCL-8 is generated as a precursor of 99 amino acids and is secreted after cleavage of a signal sequence of 20 residues. Several N-terminal processed forms are produced by proteolytic cleavage after secretion from peripheral blood monocytes, leukcocytes, and endothelial cells, but possibly other cells too.

Received: August 30, 2010
Published: February 07, 2011

ACS Publications © 2011 American Chemical Society

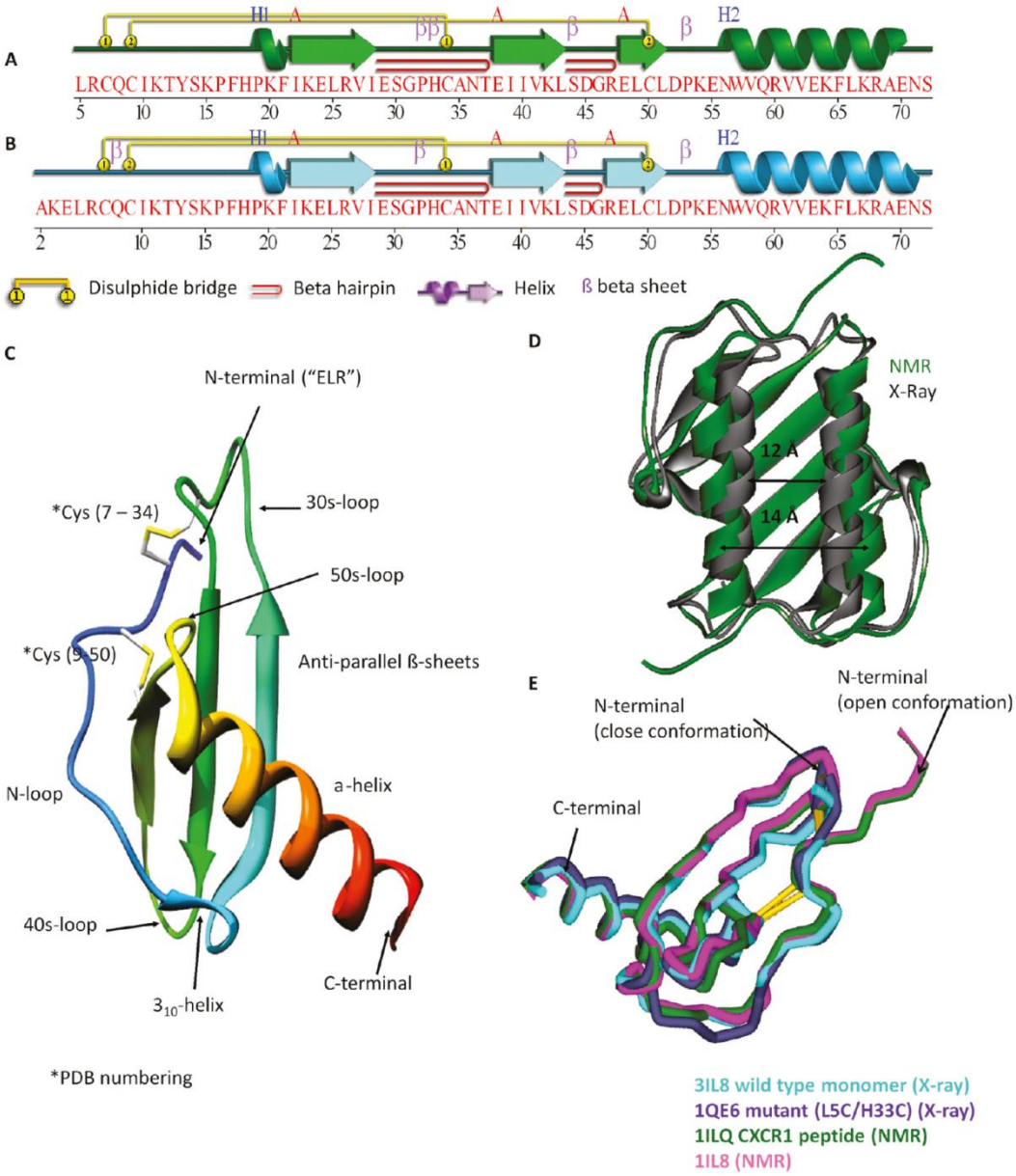


Figure 1. (A) Secondary structure representation of the structure of CXCL-8 (PDB code: 3IL8). (B) Secondary structure representation of the structure of CXCL-8 (PDB code: 1IL8). H1 indicates a 3_{10}-helix located before the first β-strand, and H2 indicates the C-terminal helix, while the symbol "β" represents the presence of β-sheets in both parts A and B. The disulfide connectivity in the N-loop is represented by yellow lines. (C) Ribbon diagram of the CXCL-8 structure. (D) C-terminal helices: closer in the X-ray structure compared to the NMR structure. (E) Superimposition of the backbone of the various NMR and X-ray structures of CXCL-8 revealing differences in the conformation of the N-loop.

CXCL-8 possess conserved amino acids that are important for its tertiary structure, such as the Greek Key motif shown in Figure 1, a single-turn helix (3_{10}-helix), three antiparallel β-strands, and a C-terminal α-helix. The Greek Key motif, which features four β-sheets, is characteristic of chemokines, but there are only three β-sheets in monomeric CXCL-8.[19] CXCL-8 has two conserved disulfide bridges between Cys7 and Cys34 and between Cys9 and Cys50 (PDB numbering). These four conserved cysteines stabilize the structure of CXCL-8.[20,21] Cys7 and Cys9 are situated close together near the N-terminus of the mature protein, while Cys34 resides at the center of the protein and Cys50 is located before the C-terminal helix. A loop of approximately 10 amino acids is located after Cys7 and Cys9, known as the N-loop. These helices and strands are connected by turns called 30s, 40s, and 50s loops. Cys61 and Cys77 are located, respectively, in the 30s and 50s loops.[19] The C-terminal α-helix plays a functional role in heparin binding and receptor interaction.[22−24]

The structural analysis of CXCL-8 by NMR spectroscopy[25−27] and X-ray diffraction[28,29] has revealed that CXCL-8 exists as a homodimer at high concentrations. Mutant and chimeric structures of CXCL-8 have also been reported.[30,31] The CXCL-8 dimer consists of two α-helices oriented in an antiparallel fashion on top of six β-sheets with the interactions at the dimerization interface consisting of main-chain hydrogen bonding (amino acids from position 23 to 29) between a β-sheet in each monomer. The structure of the monomer is similar to that of the monomeric subunit in the dimer, except that the C-terminal helical residues that are structured in the dimer are unstructured in the monomer.[27] There have been conflicting reports about the inter-relationship between CXCL-8 monomer−dimer equilibrium and their differential binding affinities for CXCR1, CXCR2, and glycosaminoglycans (GAGs), which appear to depend on the in vivo and in vitro conditions.[32−35] It has been reported that the dimerization constant for wild type CXCL-8 has a value of about 10−20 μM under physiological ionic strength, pH, and temperature conditions,[36,37] with this value being highly sensitive to specific solution conditions.[38] A dimerization constant of 120 nM has also been reported.[39] In addition to CXCL-8 homodimerization, CXCL-8 can also form heterodimers with other chemokines, such as platelet factor 4 (PF4) or CXCL-4.[40]

Few but important differences have been observed between the X-ray and NMR structures of CXCL-8. The secondary structure representations of the sequence of CXCL-8 in both the X-ray (PDB code: 3IL8) and NMR (PDB code: 1IL8) structures are shown in Figure 1. The principal difference between the NMR and X-ray structures of the CXCL-8 dimer is the distance between the α-helices (Figure 1D). The center-to-center distance between the antiparallel helices is about 12 Å in the X-ray structure and about 14 Å in the NMR structure.[28] Previous molecular dynamics (MD) simulations in implicit solvent have suggested that the helices in the NMR structure can move closer together, resulting in an interhelical distance that more closely resembles that observed in the X-ray structure.[41] Knowing which interhelical distance is more realistic is crucial for predicting and understanding the binding of ligands to CXCL-8.

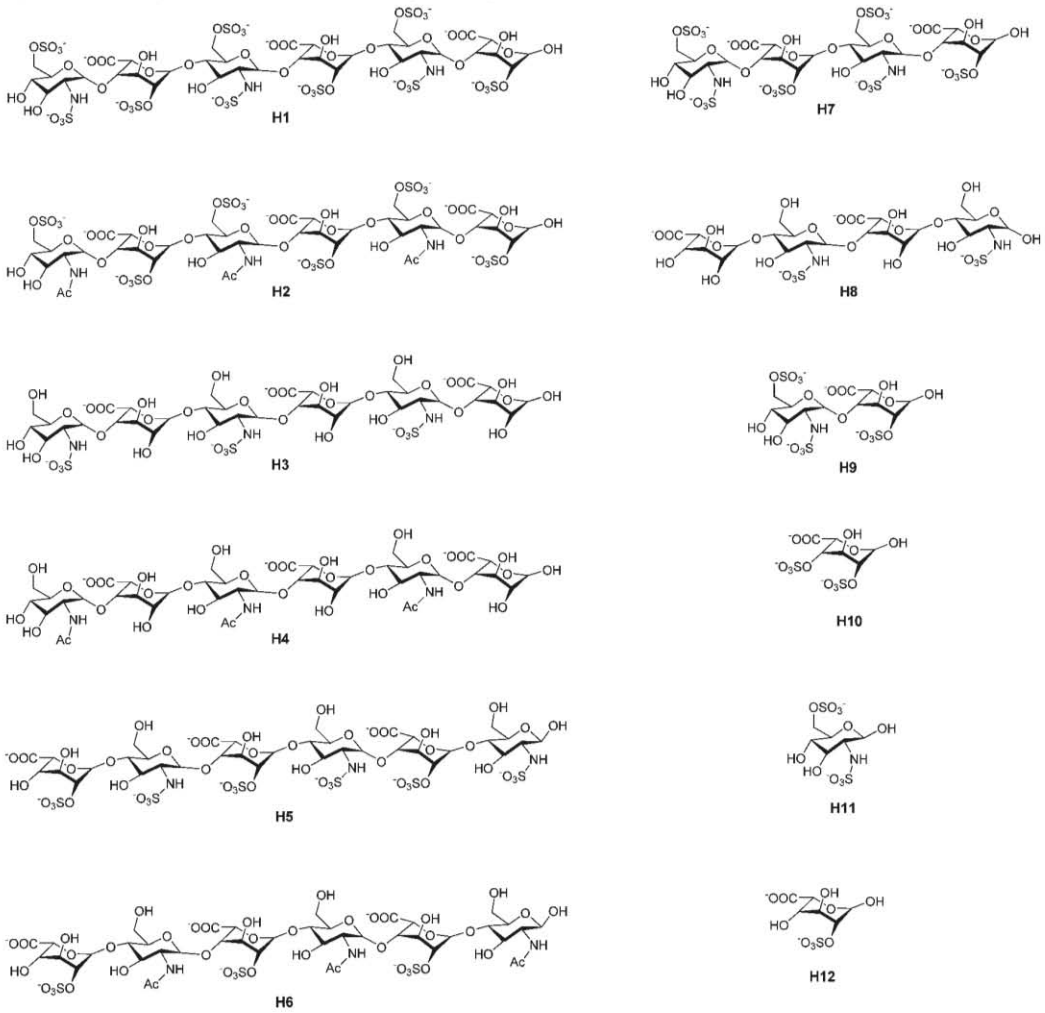
Another difference between the X-ray and NMR structures is the presence of different conformations of the N-loop and N-terminal up to the first cysteine.[26,28] Two distinct conformations (Figure 1E) can be defined: closed (with the N-loop in close proximity to the α-helix) and open (with the N-loop further away from the α-helix and near Arg47).[25,28,30,31] In the closed conformation, a hydrophobic pocket between the N-loop and Arg47 is exposed, whereas a proline residue occupies this region in the open conformation.[23]

Another important difference between the NMR and X-ray structures is the position and protonation of histidine residues. In the NMR (solution) structures, both His18 and His33 have low pK_a values [3.7 and 4.9, respectively].[26] In the X-ray structure, the Nε2 atom of His33 donates a hydrogen bond to the carbonyl in the backbone of Glu29, whereas in the low resolution NMR structure the side-chain of His33 is incorrectly placed and Arg6 has incorrect (nonplanar) χ-5 dihedral angle values. Hence, we focused our analysis on the more reliable X-ray structure of CXCL-8. In both structures, His18 makes the same hydrogen bond with the amide nitrogen of Lys20.

CXCL-8 carries out its regulatory role by binding and activating two homologous G-protein coupled receptors (GPCRs),[42−44] namely CXCR1 (IL-8 receptor A or IL-8 receptor type 1 or IL-8Rα) and CXCR2 (IL-8 receptor B or IL-8 receptor type 2 or IL-8Rβ) in humans, with affinities of 3.6 nM.[45] Mice express only the CXCR2 receptor on neutrophils. The CXCL-8 monomer and dimer differentially activate and regulate CXCR1 and CXCR2 receptors. CXCR1 binds with high affinity only to CXCL-8 while CXCR2 binds to multiple chemokines, such as the epithelial cell-derived neutrophil activator (ENA-78), neutrophil activating peptide-2, and the GRO subfamily of proteins (GRO-α, GRO-β, and GRO-γ).[46] CXCL-8 also binds to the promiscuous Duffy antigen/receptor for chemokines (DARC) found on the surface of erythrocytes, with an affinity of approximately 20 nM.[47] Mutagenesis experiments have identified the specific interactions between the N-loop of CXCL-8, including the triad Glu-Leu-Arg ("ELR"), and the N-terminal domains (site I or primary site),[48] and the III and IV extracellular loops (site II or secondary site)[23] of the chemokine receptors, which determine the receptor trafficking profiles. Selectivity and signaling specificity is achieved by the N-terminal domains of the receptors and their ligands.[49−51] The NMR structure of wild type CXCL-8 dimer (and not the functionally relevant monomer) complexed with a peptide analogue of CXCR1 (a receptor peptide) has been published.[25] This receptor fragment shows relatively potent inhibition of CXCL-8 receptor binding (K_i = 7 μM).[52] The NMR structure revealed the interaction to be primarily hydrophobic and mediated by side-chains involving amino acids Gln8, Ile10, Lys11, Tyr13, Phe17, Phe21, Ile40, Leu43, Arg47, and Leu49 present in the N-terminal region of CXCL-8,[25] whereas the side-chains of Asp5 and the amine capping residue NH_2 of the receptor peptide make hydrogen bonds with Lys20 and Gln8 of CXCL-8, respectively. Mutational studies of CXCR1 have identified Arg199 and Arg203 on the third extracellular loop and Asp265 on the fourth extracellular loop as being important for CXCL-8 binding,[53] whereas the receptor binding site on the surface of CXCL-8 comprises residues on the third β-sheet (Glu48 to Cys50), the turn preceding the third β-sheet (Ser44), one residue on the C-terminal α-helix (Val61) and residues starting from Lys15 in the N-terminal region.[54] The interactions of CXCR1 domain with the CXCL-8 monomer were distinctly different from the dimer.[54,55] On the basis the binding of a monomeric CXCL-8 to CXCR1 N-domain, only N-loop residues of CXCL-8 were found to be involved in direct binding whereas residues in any other region (such as Val61 and Ser44) is due to indirect coupling interactions in the dimer. Molecular modeling studies of the interaction of CXCL-8 with CXCR2[56] have suggested that the N-terminus of CXCR2 binds to the N-terminus of CXCL-8 (site I or primary site), while extracellular loops III and IV of CXCR2 bind to the C-terminal helix of CXCL-8 (site II or secondary site).

CXCL-8 is known to bind heparin/heparan sulfate (HS) present on the endothelial cell surface and basement membrane.[22] Heparin/HS play an important role in the promotion of CXCL-8 dependent transmigration of neutrophils and the protection of the tissue microenvironment from cleavage by enzymes, such as elastase released from migrating cells. The addition of HS or heparin to CXCL-8 has also been shown to prevent CXCL-8 from unfolding, thereby suggesting a role for glycosaminoglycans (GAGs) on CXCL-8 stability and protection from proteolytic degradation in vivo.[39] Amino acids involved in heparin binding include Arg60, Lys64, Lys67, and Arg68, which are located within the C-terminal α-helix, as well as His18 and Lys20, which are located in the proximal loop.[57] Dimeric CXCL-8 requires 5−6 trisulfated saccharides (sulfated domain) to bind to each monomer,[58] connected by a flexible linker consisting of 12−14 sugar residues (*N*-acetylated) [Figure S1 of the Supporting Information]. This heparin fragment is believed to adopt a horseshoe arrangement wherein each sulfated domain interacts with the antiparallel CXCL-8 dimer.[59] It has been reported that homodimerization of CXCL-8 is required for heparin binding, whereas monomeric CXCL-8 is capable of binding to its

Table 1. Synthetic Library of Amine-Terminated Sugar−Heparin Derivatives



receptors without loss of affinity.[60−63] Earlier molecular modeling studies on the binding of heparin to monomeric and dimeric CXCL-8 correlated well with experimental data, identifying the putative amino acids and the length of GAG fragments required for interaction with CXCL-8.[64−66] The ranking of these residues in terms of their contributions to GAG binding was determined to be Lys64, Arg60, Lys20, Lys67, Arg68, and His18.[64]

The GAG-binding region is essentially distinct from the receptor-binding region, suggesting that CXCL-8 may bind to its receptor in the presence of GAGs.[57] Conflicting reports have emerged regarding the length of heparin fragments required for binding to CXCL-8.[39,58,64] The dissociation constant (K_d) of heparin fragments to CXCL-8 has been investigated using isothermal titration calorimetry and determined to be 0.39−2.63 μM, with five saccharide units required to bind to each monomer.[58] In the same study, the binding affinity of heparin to CXCL-8 was shown to increase with increasing chain length. In addition, O-sulfation in heparin fragments was found to be more important than N-sulfation for binding to CXCL-8. Nonetheless, another study found the existence of a periodic pattern for the dissociation constants of heparin oligosaccharides with respect to chain length.[39] GAG disaccharides were identified to be the minimum length required for binding to CXCL-8. Isothermal fluorescence titration was used to estimate the binding affinities of heparin/HS oligosaccharides of various lengths, and these were found to be dependent on the oligomerization state of CXCL-8. High affinity binding of heparin/HS was reported for monomeric CXCL-8 while low affinity was reported for the dimeric form.

Soluble heparin/HS are known to inhibit binding of CXCL-8 to its receptors, blocking its biological activity. By contrast, cell surface heparin/HS help present chemokines to their GPCRs by increasing the local concentration of protein.[58] Heparin-based molecules from a synthetic library of amine-terminated sugars have also been reported to block the formation of chemokine gradients on the cell surface.[67] We henceforth refer to these oligosaccharides as H1 to H12 [Table 1]. Experiments showed that CXCL-8 binds oligosaccharides H1, H2, H5, H6, H7, H9, and H10. In addition, the binding of hexasaccharide H6 suggested that the 2-O-sulfate groups in alpha-L-iduronate (IdoA)

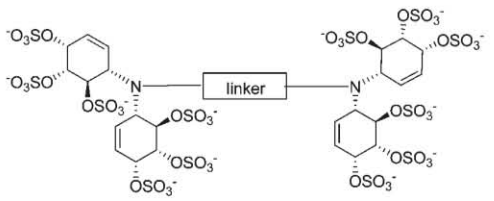
units play an important role in this interaction. The binding affinity of synthetic oligosaccharides as short as disaccharide H9 was attributed to multivalent and cooperative interactions present on the chemokine surface, as opposed to the horseshoe arrangement proposed earlier.[59]

Several approaches have been pursued in the last two decades in order to identify molecules which can modulate the biological activities of CXCL-8.[9,68] Peptide-mimetics based on the CXCL-8 scaffold have been reported to block the actions of CXCL-8. In view of the critical role of the N-terminal domain, antagonists of CXCL-8 have been developed either by truncating the amino-terminal of CXCL-8 or by modifying the amino acids in the ELR+ motif at the N-terminus. These modified CXCL-8 derivatives bind both CXCR1 and CXCR2 and inhibit CXCL-8 mediated neutrophil responses.[69−71] Alternatively, mimicking the receptor rather than the ligand has also been reported.[72] The identification and role of small molecules capable of selectively blocking CXCL-8 receptors has also been reviewed.[9,68,73] Compared to low molecular weight chemokine receptor antagonists, which generally target leukocytes, another novel approach is to design molecules to interfere with wild-type chemokine signaling at the endothelium by masking endothelial GAGs with non-GPCR-binding CXCL-8. Recently, a dnCXCL-8 (dominant negative CXCL-8) functional antagonist was shown to exhibit significantly increased binding affinity for GAG structures on endothelial surfaces, which can significantly compete with wild-type CXCL-8 from endothelial GAGs.[74,75] In addition, the natural GPCR binding motif to CXCR1/2 was knocked out in dnCXCL8, resulting in a decoy protein.

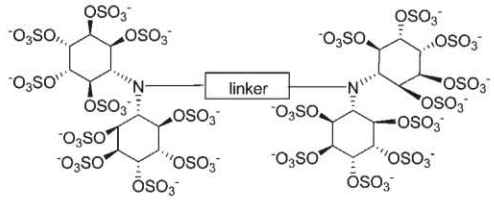
Carbohydrate mimetics that bind to CXCL-8 and inhibit its biological activity have been reported.[76] These molecules are alkyl-linked tetrameric cyclitols [Table 2] that require at least a 7-carbon long spacer in order to bind to CXCL-8. Cyclitols are pseudosugars in which the pyranosyl-ring oxygen has been replaced by a methylene unit. Such compounds are highly stable and do not suffer from glycosidic cleavage. Four of the 15 reported sulfated cyclitols, i.e. numbers 8, 9, 10, and 10a, blocked effectively the CXCL-8-heparin interaction, with the remaining 11 compounds showing no activity. Cyclitol 8 displayed an IC_{50} of 50 nM, and cyclitols 9 and 10 were also quite strong inhibitors, although not as effective as compound 8. The highly sulfated cyclitol 10a showed comparable inhibition to 10.

In this paper, we report a series of structural bioinformatics and molecular modeling studies aimed at characterizing the interactions of CXCL-8 with its various binding partners. We have analyzed the evolutionary conservation of amino acid positions in human CXCL-8 by comparing 39 protein sequences to identify structurally and functionally important regions for interactions with its receptors and GAGs. We have also carried out molecular dynamics (MD) simulations in aqueous solution of the dimer of CXCL-8 with starting geometries from both its X-ray and NMR structures and monitor the changes in the secondary structure of the CXCL-8 dimer. These analyses were further extended to identify the region which contributes mostly to the conformational flexibility of the CXCL-8 dimer. MD simulations based on structural data of a complex of CXCL-8 to a receptor peptide fragment identified extended interaction sites within the protein−protein interface upon conformational changes. We have also performed molecular docking simulations of various heparin oligosaccharides to the CXCL-8 monomer and dimer. Since no structural determinations have been reported for bound complexes of cyclitols with CXCL-8, we have also modeled their interactions with CXCL-8 oligomers.

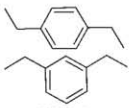
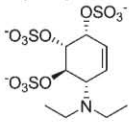
Table 2. Structures of the Various Sulfated Linked Cyclitols



1



1a

Molecule	Linker
2-10	$-(CH_2)_n-$ where n = 2-10
6a	$-(CH_2)_6-$
10a	$-(CH_2)_{10}-$
11	
12	
13	
14	

MATERIALS AND METHODS

Sequence and Structural Analyses. A PSI-BLAST[77] search was performed against the SWISS-PROT database.[78] The query sequence corresponds to the CXCL-8 human sequence (P10145). The resulting sequences, species, SWISS-PROT[78] accession numbers, and sequence identity with reference to human CXCL-8 are listed in Supporting Information Table S1.

A ConSurf[79] analysis of the evolutionary conservation profile of CXCL-8 (PDB code 3IL8) was carried out to identify putative functional regions. The phylogenetic tree produced by Consurf was plotted using FigTree 1.2.2.[80] More details about this method are provided in the Supporting Information.

Normal Modes Analysis (NMA). NMA was used to predict the low frequency, high amplitude interdomain equilibrium motions of the CXCL-8 dimer using the Elastic Network Model (ENM), as available through the ElNémo web server.[81] PDB

structures 3IL8 and 1IL8 were used as input for these calculations. The lowest frequency normal modes were computed with the minimum and maximum perturbations set to −150 and 150 DQ, respectively, to amplify molecular motions.

Electrostatic Potential Surface. Electrostatic potential calculations were done using the DELPHI program implemented in DS Modeling 2.1 (Accelrys, Inc.) using the atomic partial charges assigned by CFF,[82] with a protein interior dielectric constant of 1, a solvent dielectric constant of 80, and an ionic strength of 0.145 M.

Sulfate Binding Motif Search. The presence of sulfate ions (originating from the crystallization buffer) on the surface of a protein structure determined by X-ray diffraction may suggest the presence of GAG binding sites, as is the case with mutant CXCL-8,[30] SDF-1α,[83] MCP-1,[84] and RANTES.[85] Sequence analysis combined with fold recognition methods can also help to predict sulfate binding motifs in GAG-binding proteins.[86] A search for structural homologues of CXCL-8 was performed using the SSM server[87,88] to look for structures with similar folds. PDB structure 1QE6[30] was used as a query. Pairwise comparison and 3D alignment of protein structures were carried out considering a structural homology of 25% or more. The results were manually screened for structures which had been crystallized with sulfate ions or sulfated ligands.

Molecular Docking. The chemical structures of heparin H1−H12, 24-mer fragment, and cyclitols were sketched in Chemdraw Ultra 11.0[89] and saved in MDL molfile format. The program OMEGA v.2.3.2[90] was used to convert all heparin and cyclitol structures to 3D multiconformer structures and to add hydrogen atoms and partial charges using the MMFF force-field.[91−95] The maximum number of conformers (maxconfs) was set to 50. OMEGA was used because of its ability to enumerate ring conformations and invertible nitrogen atoms. This method is useful in enumerating ring conformations in heparin, particularly sulfated alpha-L-iduronate (IdoA2S) residues which may adopt skew boat ($^2S_0$) and chair ($^1C_4$) conformations.[96]

The FRED[97] molecular docking engine takes a multiconformer database of one or more ligands generated by OMEGA, a target protein structure, a box defining the active site of the protein on the basis of a ligand−protein complex and several optional parameters as input. The binding pocket was defined using the ligand-free CXCL-8 X-ray structure (PDB code: 3IL8) and the NMR structure (PDB code: 1IL8) using FRED-Receptor (Eyes Open, Inc.). The protein was prepared using the MMFF force-field in FRED. Three different binding sites were defined: site 1, parallel to the alpha helices (monomer); site 2, perpendicular to the alpha helices (dimer); and site 3, blind docking wherein the whole CXCL-8 dimer was enclosed in the box. FRED[97] was used because of its ability to dock large flexible ligands in large binding sites.[98] The inner contour (a shape complementary to the active site, used during the exhaustive search) was disabled as the binding sites in GAG-binding proteins consist of clefts or sets of juxtaposed surface residues rather than pockets.[99] The ligand conformers were kept flexible whereas protein structure was treated as rigid during the docking simulations. The number of poses (num_poses) to be returned by the exhaustive search was set to 300, defined as the top scoring poses selected from the list of all poses and scored by the scoring functions specified by the exhaustive scoring. The number of alternative poses (num_alt_poses) was set to 100, defined as additional poses to the top consensus structure poses.

In the optimization step, four scoring functions were used: ChemGauss,[97] ChemScore,[100] PLP,[101] and ScreenScore.[102] Binding poses obtained from FRED using the consensus score were extracted. The Gaussian shape fitting function[103] was not used. AutoDock has been used successfully to dock and score heparin−protein complexes,[66,86] and hence, the consensus binding poses were redocked with AutoDock 4.0[104] using AutoDock's Lamarckian genetic algorithm with a population of 200 individuals and 50 × 10^6 energy evaluations. The grid box was defined with a constant grid spacing of 0.37 Å around each ligand molecule using the previous obtained consensus binding pose. The AutoDock 4.0 scoring function[105] was used to predict inhibition constants (K_i).

There is a provision in AutoDock 4.0 to accommodate flexibility in the protein; however, this feature was not used in our study as there are too many flexible basic residues in the binding site that interact with the flexible ligands, such as GAGs. Furthermore, AutoDock tends to take a large amount of time to sample enough conformational space for molecules such as heparin and cyclitols with a large number of degrees of freedom (translational, rotational, and torsional). This can lead to the failure of the scoring function to accurately predict binding affinities. Hence, docking by FRED followed by a Lamarckian genetic algorithm (LGA) search in Autodock were used to determine the key interactions between GAG molecules and CXCL-8.

Molecular Dynamics Simulations. MD simulations were carried out for the CXCL-8 dimer (PDB code: 3IL8, 1IL8) as well as for complexes with the receptor peptide (PDB code: 1ILQ) and the 24-mer heparin fragment. The monomeric CXCL-8 is the relevant form for CXCR1 binding and function, and the monomer coordinates of the dimer structure (PDB code: 1ILQ) were only used for the computational studies of CXCL-8 with the receptor peptide fragment.

Simulations starting with the CXCL-8 NMR and X-ray structures were run at 300, 325, and 337 K for 100 ns each, including an equilibration period of 10 ns. Simulation of the NMR structure was also carried out at 313 K for 100 ns for direct comparison with the NMR determination, which was carried out at 40 °C.[26] Simulations were run at different temperatures to investigate if the conformational differences between the NMR and X-ray structures arise due to the influence of thermal motion at higher temperatures. The CXCL-8-receptor peptide and CXCL-8−24mer heparin fragment complexes were also simulated for 50 ns at 308 K. No restraints or constraints were used during the equilibration and production runs for any system, except for the CXCL-8-cyclitol complexes, as described below.

MD simulations were also performed for complexes of CXCL-8 with cyclitols 7, 8, 9, 10, and 10a. During heating and equilibration, distance restraints (with a force constant of 2.0 kcal/(mol Å^2)) were applied to all the atoms in the protein that interact with the each ligand to prevent it from leaving the binding site. The CXCL-8-cyclitol complexes were deemed to have equilibrated after 1.5 ns with these distance restraints. The production phases of the simulations (in the absence of any restraints) were then run at 300 K for a further 8.0 ns for the CXCL-8-cyclitol complexes. The trajectories obtained from these simulations were used in the calculation of free energies of binding (see below), for which long simulations are not required.[106]

The Parm99SB[107] force field in AMBER 9.0[108] was used for the proteins in all simulations, with the GAFF[109] parameter set used for cyclitols. The GLYCAM06[110] parameter set was used for the heparin fragment. Partial atomic charges for cyclitols and the individual units of heparin were obtained using the restricted electrostatic potential (RESP) method[111] using R.E.D.-III tools.[112] Prior to charge derivation, all ligands were subjected to a full quantum mechanics geometry optimization with a HF/6-31G* basis set using Gaussian 03.[113]

All energy minimizations and MD simulations were performed with AMBER 9.0.[108] The N- and C-termini of CXCL-8 and the receptor peptide were capped with residues ACE (acetyl beginning group) and NME (N-methylamine ending group), respectively, to remove charges in the end groups. Capping the C-terminus of CXCL-8 was done to further ensure an energetically favorable helix termination. A cubic box of TIP3P water molecules[114] was added to solvate the protein and its ligand-bound complexes, keeping a minimum distance of 12.0 Å between each face of the box and the protein. The numbers of water molecules added to the X-ray (PDB code: 3IL8) and NMR (PDB code: 1IL8) structures of CXCL-8 were 9367 and 11899, respectively. In the case of the complex of the CXCL-8 dimer with the receptor peptide (PDB code: 1ILQ), 38518 water molecules were added. The number of water molecules added to the complexes of CXCL-8 with cyclitols 7, 8, 9, 10, and 10a was 9295. The number of water molecules added to the complex of CXCL-8 with a 24-mer heparin fragment was 10758. Net charges in the protein and/or heparin fragments were neutralized by adding an appropriate number of counterions (Na^+ or Cl^-).

Free Energy Calculations. Free energy calculations using the MM/PBSA method have been reported for CXCL-8 homo and heterodimers using the CHARMm force-field.[115] We calculated the free energies of binding of cyclitol-CXCL-8 and heparin 24mer-CXCL-8 complexes using the MM/PB(GB)SA method. The MM/PBSA module of AMBER 9.0 was used to compute the above-mentioned components of the free energy. More details of the MM/PB(GB)SA methods can be found in the Supporting Information. For the cyclitol-CXCL-8 complexes, 800 snapshots of the coordinates of the system were taken at 10 ps intervals from the last 8.0 ns of the production runs. For the complexes of CXCL-8 with the 24-mer heparin fragment and the receptor peptide, 1500 snapshots of the coordinates of the systems were taken from the last 30 ns of the simulations. All solvent molecules and counterions were removed prior to analysis. The snapshots were also analyzed with the modified generalized Born (GB) solvation model,[116] modified for use with the PARM94 parameters to obtain energies of solvation. Poisson−Boltzmann calculations were also used to obtain solvation energies, with an ionic strength of 0.14 and dielectric constants (ε) of 1 for the solute and 80 for the solvent. A probe solvent radius of 1.4 Å and the PARSE atomic radii parameter set[117] were used to determine the molecular surface. Different surface parameters were used: in the case of GB calculations, $\gamma = 0.0072$ kcal/Å^2 and $\beta = 0.0$ kcal/mol, and in the case of PB calculations, $\gamma = 0.00542$ kcal/Å^2 and $\beta = 0.92$ kcal/mol.[117,118]

The vibrational, rotational, and translational entropies of the systems were computed by performing normal modes analyses[119] using the Nmode module of AMBER on snapshots collected every 200 ps, resulting in 40 snapshots for the 8.0 ns simulations for various cyclitols complexed with the CXCL-8 dimer. For the complexes of CXCL-8 with the 24-mer heparin fragment and the receptor peptide, snapshots were collected every 500 ps, resulting in 60 snapshots for the last 30 ns of the simulations. Prior to these normal modes analyses, the selected snapshots of the complex, protein, and ligand were subjected to a full conjugate gradients energy minimization using a distance-dependent dielectric ($\varepsilon = 4r$) and a convergence criterion of 0.0001 kcal/mol. The reported vibrational, rotational, and translational entropies were calculated as averages over all selected snapshots.

Visualization. All the 3D protein and ligand structures were edited and visualized in Accelrys Discovery Studio 2.1 (Accelrys, Inc.). Consurf results and MD simulation trajectories were visualized using UCSF Chimera.[120] The molecular solvent accessible surface area (SASA) and the buried/contacting surface for individual residues were calculated using NOC 3.0[121] with a solvent (water) probe radius of 1.4 Å.

RESULTS AND DISCUSSION

Sequence and Structure Analysis. Several studies have been reported for the CXCL-8 family based on a combined sequence, structure, and phylogenetic analysis to understand its structure−function relationship, interpret the evolutionary relationship across family members, understand the receptor specificity exhibited by ligands, locate the functionally important motifs responsible for dimerization, and characterize the relationship between GAG and receptor binding sites.[65,122−124]

Although multiple sequence alignments (MSA) are a powerful source of information for predicting functional regions in biological sequences, it is necessary to correlate mutation and evolutionary conservation data with the corresponding location in a three-dimensional structure, since residues that are distant in sequence may be found in close proximity in a folded protein. Furthermore, binding surfaces involved in biochemical functions experience different selection pressures from other regions on the surface of proteins.[125] The Consurf server[79] and ConSurf-DB[126] use the 3D-structure of a protein in order to represent such functional regions. In this study the Consurf analysis was carried out using chain A in PDB structure 3IL8 (the PDB numbering of this structure was henceforth used). The MSA that was constructed using the default parameters of the ConSurf DB server contained 50 homologues, which include CXCL-4, 5, 6, 13, 15, and 19 family ligands as well as PF4 (platelet factor 4). The fish analogues of CXCL-8 are not included in the hit list after searching ConSurf DB with default PSI-BLAST parameters. Consequently, a query with the CXCL-8 structure using default parameters in ConSurf DB results in the incorrect assignment of conservation scores for the majority of residues (Supporting Information Figures S2 and S3).

In order to study the phylogenetic relationships between the distinct members of the CXCL-8 family, the MSA of the 39 CXCL-8 sequences listed in Supporting Information Table S1 was constructed using MUSCLE [62]. The ELR motif, residues in the second β-sheet and the cysteine residues involved in disulfide bridges, appear to be highly conserved in both default Consurf and manually curated analyses (Supporting Information Figures S2 and S3 and Figure 2). The ELR motif has been substituted with other amino acids in *Chimaera phantasma, Paralichthys olivaceus, Latris lineate*, and *Latris lineate* (Supporting Information Figure S4). Residues His18 and Lys20 showed conservation scores of 4 and 6, respectively, upon inclusion of CXCL-8 sequences from fish. There was an additional serine found in the second β-sheet of *Dasypus novemcinctus* (Supporting Information Figure S4). Residues in the first β-sheet which form the dimer interface appear to be only partially conserved based on the ConSurf analysis of the CXCL-8 monomer structure. Analyses using both methods suggest that residues in the third β-sheet and the GAG binding site in the C-terminus are highly variable. Residues Arg60 and Lys64 both had a conservation score of 1, while Lys67 and Arg68 had conservation scores of 3 and 4, respectively. Consequently, it appears that human CXCL-8 may have unique features that distinguish it from chemokines expressed in other species with respect to interactions with molecules at the cell-surface, such as sulfated GAGs.

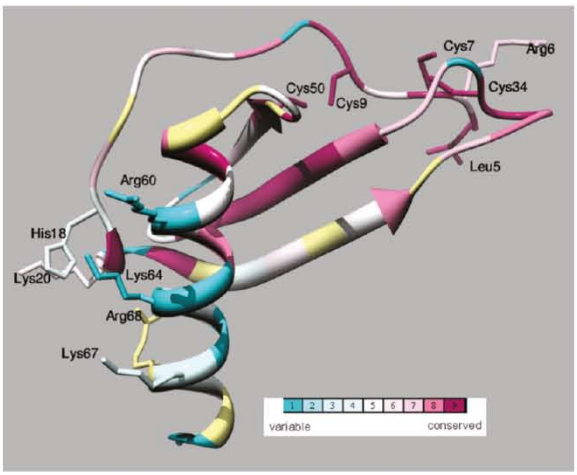


Figure 2. Representation of the structure of CXCL-8 (PBD code: 3IL8) upon multiple sequence alignment with other known CXCL-8 proteins. The structure is colored according to the Consurf color schemes, from turquoise to burgundy, representing evolutionary conservation scores from variable through conserved, respectively.

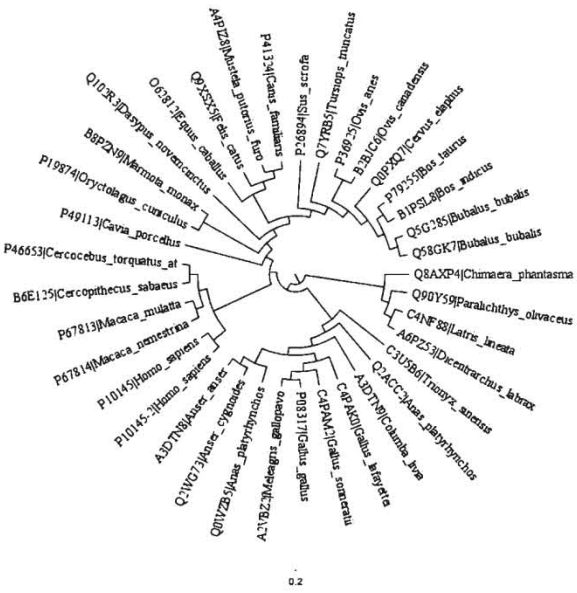


Figure 3. Polar phylogenetic trees of CXCL-8 sequences computed from a multiple sequence alignment using MUSCLE. The phylogenetic tree produced by Consurf was plotted using FigTree 1.2.2.[80]

A molecular phylogenetic tree was constructed to analyze the evolutionary relationships of CXCL-8 proteins (Figure 3). Sequence homology reveals that human CXCL-8 is very closely related to macaque, and hence, they are clustered on the same clade (Figure 3). The CXCL-8 encoding region of fish species *Chimaera phantasma*,[127] *Paralichthys olivaceus*,[128] *Latris lineate*,[129] and *Dicentrarchus labrax*[130] is clustered into one lineage. The CXCL-8 sequences of the fish species studied lack the ELR motif which, in humans, induces a chemotactic effect on neutrophils and an angiogenic effect,[15] suggesting that CXCL-8 may not possess these functions in fish. These results also suggest that avian CXCL-8 proteins evolved distinctly from those of mammals.[131] In this study, we undertook an evolutionary survey of sequence substitution patterns of CXCL-8 but not its receptors. The evolutionary analyses of chemokines and their receptors have already been reported,[132−135] suggesting that as the N-terminus and transmembrane portions of chemokine receptors are involved in ligand/receptor binding and recognition, amino acid substitutions in the N-terminus of the receptor as well as in the ligands in various species could potentially modify the ligand−receptor interaction, i.e., the specificity and/or affinity as compared to human CXCL-8/CXCR1 interactions. For example, the mouse CXCL-8 receptor does not bind human CXCL-8 with high affinity.[136] Such divergence may have resulted in important differences in the function of chemokines between species.

The electrostatic potential of proteins arising from charged sidechains plays a role in protein folding and stability and specific protein−protein/ligand recognition. In general, electrostatic surface representation aids in the prediction of protein−ligand interactions based on the concept of charge complementarity. Supporting Information Figure S5 shows the electrostatic surface representation of CXCL-8 using its X-ray structure (PDB code: 3IL8). Positive electrostatic potential surfaces are shown in blue, and negative electrostatic potential surfaces are shown in red. A positively charged area of the protein surface can be seen to be created by the C-terminal α-helix together with the loop connecting the extended N-terminal strand region with the first β-strand of the sheet, wherein the basic residues alongside the alpha helices can interact with negatively charged ligands like heparin. In this study, CXCL-8 is taken to be at acidic pH, wherein all histidines are considered protonated.[137]

Sulfate Binding Motif Search. Various approaches have been used to identify heparin/GAG-binding sites on the surface of proteins. These have been based on amino acid composition,[138,139] consensus sequences,[140] secondary structure,[141] spatial distribution of basic amino acids,[142] and the electrostatic surface properties of proteins.[143] GAG-binding sites generally consist of a cluster of basic residues on the protein surface but not necessarily in a continuous sequence. A survey of sulfate-binding regions from the PDB database can reveal several positively charged regions on a protein surface which might form the binding site of sulfated GAGs.[86] In this work, the PDB database was surveyed for sulfate-binding motifs using SSM searches to look for proteins with a similar fold as CXCL-8 and that have been crystallized with sulfate ions or sulfated ligands. The mutant structure of the CXCL-8 dimer (PDB code: 1QE6)[30] was crystallized with two sulfates: one located near residues Ser1 and Arg6 in the N-terminus and another one located near residues His18, Pro19, and Lys20. Table 3 lists details of different structural homologues of CXCL-8 along with their sulfate binding regions. It appears that most of the sulfates and sulfated ligands could be mapped to residues in the vicinity of His18, Lys20, Lys64, and Arg68 of the C-terminal helix. One of the examples is shown in Figure 4A. These residues are indeed already known to play an important role in heparin binding.[57] In addition to the heparin binding regions present on CXCL-8, we found sulfate-binding motifs in the vicinity of Arg47, which suggests that this residue may further extend the length of the putative heparin binding site.

Some of the sulfated ligands complexed with structural homologues of CXCL-8 were found to be associated with residues on the β-sheet surface of CXCL-8. One of these structural homologues is CXCL-12 (PDB code: 2NWG), bound with two heparin disaccharides. One of the disaccharides would

Table 3. Prediction of Sulfate Binding Motifs

PDB code	protein	amino acids of target structures involved in sulfate binding with reference to the CXCL-8 structure (indicated in bold)			rmsd (Å) calculated between the Cα-atoms of matched residues in the best 3D superposition of the target structures and the reference CXCL-8 structure (1QE6)	% sequence identity between the target and CXCL-8	% structural structure element match of the target protein with respect to the CXCL-8 structure (1QE6)
1F9P	connective tissue activating peptide- III (CTAP-III)	His35	Pro36	Lys 81	1.28	35	64
		His18	**Pro19**	**Lys64**			
2NWG[175]	stromal cell-derived factor-1alpha/ CXCL-12 (SDF-1α)	Arg 20	Ala19	Ala21	2.65	29	100
		His18	**Phe17**	**Lys20**			
		Asn30	Lys64				
		Glu29	**Arg68**				
		His25	Lys27	Asn45			
		Glu24	**Arg26**				
		Arg41					
		Lys42					
1QG7[176], 1A15[83]		Ala19	Arg20	Ala21	2.70	22	100
		His18	**Pro19**	**Lys64**			
2HCI[177]	human macrophage inflammatory protein-3alpha (MIP-3α)	Leu15	Trp55		1.71	27	80
		Phe17	**Trp57**				
		Thr54	Trp55	Gln53			
		Gln56	**Trp57**	**Glu55**			
		Cys32	Gln26	Leu27			
		Cys34	**Ile28**	**Glu29**			
		Pro51	Leu27	Ile37			
		Val67	Asn35				
		Pro53	**Glu29**	**Ile39**			
		Ala69	**Thr37**				
		Gly31	Asp33	Cys32			
		Ala28					
		His33	**Ala35**	**Cys34**			
		Ser30					
2RA4[178]	human monocyte chemoattractant protein 4 (MCP-4/CCL-13)	Arg24	Lys18	Lys48	1.6	29	80
		Phe21	**Lys15**	**Arg47**			
		Lys18	Ser21	Arg24			
		His18	**Pro19**	**Phe21**			
		Lys48	Thr44	Arg24			
		Arg47	**Leu43**	**Phe21**			
1B3A[179]	anti-HIV protein aop-rantes	His23	Arg47	Thr43	1.76	23	80
		Lys45	Arg44				
1EQT[85]	Met-rantes	**Phe21**	**Arg47**	**Leu43**			
		Asp45	**Ser44**				
1U4R[180]	human rantes mutant 44-aana-47	Ser31	Gly32	Lys33	1.59	21	80
		Glu29	**Ser30**	**Cys33**	1.96	19	80
		Phe28	Pro9				
		Arg26	**Cys7**				

be superimposed near His18 and Lys20 of CXCL-8, whereas the other heparin disaccharide binding site consists of a BBXB motif (in this case Lys-His-Leu-Lys) in the 40s loop near the dimerization interface connecting the two β-sheets. There are several reasons why it is unlikely that a heparin saccharide can bind to CXCL-8 on the β-sheet surface. First, the histidine in the BBXB motif of CXCL-12 has been substituted by Glu in CXCL-8 (Figure 4B). Second, superimposition of the backbone of CXCL-12 onto that of CXCL-8 using a SSM search results in a rmsd greater than 2.5 Å, with a larger deviation in the sulfate binding motif region. Third, the CXCL-8 dimer displays the same positively charged area on the surface of the α-helices, except in the SDF-1α dimer, where the highest positively charged area is on the back of the β-sheet formed by the dimer interface, as reported before using a cluster analysis of chemokine isosurfaces.[65] CXCL ligands such as macrophage inflammatory protein-3alpha (MIP-3α) and RANTES bind sulfates and sulfated ligands in the region formed by a long loop connecting the third β-strand and the C-terminal α-helix. However, due to the lack of proper sequence identity in the β-sheet regions on the target and CXCL-8 structures, it is not possible to establish a correlation with the sulfate binding region. Human monocyte chemoattractant protein 4 (MCP-4), which makes a CC type of dimer, displays an arrangement of sulfates similar to the CXCL-8 cluster.

MD Simulations of the Complex of CXCL-8 with the Receptor Peptide Fragment. We have also aimed to characterize the interactions of the extracellular fragment of the CXCR-1 receptor, one of the binding partners of CXCL-8, and its associated conformational changes using MD simulations to try to provide the basis for the future structure-based design of chemokine antagonists.

Recent NMR determinations show that the N-terminal domain of the receptor interacts with the entire N-terminal of CXCL-8 through a combination of polar, electrostatic, and hydrophobic/packing interactions and that this binding triggers propagated conformational changes throughout CXCL-8 required for binding affinity and function,[55] in contrast to previous

dx.doi.org/10.1021/ci1003366 |J. Chem. Inf. Model. 2011, 51, 335–358

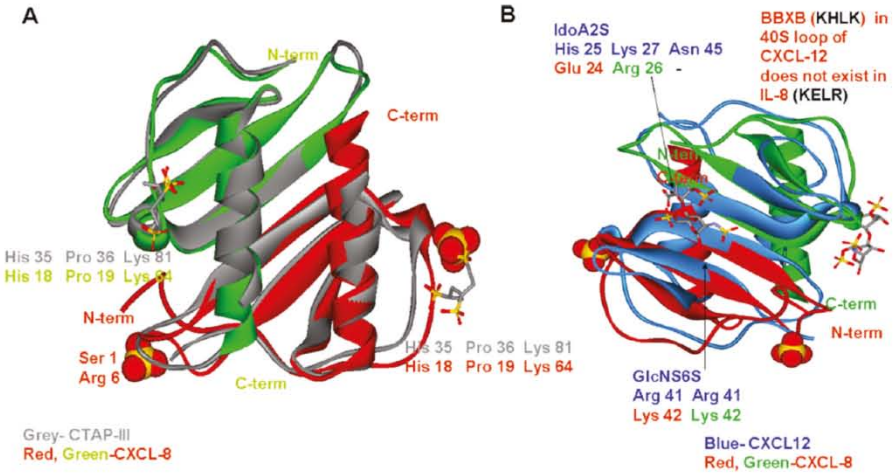


Figure 4. Example of sulfate binding motifs mapped on the CXCL-8 structure 1QE6 based on the superimposition of the target structures (A) CTAP-III and (B) CXCL-12, as obtained by SSM search. Corresponding residues in the target structures are aligned with reference to amino acids in CXCL-8. The protein surface created by the C-terminal α-helix together with the loop connecting the extended N-terminal strand region with the first β-strand of the sheet can interact with sulfates and sulfated ligands. It does not appear possible for a heparin saccharide to bind to CXCL-8 on its β-sheet surface.

suggestions that a subset of highly clustered hydrophobic N-loop residues defines the binding region.[23,144] Furthermore, the N-termini of the receptor peptide and CXCL-8 show characteristics of unstructured or minimally structured domains and hence can form a larger binding region than the folded protein with its restricted conformational freedom. In order to better understand the dynamical behavior of the complex of CXCL-8 with the receptor peptide, we performed unrestrained MD simulations starting with the reported minimized mean NMR structure (PDB code: 1ILQ) of the synthetic receptor peptide (referred to as RP hereafter) and CXCL-8 (Figure 5A), wherein Pro9 and Pro10 of the RP interact with Tyr13, Phe17, Phe21, Leu43, Arg47, and Leu49 of CXCL-8, and Tyr15 of the RP has contacts with Ile10, Lys11, Tyr13, and Leu49 of CXCL-8.[25] In addition to hydrophobic contacts, the ensemble of structures of the RP (PDB code: 1ILP) suggested the presence of intermolecular charge−charge interactions between Asp12, Glu13, and Asp14 of the RP and Lys15, Arg47, and Lys11 of CXCL-8, respectively.

Residues Pro9 and Pro10 of the RP retain their interactions with Tyr13, Tyr15, Phe21, and Arg47 during the MD simulations. Calculations of the SASA and buried surface for individual residues suggested that residues Phe17, Pro19, Ile22, Val27, Ile28, Ile39, Val41, Leu49, Val61, Val62, Phe65, and Ala69 of CXCL-8 are buried. Buried residues Phe17 and Leu49 were not directly involved in interactions with the RP during the simulations, which is consistent with a structural model recently proposed.[55] The intermolecular charge−charge interactions between Glu13 and Asp14 of the RP and Arg47 and Lys11 of CXCL-8, respectively, were observed during the MD simulations; however, the ionic interaction between Asp12 of RP and Lys15 of CXCL-8 were not observed.

Visualization and DSSP analyses of trajectories showed that residues Met1 to Met8 of the RP retained their disordered secondary structure during the MD simulations, consistent with the NMR-determined structural ensemble (Figure 5B). In the initial structure (prior to the production stage of the simulations), the side-chain of Asp5 of the RP makes a salt bridge with Lys20 of CXCL-8. The electrostatic interactions between the side-chain of Asp5 with the Nε1 of His18 of CXCL-8 were observed throughout the production run while the transient electrostatic interactions between the side-chain of Asp6 of RP and the side-chain of Lys64 located in the C-terminal helix of CXCL-8 were observed after 44 ns (Figure 5C). Studies of C-terminus deletion mutants and a CXCL-8 chimera containing helices suggested that the C-terminal α-helix is not directly involved in CXCL-8 receptor binding but helps to stabilize the tertiary structure.[21,145] In contrast, the flexible N-terminus of the RP was found to interact with residues of the C-terminal helix of CXCL-8 during MD simulations. This mechanism of interaction is similar to the one observed in HLA-A2, a structural homologue of CXCL-8, where the antiparallel C-terminal helices of the dimer form the receptor binding site.[26] The nitrogen in the pyrrolidine ring of Trp2 and the backbone carbonyl of Pro10 from the RP make hydrogen bonds with the backbone carbonyl group of Asn71 and the side-chain of Arg47 of CXCL-8, respectively. The hydrogen bond between Trp2 of RP and Asn71 is detected only after 45 ns during the simulation (Figure 5C). In our simulations, a hydrogen bond is also observed between the backbone carbonyl of Tyr15 of the RP and the backbone NH of Cys50 of CXCL-8, suggesting that it plays an important role in receptor binding, as reported in a study of synthetic CXCL-8 analogues.[21] The interactions of Pro17 at the C-terminus of the RP with Gln8, Ile10 and Ile40 of CXCL-8 were retained during the simulations. The interaction between the sidechain of the amine capping residue of the RP with the backbone of Gln8 of CXCL-8 was also retained during the simulation. The time evolution of the backbone rmsds of the whole CXCL-8 dimer, the C-terminal helices of CXCL-8, and the RP fragment are shown in Supporting Information Figure S6, wherein the largest rmsd contribution comes from the flexible RP. The conformation of the C-terminus of the RP is preserved throughout the simulation. MD simulation of the CXCL-8-receptor fragment complex suggest similar conformational changes throughout the CXCL-8 protein (as discussed below in the simulations of the CXCL-8 dimer) and the bound receptor peptide, and also reveal aromatic, hydrophobic, and electrostatic interactions between CXCL-8 and the receptor fragment. These interactions might further be explored for the rational design of novel peptides which

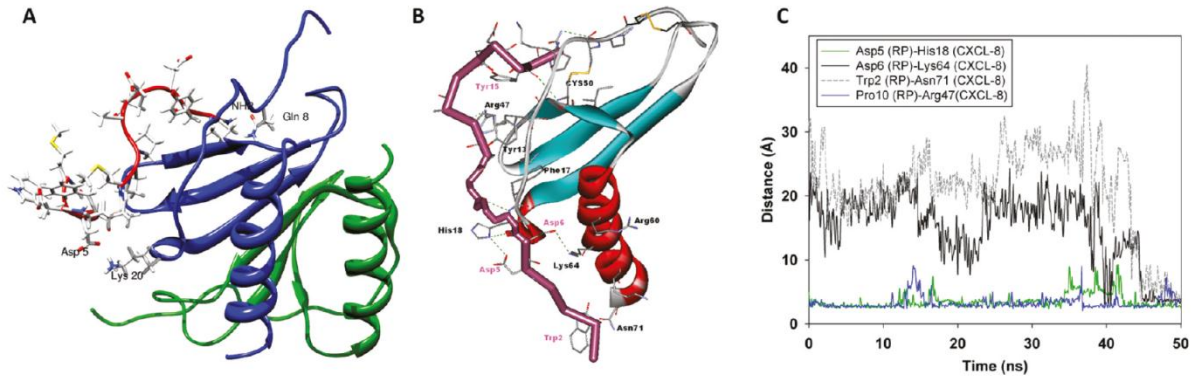


Figure 5. (A) NMR structure of wild-type CXCL-8 in complex with a peptide analogue of CXCR1. The NMR structure reveals the interaction of the receptor peptide and CXCL-8 to be primarily hydrophobic. Two hydrogen bonds are present. The sidechains of Asp5 and the amine capping residue NH_2 of the receptor peptide make hydrogen bonds with Lys20 and Gln8 of CXCL-8, respectively. (B) Snapshot taken from the 50 ns simulation of the CXCL-8—receptor peptide complex. The interactions are highlighted by green dotted lines. The peptide is represented by sticks in magenta, and the CXCL-8 structure is shown according to secondary structure. Only a few residues and monomer of CXCL-8 are shown for clarity. The labels in black indicate residues from the CXCL-8 structure, and the labels in magenta correspond to the synthetic peptide. (C) Plot of the distance between the interacting residues of receptor peptide and CXCL-8. The electrostatic interactions between the sidechain (carboxyl atoms) of Asp5 with the Nε1 of His18 of CXCL-8 were observed throughout the production runs. The electrostatic interactions between sidechain (carboxyl atoms) of Asp6 of RP and sidechain (nitrogen atom) of Lys64 located in the C-terminal helix of CXCL-8 and the hydrogen bond between nitrogen in the pyrrolidine ring of Trp2 of RP and backbone carbonyl of Asn71 of CXCL-8 were transient. The hydrogen bond between backbone carbonyl of Pro10 from RP and the sidechain of Arg47 of CXCL-8 was also observed over the course of the trajectory.

can inhibit formation of the protein complex and serve as a model for the structure-based design of small-molecule analogues to treat inflammatory diseases.

Binding-induced dissociation studies of a trapped disulfide-linked dimer and the wild-type dimer indicated that binding to the receptor N-terminal domain induces dissociation of the dimer—receptor complex into a monomer—receptor complex due to conformational flexibility and that the receptor-bound dimer is thermodynamically unfavored while the receptor-bound monomer is the favored state.[55] In our MD simulations, we could not observe any disruption caused by the N-terminal receptor peptide at the dimer interface, which has extensive H-bonding and packing interactions. It is possible, however, that upon dissociation more residues from CXCL-8 become available to interact with the negatively charged N-terminus of the RP, making the monomer a high-affinity ligand for the CXCR1 receptor.

The free energies of binding predicted from a 30.0 ns MD simulation trajectory using the MM/PB(GB)SA method are reported in Supporting Information Table S2. Large changes in the free energies can be attributed to either solvation effects or electrostatic effects. The change in free energy is dominated by the electrostatic component as opposed to hydrophobic interactions. This is due to the presence of a large number of charged residues (Lys11, His18, Lys20, Arg47, Lys64) of CXCL-8, which provide a sufficient number of electrostatic interaction partners for the receptor peptide. These in silico binding energies cannot be compared quantitatively to experimentally obtained dissociation constant (K_d) values[25] but can be useful to evaluate the relative binding affinities of mutant peptides that may become available in the future. Recently, an empirical approach using the all-atom free-energy force-field PFF02[146,147] was employed to analyze the influence of mutations on the complex of CXCL-8 with the N-terminal peptide of its cognate receptor CXCR1.[148] Seventeen of the individual residues of the N-terminal peptide of CXCR1 receptor were mutated to alanine. Mutations in Met1, Trp2, Gly7, Met8, Pro9, Pro10, Tyr15, Ser16, and Pro17 were shown to have little effect on the interaction energy. In contrast, substantial changes in the interaction energy were seen with substitutions in most of the charged amino acids: Asp3, Phe4, Asp5, Asp6, Asp12, Glu13, and Asp14, the largest being for Asp12. Those calculations were carried out keeping the chemokine backbone rigid while the RP was annealed within its binding region, resulting in an energy-optimized conformation that had a 2.25 Å rmsd from the experimental structure. These results cannot be directly compared with our MD simulations using the MM/PB(GB)SA method because of the fluctuations in the complex and high flexibility in the N-terminus of the RP (rmsd greater than 4 Å), whereas the above Monte Carlo free energy simulations neglect the entropic contributions. Empirical bioinformatics approaches and the MM/PB(GB)SA method benefit from their computational efficiency but are not the methods of choice when large conformational changes in proteins are involved.[149] Consideration of structural flexibility may be better treated through the use of structural ensembles, as in the CC/PBSA method.[150]

MD Simulations of the CXCL-8 Dimer. Earlier 1.0 ns long NVE MD simulations at 300 K of the NMR-determined structure of CXCL-8 using explicit and implicit solvent models predicted the relative motion of the antiparallel α-helices in the dimer.[41] The continuum solvent simulation showed slightly larger deviations from the initial NMR-determined structure, but in general, the implicit and explicit solvent simulations similarly predicted the relative approach of the α-helices, making the structure similar to that determined by X-ray diffraction. These simulations may not have been either long enough to capture the alternating conformational behavior of the α-helices in the CXCL-8 dimer and predict with accuracy a realistic interhelical distance. Moreover, the NMR structure was determined at an acidic pH at 313 K. Consequently, we have performed NPT MD simulations of the CXCL-8 dimer starting with the NMR-determined structure at 300, 313, 325, and 337 K for 100.0 ns with all histidines protonated. Simulations starting with the X-ray structure of CXCL-8 were carried out at 300, 325, and 337 K

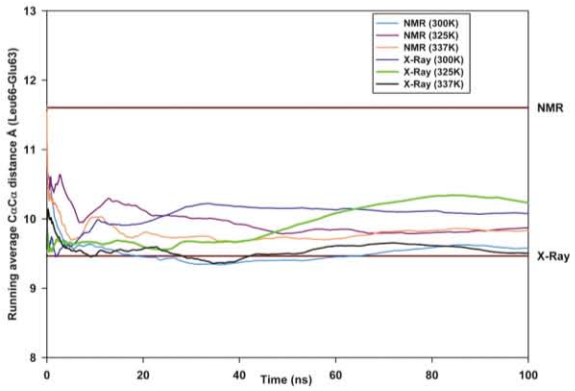


Figure 6. (A) Plots of the running averages of the distance between the Cα atoms of Leu66 on one monomer and Glu63 on the other monomer. The structure of the CXCL-8 dimer was taken from X-ray and NMR structures and was simulated at different temperatures. The reference lines are shown to indicate the reference initial conformations having Cα−Cα distances of 9.47 Å (X-ray) and 11.61 Å (NMR). The coordinates of the CXCL-8 dimer were taken from X-ray and NMR structures and were simulated at different temperatures.

under the same conditions. These simulations were aimed at investigating the interhelical dynamics.

The distance between residues Leu66 of one monomer and Glu63 of the other is approximately 9.47 Å in the X-ray structure and 11.61 Å in the NMR structure. Analysis of the structures in the NMR-determined ensemble (PDB code: 2IL8) revealed an interhelical distance ranging from 11.0 to 12.74 Å. This interhelical distance was monitored during each of our simulations and the running averages are plotted in Figure 6. Although the interhelical distance would be expected to be similar to the one observed in the NMR structure in solution, the simulations in explicit water at different temperatures show that the interhelical distance did not converge to a value close to 11.0 Å. Simulations of the NMR structure at 300 (Figure 6) and 313 K (not shown) reveal that the interhelical distance closely resembled that observed in the X-ray structure. MD simulations at higher temperatures (325 and 337 K) also showed the two antiparallel helices moving closer together and reaching a distance of 9.87 Å, whereas simulations starting from the X-ray structure at 300 and 325 K showed the two helices moving further away from each other. The interhelical distance of the X-ray structure at the end of the simulation at 337 K was close to the initial starting distance of 9.5 Å. This high temperature simulation suggests that increased sampling (as achieved by using a higher temperature or by running a significantly longer simulation) results in a dynamic interhelical distance that nonetheless tends to converge to the distance found in the X-ray structure.

While it is possible to consider alternative measurements of the interhelical distance between equivalent positions along the length of the helices, we considered the distance between residues Leu66 of one monomer and Glu63 of the other for consistency with previous simulations.[41] The interhelical distances between equivalent positions along the length of the helices measured using the Interhlx program[151] were found to be 13.7 and 17.5 Å in the initial X-ray and NMR structures, respectively, which converged to 11.76 (X-ray) and 13.19 Å (NMR) at the end of our simulations at 300 K. Regardless of the choice of interhelical pairs, it is clear that the interhelical distance measured in the NMR structure during the simulations resembled the interhelical distance observed in the X-ray structure.

The relative orientations of the α-helices with respect to each other were measured using the Interhlx program.[151] The interhelical angles (tilt angle between two different helices) were found to be similar: 165.48° and 169.0° in the initial X-ray and NMR structures, respectively. The average interhelical angles at the end of the simulations at 300 K were found to be 159.40° and 167.0° for the X-ray and NMR structures, respectively.

Nuclear Overhauser effect (NOE) restraints involving residues in the region 23−36 and the C-terminal helices were significantly different from those expected in- the X-ray structure of the CXCL-8 dimer.[28,152] This difference in the region 23−36 is, however, of secondary importance for determining receptor binding specificity[63] because the specificity resides in the N-loop region of CXCL-8. The differences in the interhelical region could nonetheless be exploited for the design of inhibitors. Our simulations suggest that the interhelical distance in the X-ray structure can be used to model interactions with molecules such as cyclitols (see further below).

The rmsds of all atoms, backbone atoms, and backbone atoms of the helices and the N-terminus were monitored in each of the simulations of dimeric CXCL-8. The average rmsd of all backbone atoms of the CXCL-8 dimer in MD simulations starting from the X-ray structure was consistently smaller than in the simulations starting from the NMR structure (Figure 7A). The monomers within the dimer were also found to be relatively stable with RMSDs, not exceeding 3.0 Å during the 100 ns runs (Figure 7B). We also monitored the distance between the backbone NH of Gln8 and the Nε2 atom of His33 in both the NMR and X-ray structures of the CXCL-8 dimer, as reported in previous MD simulations.[41] ^{1}H NMR spectroscopy has confirmed the presence of a hydrogen bond between the backbone amide of Gln8 and Cys9 and the carboxylate in Glu38; however, these interactions, although present in the N-terminus, are not relevant for neutrophil activation.[153,154]

Findings from both our MD simulations and those earlier reported[41] rely on the accuracy of the force fields used. Use of the AMBER force field ff94[155] (as used in the earlier MD study of CXCL-8)[41] and its variants[156] is known to lead to the overstabilization of helical systems and the adoption of stable helices for sequences that have been experimentally determined not to have helical structures.[157] On the other hand, use of the AMBER ff96 force field has been observed to overestimate β-strand propensity.[158] AMBER ff99SB, used in this study, is known to have some limitations for the modeling of the sidechains of Ile, Leu, Asp, and Asn in helical peptides.[159−161] Nevertheless, other works have shown good agreement between simulations using ff99SB and NMR structural and relaxation data compared to other AMBER variants,[89,160,162,163] making it an appropriate choice for this work.

Changes to the secondary structure of the CXCL-8 dimer during the simulations were also monitored to help to identify the region which contributes the most to the conformational flexibility of this structure. This analysis was carried out using the DSSP method of Kabsch and Sander[164] for every snapshot collected every 500 ps, as discussed above, monitoring the simulations at 300 K starting from both the NMR and X-ray structures, and is shown in Supporting Information Figures S7 and S8, respectively. The results were mapped onto the corresponding 3D-structures, as shown in Figure 8, where the regions

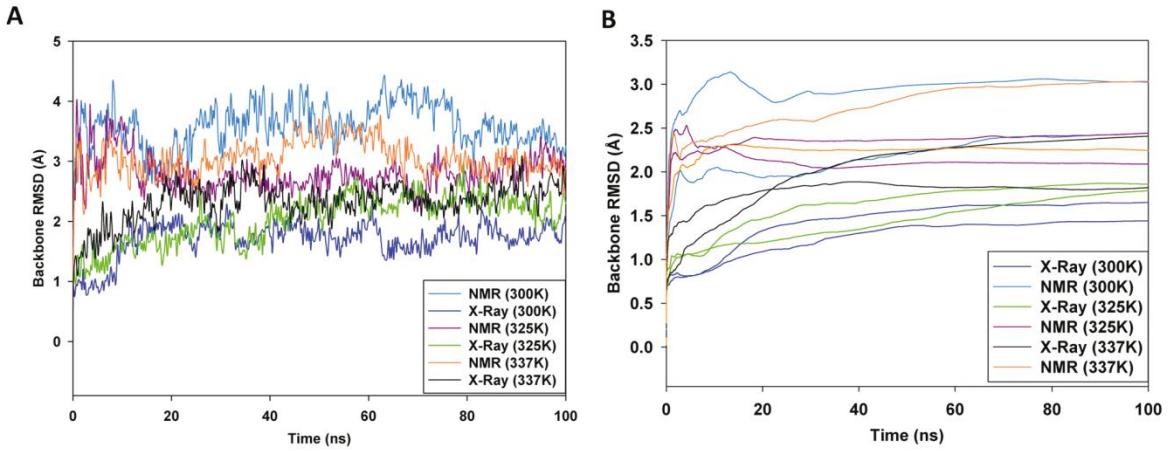


Figure 7. (A) Backbone average rmsd of the CXCL-8 dimer at different temperatures in simulations starting with the NMR and X-ray structures. All backbone rmsd in MD simulations starting from the X-ray structure were consistently smaller than those starting with the NMR structure. (B) Backbone average rmsd of the individual monomers present in the CXCL-8 dimer at different temperatures in simulations starting with the NMR and X-ray structures.

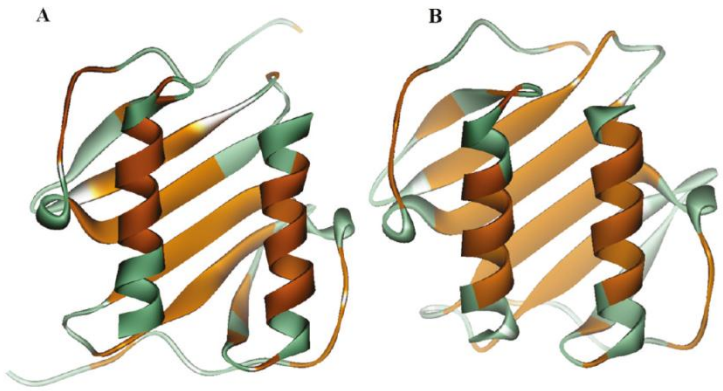


Figure 8. DSSP analysis monitored during 100 ns simulations at 300 K starting from (A) NMR and (B) X-ray structures mapped on the CXCL-8 3D structure. The regions which experience significant conformational change during the MD simulations are shown in cyan. The regions that preserve their conformation are shown in gold. Significant changes in the secondary structure are observed in the N-terminal loop, the 3_{10}-helix, the 30s, 40s, and 50s loops, and the third β-sheet. The secondary structure of β-sheets connecting the two monomers were dynamically conserved during the simulations.

that experience significant conformational changes during the simulations are shown in cyan and the regions that preserve their conformation are shown in gold. The secondary structure of the β-sheets connecting the two monomers was "dynamically" conserved during the simulations, most likely due to the presence of intermolecular hydrogen bonds. Significant changes to the secondary structure were observed in the N-terminal loop, the 3_{10}-helix, the 30s, 40s, and 50s loops and the third β-sheet, in agreement with experimental findings showing that the structural differences seen between the X-ray and NMR structures are related to the existence of various conformations of the N-terminal loop and the 31 to 36 loop, both of which display different interactions in the two structures with respect to the ends of the central β-strands.[152] Furthermore, our simulations also indicate that the conformational flexibility in these regions might contribute to a shearing movement in the C-terminal helices of CXCL-8. This suggests that flexibility in the CXCL-8 dimer can be attributed to the concerted rather than independent motions in regions of secondary structure, as reported previously.[165]

NMA was used to analyze the motions of the α-helices in the NMR and X-ray structures of the CXCL-8 dimer. The low frequency, high amplitude vibration modes characterized by this approach were used to reveal the large scale movements associated with the relative motion of the two monomers. The output of these calculations in PDB format corresponding to the sixteenth and eighteenth lowest frequency modes are provided in the Supporting Information as ci1003366_si_002.pdb and ci1003366_si_003.pdb. These modes were selected because they exhibited the highest degree of collectivity, as a measure of the fraction of residues that are significantly affected by a given vibrational mode. Visual inspection of these vibrations revealed that the conformational flexibility in the 3_{10}-helix and third β-sheet accompanies the shearing movement in the C-terminal helices. This is consistent with our previous analysis of changes to the secondary structure during the MD simulations. On the basis of the results of our MD simulations and secondary structure and NMA analyses, it seems reasonable to propose the existence of shearing movements in CXCL-8 and, in particular, low-frequency conformational

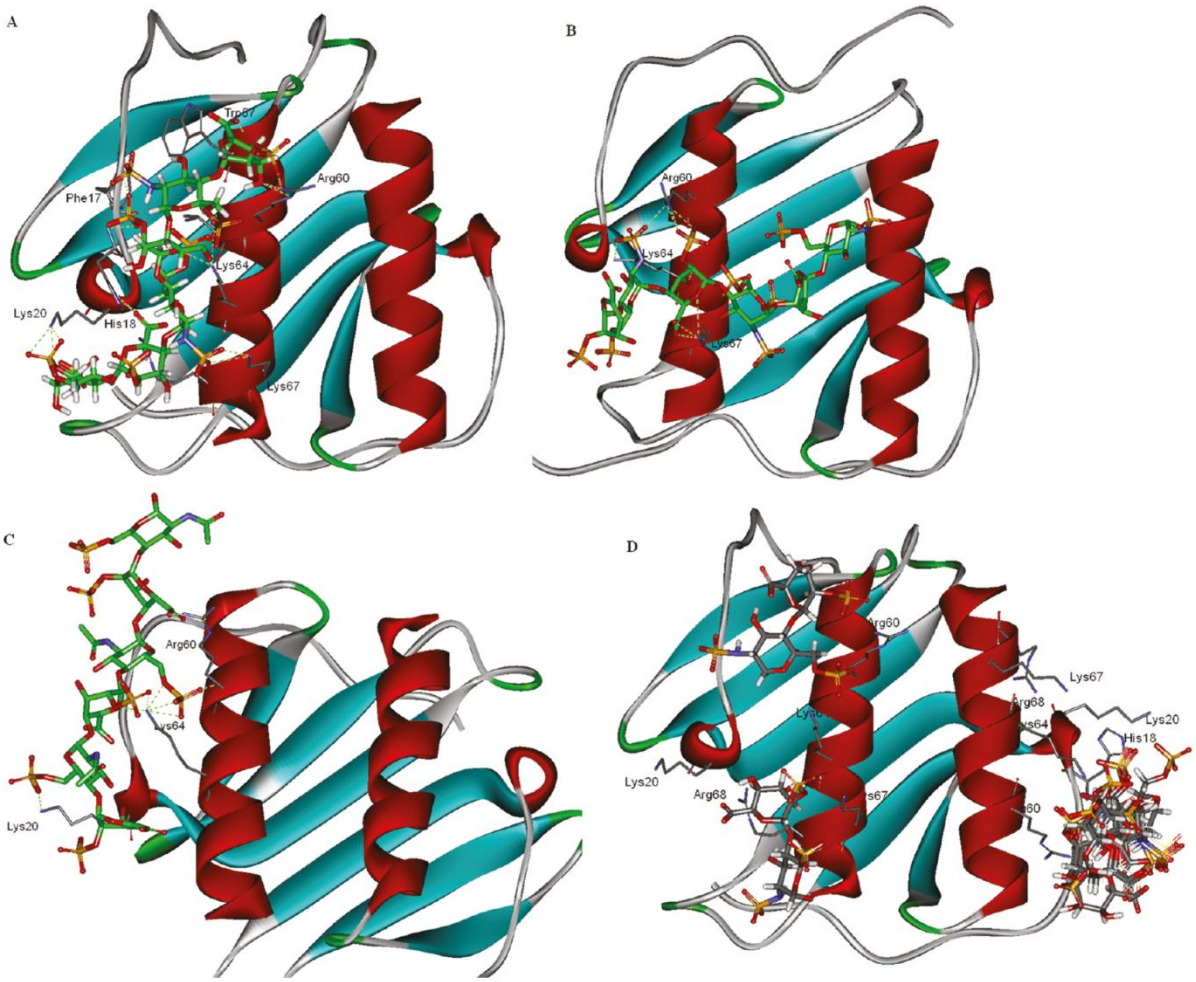


Figure 9. Binding mode and interactions of heparin with CXCL-8 predicted by docking simulations. Interacting residues are represented as sticks. (A) Interaction of Heparin H1 parallel to the C-terminal helical axis. (B) Interaction of Heparin H1 perpendicular to the C-terminal helical axis. (C) Interaction of Heparin H2 with the CXCL-8 monomer, indicating that O-sulfation is important for binding. (D) Cluster of docked conformations of disaccharide H9 observed on the GAG binding surface of CXCL-8, indicating the presence of a multivalent binding mode.

motions in the C-terminal helices. This conformational behavior might be exploited for the rational design of inhibitors bridging the dimer structure of CXCL-8.

Molecular Docking of Heparin Fragments. Free energy calculations using the MM/PB(GB)SA approach have also been successfully used to rank the relative binding affinities of several glycan−protein complexes.[166] However, MM/PB(GB)SA calculations are known to overestimate the free energies of binding of heparin-protein complexes[167] as none of the current PB or GB implicit solvation methods have been developed to specifically treat the desolvation energies and polarization effects arising from the polyanionic nature of GAGs. Studies have shown that AutoDock is effective in studying the docking of highly charged large molecules such as GAGs.[66,86] Consequently, docking of the heparin analogues reported in Table 1 was considered to be a more feasible approach to predict free energies than MM/PB(GB)SA. These analogues were blind docked (site 3) to the whole surface of CXCL-8. During the docking simulations, clustering of docked poses at site 1 (parallel to the helical axis in the monomers) and site 2 (perpendicular to the helical axis in the dimer) were observed (Figure 9A and B). The interactions of various heparin analogues with CXCL-8 and the binding affinities predicted by AutoDock for the lowest energy poses are reported in Table 4.

Previous docking simulations reported that the lowest energy conformation of a heparin hexasaccharide has a perpendicular orientation and bridges the gap between the two antiparallel α-helices.[66] A consistent set of residues (His18, Lys20, Lys64, Lys67, and Arg68) were identified as possible interaction partners for the heparin oligosaccharides in our docking simulations. In contrast to the previous study,[66] all the low energy conformations of sulfated heparin were observed to cluster parallel to the C-terminal helical axis of the X-ray structure of CXCL-8 (Figure 9). No docking simulations were performed with the NMR structure of monomeric CXCL-8 (PDB code: 1IKL), where the amide of Leu25 was modified into N-methyl (L25NMe).[27] The main difference between this monomer and

Table 4. Docking of Heparin Fragments to CXCL-8[a]

heparin no.	AutoDock energy of binding (kcal/mol)	K_i	lowest energy binding modes	interactions with residues of CXCL-8 within 3.5 Å
H1	−10.19	33.82 nM	parallel	sidechains of His18, Lys20, Arg60, Lys64, Lys67, and Arg68, backbone of Phe17 and Trp57
	−4.72	346.62 μM	perpendicular	sidechains of Arg60, Lys64, and Lys67
H2	−8.41	681.43 nM	parallel	sidechains of Lys20, Arg60, and Lys64
H3	−3.99	1.19 mM		sidechains of Arg60, Lys64, backbone of Pro16, Phe17
H4	−4.03	1.11 mM		sidechain of Arg60, backbone of Phe17
H5	−3.33	3.65 mM	parallel	sidechains of His18, Arg60, Lys64, and Arg68
H6	−4.58	438.21 μM	parallel	sidechains of His18, Arg60, and Lys64, backbone of Arg68
H7	−5.82	54.09 μM	parallel	sidechains of His18 and Lys64
H8	−3.26	4.07 mM		backbone of Pro16 and Phe17
H9	−4.89	259.67 μM		sidechains of His18, Lys20, Arg60, Lys64, Lys67, and Arg68
H10	−4.41	587 μM		sidechain of Lys 64 or sidechain of Arg47
H11	−4.25	763 μM		sidechain of Glu63
H12	−4.26	754 μM		sidechain of Glu63

[a] Parallel binding mode: heparin axis is parallel to the C-terminal helical axis of CXCL-8. Perpendicular binding mode: heparin axis is perpendicular to the C-terminal helical axis of CXCL-8.

the X-ray and NMR structures of the dimer is in the GAG-binding site in the C-terminal helix, which leads to significant differences in the results of docking.

In this present study, docking simulations of a heparin disaccharide resulted in many binding pose clusters in the vicinity of basic residues, suggesting that heparin disaccharides can bind the CXCL-8 monomer or dimer in multivalent mode. Heparin hexasaccharides H1 and H2 were predicted to make the majority of their interactions with the heparin binding site (basic residues Lys20, Arg60, and Lys64) through their *O*-sulfate groups (Figure 9A and C), whereas the *N*-acetylated group was not found to interact with the heparin binding site. Hence, on the basis of these docking simulations of the two heparin hexasaccharides reported here, it seems that O-sulfation is more important than N-sulfation for the binding to CXCL-8, consistent with experimental data.[58] Only a monosaccharide unit in heparin fragments H3, H4, and H7 was predicted to make interactions with the GAG binding site, whereas a disaccharide unit present in H5 was predicted to make electrostatic interactions with basic residues in the GAG binding site. Similar docked poses were predicted with the NMR structure.

The predicted measures of binding affinity determined by the AutoDock scoring function were not of the same magnitude as the experimentally determined activity of molecules H5, H6, H7, and H9, although the free energies predicted could differentiate and rank accurately compounds with different degrees of sulfation. Ligand-specific scoring functions such as BALLDock/SLICK[168,169] and others[170] focused on carbohydrates (simple sugars) have been developed. These scoring functions will need to be calibrated before they can be used for predicting free energies of binding to proteins of sulfated GAGs.

There does not appear to be a simple correlation between the length of the heparin fragments and their free energies of binding in our docking studies, and hence, it was not possible to resolve the controversy regarding the optimum and minimum lengths of the oligosaccharide required for binding to CXCL-8 (vide supra).[39,58,64] These docking studies suggest that sulfated heparin saccharides with six or less units bind to monomeric CXCL-8 rather than the dimeric structure because heparin fragments (H1−H12) were seen to bind parallel to the helical axis in the monomers only.

Molecular Docking and Free Energy Calculations of Complexes of Cyclitols with CXCL-8. MD simulations of the CXCL-8 dimer reported above advocated the use of the X-ray structure of the CXCL-8 dimer for docking simulations. Hence, the cyclitols reported in Table 2 were docked to sites 1 and 2 of the CXCL-8 X-ray structure. The interactions of various cyclitols with the CXCL-8 monomer (site 1) predicted by these docking simulations are reported in Table 5. It can be seen that the free energy scores and the binding modes predicted by docking to the CXCL-8 monomer cannot differentiate between the active cyclitols 8, 9, 10, and 10a and other inactive ligands (2−7 and 11−14). For example, cyclitol 3 is predicted to bind with higher affinity to monomeric CXCL-8, in contrast to experimental data that shows that this molecule is incapable of blocking heparin−CXCL-8 interactions.[76] These results suggest that cyclitols may in fact interact with CXCL-8 through another binding site.

The interactions of cyclitols 2−14 in binding site 2 of dimeric CXCL-8 are reported in Table 5. While the scoring function energies cannot be correlated to earlier-reported SAR studies,[76] short cyclitols (2−6) are correctly predicted not to bind to CXCL-8. Visualization of the binding modes shows that cyclitols with linker size >7 are able to bridge the two monomers A and B of the CXCL-8 dimer. The hydrophobic linker of size 8 lies perpendicular to the two monomeric chains A and B, and the sulfates make electrostatic interactions with basic residues in the C-terminal α-helices of both monomers, which form the GAG binding site (Figure 10). Similar binding modes were observed for cyclitols 9, 10, and 10a. The sulfates in cyclitols 2−7 and cyclitols 11−12 appear to have a higher number of interactions with the basic residues of one monomer compared to other monomer because the hydrophobic linker in these cyclitols is too short to bridge the two monomers. The sulfonate group (SO_3^-) in cyclitols 13 and 14 might reduce binding due to the presence of a negatively charged region near Glu63. The present docking simulations suggest that not all sulfates on the 6-membered cyclitol ring interact with residues of the CXCL-8 dimer. We were not able to obtain a better binding mode with the NMR structure of the CXCL-8 dimer due to differences in the

Table 5. Docking of Cyclitols to Monomeric and Dimeric CXCL-8

cyclitol no.	AutoDock energy of binding (kcal/mol)	K_i	number of interacting sulfates	interactions between cyclitol sulfates and the sidechains of residues of CXCL-8 within 3.5 Å
		docking of cyclitols to CXCL-8 monomer (site 1)		
2	−8.09	1.17 μM	2	Arg60, Lys64
3	−9.94	51 nM	5	Lys20, Lys64, Lys67, Arg68, Asn71
4	−5.39	111.74 μM	4	Lys20, Trp57, Arg60
5	−2.64	11.54 mM	2	Asn56, Lys64
6	−6.45	18.86 μM	3	Lys20, Arg45, Lys64
7	−6.88	9.06 μM	5	His18, Lys67, Lys64, Asn71
8	−6.10	33.76 μM	5	His18, Arg60, Lys64, Lys67, Asn71
9	−4.01	1.14 mM	6	His18, Arg60, Lys64, Lys67, Asn71
10	−4.37	626 μM	4	Lys64, Lys67, Arg68, Asn71
11	−7.57	2.82 mM	3	His18, Arg60, Lys64
12	−5.82	54.30 mM	2	Ser44, Arg47
13	−7.31	4.39 mM	3	His18, Lys20, Lys64
14	−4.09	999.56 mM	3	Phe17, His18, Lys64
6a	−7.14	5.86 μM	3	Lys64, Lys67
10a	−6.67	13.01 μM	6	His18, Lys20, Arg60, Lys64
		docking of cyclitols to CXCL-8 dimer (site 2)[a]		
2	+0.28		4	Arg60(A), Lys67(A), Arg60(B)
3	+0.34		3	Lys67(A), Asn56(B), Lys67(B)
4	−0.68	317.86 mM	5	Arg60(A), Lys64(A), Lys67(A), Arg60(B), Lys 67(B)
5	+1.24		3	Arg60(A), Lys64(A), Lys67(A)
6	+0.57		4	Arg60(A), Lys64(A), Lys67(A), Lys67(B)
7	−4.83	286.93 μM	4	Asn56(A,) Arg60(A), Lys67(A), Lys67(B)
8	−6.28	24.96 μM	6	Arg60(A), Lys64(A), Lys67(A), Lys67(B), Lys64(B), Arg60(B)
9	−4.86	273.86 μM	5	Lys64(A), Lys67(A), Lys67(B), Arg60(B)
10	−3.67	2.04 mM	4	Lys64(A), Lys67(A), Lys67(B), Arg60(B)
11	−1.71	55.36 mM	4	Arg60(A), Lys64(A), Lys67(A), Lys67(B)
12	−2.18	25.30 mM	3	Asn56(A), Arg60(A), Arg60(B)
13	+0.19		5	Arg60(A), Lys67(A), Asn56(B), Arg60(B),
14	+0.43		5	Arg60(A), Lys67(A), Asn56(B), Lys64(B), Lys67(B)
6a	−4.31	688.34 μM	7	Arg60(A), Lys67(A), Arg60(B), Lys67(B)
10a	+2.87		9	Arg60(A), Lys64(A), Lys67(A), Arg60(B), Lys64(B), Lys67(B)

[a] (A) and (B) refer to each monomer of the CXCL-8 dimer.

interhelical distance, the packing interaction between the monomers and the differences in the GAG binding site (vide supra).

MM/PB(GB)SA calculations were performed for cyclitols 7, 8, 9, 10, and 10a docked at binding site 2 of the CXCL-8 dimer. The complex of cyclitol 7 and CXCL-8 was taken as reference to establish the SAR between active and inactive molecules. The predicted MM/PB(GB)SA free energy components are reported in Table 6. Cyclitol 7 was observed to drift away from the binding site within 1 ns after removing constraints and then align itself near the C-terminal helix of a single monomer of CXCL-8. The ligand appears to be too short to bridge the two helices of the CXCL-8 dimer. Free energies of binding obtained with the MM/PBSA method can clearly differentiate between cyclitols 8, 9, and 10. Although these free energies of binding cannot be directly compared to experimental inhibition constants, they correctly rank cyclitol 7 as the more active molecule compared to cyclitols 9 and 10. In contrast to experimental data, the complex of cyclitol 10a with CXCL-8 is predicted to have the highest binding affinity. This appears to be due to a higher degree of sulfation, resulting in an increased nonbonded electrostatic interaction energy. Free energies of binding obtained with the MM/GBSA approach were positive in value for the interaction of cyclitol 10 with the CXCL-8 dimer, due to a positive value of GBELE (Table 6). Exclusive of entropy terms, the GBSA method correctly ranks cyclitol 8 as the most active out of all cyclitols and cyclitol 10 as having a comparable level of inhibition to cyclitol 10a, whereas the PBSA method overestimates the binding affinity of cyclitol 10a, as compared to the experimental data that shows that cyclitol 8 is the most active molecule.

CXCL-8 monomers and dimers exist in vivo and their differential binding to CXCR2 and GAGs mediates and regulates in vivo neutrophil recruitment.[32] However, the CXCL-8 monomer−dimer equilibrium is disturbed in acute or chronic inflammation and tissue injury, which can lead to the increase in dimer concentration and, consequently, to persistent neutrophil infiltration.[32] Therefore, in contrast to cyclitols, the design of small molecules that can bind to CXCL-8 dimer interface and consequently impair potential to form dimers will be important for the future rational design of drugs against a wide variety of neutrophil-mediated inflammatory diseases.

Molecular Docking and Molecular Dynamics of the CXCL-8 Dimer in Complex with a 24-mer Heparin Fragment. The 24-mer heparin fragment was divided into two sulfated hexasaccharides consisting of IdoA2S-GlcNS6S residues and two nonsulfated hexasaccharides consisting of GlcNAc-GlcA. The heparin hexasaccharide was initially docked to one of the monomers taken from the structure of the CXCL-8 dimer. Another heparin hexasaccharide was docked into the second, symmetry-related binding site of the other CXCL-8 monomer. The two heparin hexasaccharides were then extended and manually linked with a stretch of two nonsulfated hexasaccharides to form a horseshoe pattern (Figure 11). The interactions of the sulfated heparin fragments with CXCL-8 predicted by docking are reported in Table 7. Most of the interactions of the heparin fragment with CXCL-8 are predicted to be mediated through O-sulfation. It is not meaningful to measure the binding affinity to the CXCL-8 dimer of a 24-mer heparin fragment using the AutoDock scoring function. This scoring function includes a term for the number of rotatable bonds in the ligand (and/or the protein) as an approximation to the entropic cost arising from the loss of conformational degrees of freedom upon binding.[105] However, if the number of rotatable bonds is too large (in this case >60), this term becomes too large, resulting in abnormally large and positive predicted free energies of binding.

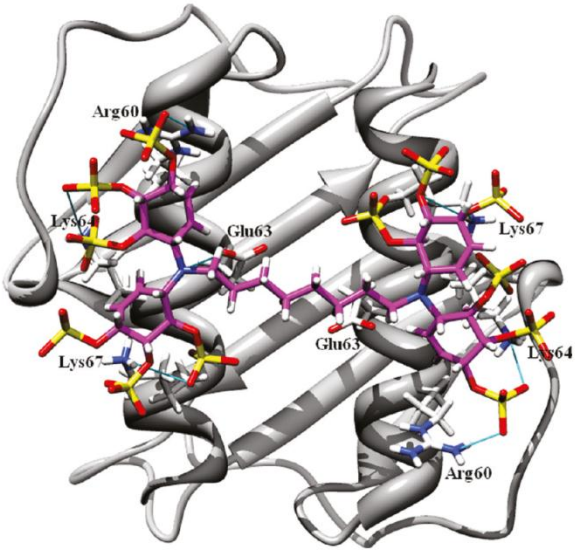


Figure 10. Binding mode and interactions of cyclitol 8 with the CXCL-8 dimer predicted by docking simulations. Interacting residues are represented as sticks. Electrostatic interactions are indicated by blue lines.

MD simulations were then carried out for the complex of the 24-mer heparin fragment with the CXCL-8 dimer. The all-atom rmsd of the protein as well as the sulfated and nonsulfated hexasaccharides were monitored during the simulations (Supporting Information Figure S9A). The nonsulfated fragments were seen to move away from the protein compared to the sulfated fragments, which form direct electrostatic interactions with the protein in this complex. Limited structural change (rmsd of 2.5 Å) was observed within the protein, while the potential energy (Supporting Information Figure S9B), density and temperature of the system were stable during the simulations. On the basis of the measured structural variance and visualization of MD trajectories, an ensemble consisting of at least four sulfated residues in the chair conformation was found to be important for binding to the C-terminal helices of CXCL-8. Our simulations are not consistent with previous MD simulations of this complex where the interaction of CXCL-8 with a single sulfated

Table 6. MM/PB(GB)SA Energy Component Analysis of the Interactions of Cyclitols with the CXCL-8 Dimer Averaged over 8.0 ns[a]

	Δ[b]CXCL-8-cyclitol 8		Δ[b]CXCL-8-cyclitol 9		Δ[b]CXCL-8-cyclitol 10		Δ[b]CXCL-8-cyclitol 10a	
	mean	std	mean	std	mean	std	mean	std
ELE[d]	−2392.28	88.62	−2496.87	99.64	−3087.14	73.98	−4951.51	102.78
VDW[e]	−28.75	5.4	−38.73	4.33	−32.88	6.53	−22.56	5.93
INT[f]	0	0	0	0	0	0	0	0
GAS[g]	−2421.04	87.82	−2535.59	100.5	−3120.02	75.97	−4974.07	103.6
PBSUR[h]	−5.68	0.3	−6.52	0.22	−5.99	0.5	−5.43	0.4
PBCAL[i]	2331.11	78.9	2455.45	90.46	3039.81	72.39	4858.06	96.3
PBSOL[j]	2325.43	78.17	2448.93	90.41	3033.82	72.09	4852.63	96.16
PBELE[k]	−61.17	16.78	−41.42	17.14	−47.33	15.69	−93.45	17.53
PBTOT[l]	−95.6	15.45	−86.67	17.04	−86.2	12.75	−121.44	16.23
GBSUR[m]	−5.68	0.3	−6.52	0.22	−5.99	0.5	−5.43	0.4
GBCAL[n]	2374.28	81.44	2494.62	93.97	3091.32	73.08	4945.81	99.5
GBSOL[o]	2368.6	81.43	2488.1	93.92	3085.32	72.79	4940.38	99.35
GBELE[p]	−18	11.74	−2.25	9.55	4.17	12.11	−5.7	13.57
GBTOT[q]	−52.43	10.72	−47.5	9.51	−34.7	8.81	−33.68	11.57
$T\Delta S^{c,r}$	−41.09		−42.16		−40.10		−46.84	
$\Delta G_{\text{binding-PBSA}}{}^{s}$	−54.51		−44.51		−46.1		−74.6	
$\Delta G_{\text{binding-GBSA}}{}^{s}$	−11.34		−5.34		+5.4		−13.16	

[a] Average over 800 snapshots of the MD simulation trajectory. [b] Δ is the average free energy of binding. [c] Entropy calculations were based on normal modes analysis using 40 snapshots. [d] Nonbonded electrostatic energy. [e] Nonbonded van der Waals energy. [f] Bond, angle, dihedral energies. [g] ELE + VDW + INT. [h] Hydrophobic contribution to solvation free energy for PB calculations. [i] Reaction field energy calculated by PB. [j] PBSUR + PBCAL. [k] PBCAL + ELE. [l] PBSOL + GAS. [m] Hydrophobic contributions to solvation free energy for GB calculations. [n] Reaction field energy calculated by GB. [o] GBSUR + GBCAL. [p] GBCAL + ELE. [q] GBSOL + GAS. [r] T = temperature, ΔS = sum of rotational, translational, and vibrational entropies. [s] Total free energy of binding.

dx.doi.org/10.1021/ci1003366 |J. Chem. Inf. Model. 2011, 51, 335–358

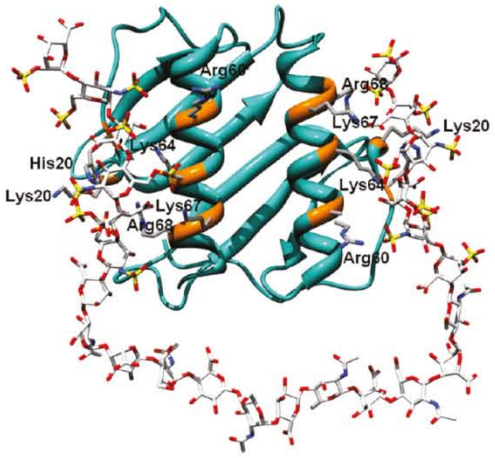


Figure 11. Interactions of the 24-mer heparin fragment with the CXCL-8 dimer. Interacting residues are represented as sticks. Only basic amino acids within 3.5 Å of the fragment are shown.

Table 7. Interactions of the Sulfated Domain of a 24-mer Heparin Fragment with the CXCL-8 Dimer

residues of the sulfated saccharides	interaction with residues of CXCL-8
IdoA2S (COO^-)	sidechain of Lys64,
IdoA2S (2-*O*-sulfate)	sidechain of Lys20, sidechain of Lys64
GlcNS6S (6-*O*-sulfate)	sidechain of Arg68, backbone of His18, sidechain of Lys64
GlcNS6S (*N*-sulfate)	

disaccharide unit from the 24-mer fragment had been retained at the end of a 5.0 ns simulation.[64] In our study, the sulfated heparin fragments retain the electrostatic interactions with the proximal loop and C-terminal helices over 50.0 ns of simulation time. This discrepancy may be due to differences in docking approaches, as the selection of a badly docked conformation of a long heparin fragment may result in the ligand moving away from the binding site during an MD simulation. On the other hand, our simulations do agree with the observation that a single disaccharide unit with 2- and 6-sulfate groups is sufficient to achieve high affinity of binding to CXCL-8 in a multivalent model, as shown in Figure 9D. The low value of the free energy of binding predicted using the MM/PB(GB)SA method (Supporting Information Table S3) suggests that the most important energy contributions come from the interactions of various monosaccharide units present in the 24-mer fragment with CXCL-8 during the simulation, rather than just the interaction of a single disaccharide unit with the CXCL-8 dimer.

The predicted free energies of binding of CXCL-8 with cyclitols, a receptor fragment, and a 24-mer heparin fragment, as documented in Tables 5, S2 and S3 (of the Supporting Information), respectively, cannot be compared quantitatively with experimentally obtained K_d values. This can be attributed to the approximations involved in the MM/PB(GB)SA method and, more importantly, to the lack of reparameterization in the current PB or GB methods to deal with the desolvation and entropic terms of highly charged sulfated carbohydrates such as GAGs.[166,167] Moreover, inaccurate representation of the internal dynamics of carbohydrates, is likely to lead to insufficient sampling of glycosidic linkage and ring conformations that contribute to the free energy of binding.[171] A further approximation arises from the choice of treatment of 1−4 nonbonded interactions in force fields during MD simulations of glycoproteins.[172] While default 1−4 scaling may be appropriate for the protein, this may be inaccurate for the treatment of nonbonded interactions in carbohydrate force fields, which would affect the accuracy of predictions of free energies. In the context of proteins, the significant difference in free energies of binding with respect to the experimental values may arise due to an incorrect balance of enthalpic and entropic contributions in simulations using the AMBER ff99SB force field.[173] In addition, estimates of the free energy, enthalpy, and entropy are also dependent on the choice of water potential.[174]

CONCLUSIONS

In this paper, various studies of CXCL-8 performed using sequence, structure, and molecular dynamics methods are reported. Evolutionary sequence conservation analysis showed that the GAG binding site is highly variable, suggesting that CXCL-8 across different species may bind GAGs with different sulfation patterns and with different affinity. Residues in the third β-sheet of CXCL-8 were found to be highly variable, while the second β-sheet appears highly conserved. Residues in the first β-sheet, which forms the dimer interface, appear to be partially conserved, based on an analysis of the CXCL-8 monomer. A survey of sulfate-binding regions in the PDB database revealed several occurrences of positively charged regions on the protein surface which form the binding site for sulfated GAGs such as heparin/HS, in agreement with the GAG binding site reported on the basis of analysis of mutagenesis data.

The interaction of CXCL-8 and its receptor CXCR-1 was examined using the experimental structure of CXCL-8 bound to a synthetic 17-amino acid peptide derived from CXCR1. A large degree of structural disorder for residues Met1−Met8 of CXCR1 was observed during MD simulations, indicating that this part of the peptide is bound loosely to CXCL-8. Our simulations suggest that, apart from hydrophobic interactions, aromatic and electrostatic interactions also play a role in mediating contacts between CXCL-8 and the synthetic receptor peptide and that it may be worth re-examining the role of these interactions experimentally. These findings may help to identify peptides derived from one of the protein binding partners with enhanced affinity and serve as starting point for the development of protein−protein inhibitors. Such molecules may have a role in inflammation by inhibiting the superfluous association of neutrophils to the endothelial cell surface, thus reducing damage to host tissue.

MD simulations of the NMR structure of the CXCL-8 dimer revealed that the interhelical distance approaches the interhelical distance observed in the X-ray structure, suggesting that the latter is a better representation of the structure of the dimer, which can be used for designing ligands such as cyclitols. The simulations also showed changes in the secondary structures of the N-terminal loop, the 3_{10}-helix, the 30s, 40s, and 50s loops, and the third β-sheet of CXCL-8 in a concerted manner, accompanying the shearing movements of these C-terminal helices in the CXCL-8 dimer.

Molecular docking of heparin fragments showed low energy docking poses of sulfated heparin being clustered in locations parallel to the C-terminal helical axis of the X-ray structure of

CXCL-8. Docking of disaccharide fragments resulted in many docking pose clusters around basic residues, suggesting that they can bind either the CXCL-8 monomer or dimer in multivalent mode. The docking simulations also suggested that most of the interactions of heparin fragments with the GAG binding site of CXCL-8 are mediated through *O*-sulfate groups. It is proposed that sulfated heparin saccharide fragments (with six or fewer sugar units) bind to monomeric CXCL-8, whereas longer heparin fragments can bridge the two dimers. The binding of cyclitols to CXCL-8 requires a dimer, and this binding pattern appears to be analogous to the proposed horseshoe binding pattern of a 24-mer heparin fragment. Our simulations provide a rationale for the optimal number of sulfates and the length of alkyl spacers required for the interaction of this type of inhibitors with the dimeric form of CXCL-8. In summary, these modeling studies suggest a possible avenue for therapeutic applications as they predict that the conformational changes of dimeric CXCL-8 may be relevant for the rational design of CXCL-8 inhibitors (carbohydrate or peptide based).

■ ASSOCIATED CONTENT

Supporting Information. Accession numbers for CXCL-8 from various species and percentage identity with respect to human CXCL-8 (Table S1), MM/PB(GB)SA energy component analysis of the interactions of the CXCL-8 dimer with the receptor peptide averaged over 50 ns (Table S2), MM/PB-(GB)SA energy component analysis of the interactions of the CXCL-8 dimer with a 24-mer heparin fragment averaged over 50 ns (Table S3), chemical structure of the 24-mer heparin fragment (Figure S1), amino acid residue conservation in the structure of CXCL-8 (PBD code: 3IL8) obtained using default Consurf parameters (Figure S2), multiple sequence alignment of CXCL-8 homologous sequences obtained using Consurf DB (Figure S3), multiple sequence alignment of CXCL-8 protein sequences collected from SWISS-PROT using a BLAST search as mentioned in Table S1 (Figure S4), ribbon diagrams highlighting those residues in the CXCL-8 dimer (taken from the X-ray structure, PDB code: 3IL8) involved in the binding to receptors and GAGs (Figure S5), time evolution of the average backbone rmsd of the whole CXCL-8 dimer, the two C-terminal helices of CXCL-8 and the receptor peptide (Figure S6), secondary structure "dynamic conservation" analysis using the DSSP method in the simulation starting from the NMR structure (Figure S7), secondary structure dynamic conservation analysis using the DSSP method in the simulation starting from the X-ray structure (Figure S8), time evolution of the average all-atom RMSD of the CXCL-8 dimer, the two sulfated and two non-sulfated regions of the 24-mer heparin fragment. The non-sulfated region of heparin is more flexible compared to the sulfated regions, which are bound to the protein in the CXCL-8–heparin complex (Figure S9). The output of normal modes calculations in PDB format corresponding to the sixteenth and eighteenth lowest frequency modes are provided as ci1003366_si_002.pdb and ci1003366_si_003.pdb. The Consurf and MM/PB(GB)SA methods are also described in detail. This information is available free of charge via the Internet at http://pubs.acs.org

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: r.mancera@curtin.edu.au. Phone: +61 8 9266 1017. Fax: +61 8 9266 2342.

■ ACKNOWLEDGMENT

N.S.G. is grateful for the award of an Endeavour International Postgraduate Research Studentship. We also gratefully acknowledge the Western Australian Interactive Virtual Environments Centre (IVEC) and the National Computational Infrastructure (NCI) Facility for access to high-performance computing.

■ REFERENCES

(1) Iizasa, H.; Matsushima, K., IL-8. In *Cytokine reference: a compendium of cytokines and other mediators of host defense*; Oppenheim, J. J., Feldmann, M., Durum, S. K., Hirano, T., Vilcek, J., Nicola, N. A., Eds.; Academic Press: New York, 2001; Vol. I, pp 1061–1068.

(2) Baggiolini, M.; Clark-Lewis, I. Interleukin-8, a chemotactic and inflammatory cytokine. *FEBS Lett.* **1992**, *307*, 97–101.

(3) Murphy, P. Neutrophil receptors for interleukin-8 and related CXC chemokines. *Semin. Hematol.* **1997**, *34*, 311–318.

(4) Hoch, R. C.; Schraufstatter, I. U.; Cochrane, C. G. In vivo, in vitro, and molecular aspects of interleukin-8 and the interleukin-8 receptors. *J. Lab. Clin. Med.* **1996**, *128*, 134–145.

(5) Baggiolini, M.; Walz, A.; Kunkel, S. L. Neutrophil-activating peptide-1/interleukin 8, a novel cytokine that activates neutrophils *J. Clin. Invest.* **1989**, *84*, 1045–1049.

(6) Norgauer, J.; Krutmann, J.; Dobos, G. J.; Traynor-Kaplan, A. E.; Oades, Z. G.; Schraufstatter, I. U. Actin polymerization, calcium-transients, and phospholipid metabolism in human neutrophils after stimulation with interleukin-8 and N-formyl peptide. *J. Invest. Dermatol.* **1994**, *102*, 310–314.

(7) Middleton, J.; Patterson, A. M.; Gardner, L.; Schmutz, C.; Ashton, B. A. Leukocyte extravasation: chemokine transport and presentation by the endothelium. *Blood* **2002**, *100*, 3853–3860.

(8) Sekido, N.; Mukaida, N.; Harada, A.; Nakanishi, I.; Watanabe, Y.; Matsushima, K. Prevention of lung reperfusion injury in rabbits by a monoclonal antibody against interleukin-8. *Nature* **1993**, *365*, 654–657.

(9) Bizzarri, C.; Allegretti, M.; Bitondo, R. D.; Cervellera, M. N.; Colotta, F.; Bertini, R. Pharmacological inhibition of interleukin-8 (CXCL8) as a new approach for the prevention and treatment of several human diseases. *Curr. Med. Chem. Anti. Inflamm. Anti. Allergy Agents* **2003**, *2*, 67–79.

(10) Duan, H.; Koga, T.; Kohda, F.; Hara, H.; Urabe, K.; Furue, M. Interleukin-8-positive neutrophils in psoriasis. *J. Dermatol. Sci.* **2001**, *26*, 119–124.

(11) Chollet-Martin, S.; Montravers, P.; Gibert, C.; Elbim, C.; Desmonts, J. M.; Fagon, J. Y.; Gougerot-Pocidalo, M. A. High levels of interleukin-8 in the blood and alveolar spaces of patients with pneumonia and adult respiratory distress syndrome. *Infect. Immun.* **1993**, *61*, 4553–4559.

(12) Carré, P. C.; Mortenson, R. L.; King, T. E.; Noble, P. W.; Sable, C. L.; Riches, D. W. Increased expression of the interleukin-8 gene by alveolar macrophages in idiopathic pulmonary fibrosis. A potential mechanism for the recruitment and activation of neutrophils in lung fibrosis. *J. Clin. Invest.* **1991**, *88*, 1802–1810.

(13) Miller, E. J.; Cohen, A. B.; Nagao, S.; Griffith, D.; Maunder, R. J.; Martin, T. R.; Weiner-Kronish, J. P.; Sticherling, M.; Christophers, E.; Matthay, M. A. Elevated levels of NAP-1/interleukin-8 are present in the airspaces of patients with the adult respiratory distress syndrome and are associated with increased mortality. *Am. Rev. Respir. Dis.* **1992**, *146*, 427–432.

(14) Lynch, J. P., 3rd; Standiford, T. J.; Rolfe, M. W.; Kunkel, S. L.; Strieter, R. M. Neutrophilic alveolitis in idiopathic pulmonary fibrosis. The role of interleukin-8. *Am. Rev. Respir. Dis.* **1992**, *145*, 1433–1439.

(15) Strieter, R. M.; Polverini, P. J.; Arenberg, D. A.; Walz, A.; Opdenakker, G.; Vandamme, J.; Kunkel, S. L. Role of C-X-C chemokines as regulators of angiogenesis in lung-cancer. *J. Leukoc. Biol.* **1995**, *57*, 752–762.

(16) Koch, A. E.; Polverini, P. J.; Kunkel, S. L.; Harlow, L. A.; Dipietro, L. A.; Elner, V. M.; Elner, S. G.; Strieter, R. M. Interleukin-8 as a macrophage-derived mediator of angiogenesis. *Science* **1992**, *258*, 1798–1801.

(17) Zlotnik, A.; Yoshie, O. Chemokines: A new classification system and their role in immunity. *Immunity* **2000**, *12*, 121–127.
(18) Baggiolini, M.; Dewald, B.; Moser, B. Interleukin-8 and related chemotactic cytokines--CXC and CC chemokines. *Adv. Immunol.* **1994**, *55*, 97–179.
(19) Fernandez, E. J.; Lolis, E. Structure, function, and inhibition of chemokines. *Annu. Rev. Pharmacol. Toxicol.* **2002**, *42*, 469–499.
(20) Rajarathnam, K.; Sykes, B. D.; Dewald, B.; Baggiolini, M.; Clark-Lewis, I. Disulfide bridges in interleukin-8 probed using non-natural disulfide analogues: Dissociation of roles in structure from function. *Biochemistry* **1999**, *38*, 7653–7658.
(21) Clark-Lewis, I.; Dewald, B.; Loetscher, M.; Moser, B.; Baggiolini, M. Structural requirements for interleukin-8 function identified by design of analogs and CXC chemokine hybrids. *J. Biol. Chem.* **1994**, *269*, 16075–16081.
(22) Webb, L. M.; Ehrengruber, M. U.; Clark-Lewis, I.; Baggiolini, M.; Rot, A. Binding to heparan sulfate or heparin enhances neutrophil responses to interleukin 8. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7158–7162.
(23) Hammond, M. E.; Shyamala, V.; Siani, M. A.; Gallegos, C. A.; Feucht, P. H.; Abbott, J.; Lapointe, G. R.; Moghadam, M.; Khoja, H.; Zakel, J. Receptor recognition and specificity of interleukin-8 is determined by residues that cluster near a surface-accessible hydrophobic pocket. *J. Biol. Chem.* **1996**, *271*, 8228–8235.
(24) David, R.; Günther, R.; Baumann, L.; Lühmann, T.; Seebach, D.; Hofmann, H.-J. r.; Beck-Sickinger, A. G. Artificial chemokines: Combining chemistry and molecular biology for the elucidation of interleukin-8 functionality. *J. Am. Chem. Soc.* **2008**, *130*, 15311–15317.
(25) Skelton, N. J.; Quan, C.; Reilly, D.; Lowman, H. Structure of a CXC chemokine-receptor fragment in complex with interleukin-8. *Structure* **1999**, *7*, 157–168.
(26) Clore, G. M.; Appella, E.; Yamada, M.; Matsushima, K.; Gronenborn, A. M. Three-dimensional structure of interleukin 8 in solution. *Biochemistry* **1990**, *29*, 1689–1696.
(27) Rajarathnam, K.; Clark-Lewis, I.; Sykes, B. D. 1H NMR solution structure of an active monomeric interleukin-8. *Biochemistry* **1995**, *34*, 12983–12990.
(28) Baldwin, E. T.; Weber, I. T.; St Charles, R.; Xuan, J. C.; Appella, E.; Yamada, M.; Matsushima, K.; Edwards, B. F.; Clore, G. M.; Gronenborn, A. M. Crystal structure of interleukin 8: symbiosis of NMR and crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 502–506.
(29) Baldwin, E. T.; Franklin, K. A.; Appella, E.; Yamada, M.; Matsushima, K.; Wlodawer, A.; Weber, I. T. Crystallization of human interleukin-8. A protein chemotactic for neutrophils and T-lymphocytes. *J. Biol. Chem.* **1990**, *265*, 6851–6853.
(30) Gerber, N.; Lowman, H.; Artis, D. R.; Eigenbrot, C. Receptor-binding conformation of the "ELR" motif of IL-8: X-ray structure of the L5C/H33C variant at 2.35 Å resolution. *Proteins* **2000**, *38*, 361–367.
(31) Eigenbrot, C.; Lowman, H. B.; Chee, L.; Artis, D. R. Structural change and receptor binding in a chemokine mutant with a rearranged disulfide: X-ray structure of e38C/C50A IL-8 at 2 Å resolution. *Proteins* **1997**, *27*, 556–566.
(32) Das, S. T.; Rajagopalan, L.; Guerrero-Plata, A.; Sai, J.; Richmond, A.; Garofalo, R. P.; Rajarathnam, K. Monomeric and dimeric CXCL8 are both essential for in vivo neutrophil recruitment. *PLoS ONE* **2010**, *5*, e11754.
(33) Leong, S. R.; Lowman, H. B.; Liu, J.; Shire, S.; Deforge, L. E. Gillece-Castro, B. L.; McDowell, R.; Hebert, C. A. IL-8 single-chain homodimers and heterodimers: interactions with chemokine receptors CXCR1, CXCR2, and DARC. *Protein Sci.* **1997**, *6*, 609–617.
(34) Rajarathnam, K.; Kay, C. M.; Clark-Lewis, I.Sykes, B. D. Characterization of quaternary structure of interleukin-8 and functional implications. *Methods Enzymol.* **1997**, *287*, 89–105.
(35) Lowman, H. B.; Fairbrother, W. J.; Slagle, P. H.; Kabakoff, R. Liu, J.; Shire, S.; Hebert, C. A. Monomeric variants of IL-8: Effects of side chain substitutions and solution conditions upon dimer formation. *Protein Sci.* **1997**, *6*, 598–608.
(36) Burrows, S. D.; Doyle, M. L.; Murphy, K. P.; Franklin, S. G.; White, J. R.; Brooks, I.; McNulty, D. E.; Scott, M. O.; Knutson, J. R. Determination of the monomer-dimer equilibrium of interleukin-8 reveals it is a monomer at physiological concentrations. *Biochemistry* **1994**, *33*, 12741–12745.
(37) Paolini, J. F.; Willard, D.; Consler, T.; Luther, M.; Krangel, M. S. The chemokines IL-8, monocyte chemoattractant protein-1, and I-309 are monomers at physiologically relevant concentrations. *J. Immunol.* **1994**, *153*, 2704–2717.
(38) Henry, B. L.; Wayne, J. F.; Paul, H. S.; Rhona, K.; Caroline, A. H.; Jun, L.; Steven, S. Monomeric variants of IL-8: Effects of side chain substitutions and solution conditions upon dimer formation. *Protein Sci.* **1997**, *6*, 598–608.
(39) Goger, B.; Halden, Y.; Rek, A.; Mosl, R.; Pye, D.; Gallagher, J.; Kungl, A. J. Different affinities of glycosaminoglycan oligosaccharides for monomeric and dimeric interleukin-8: A model for chemokine regulation at inflammatory sites. *Biochemistry* **2002**, *41*, 1640–1646.
(40) Nesmelova, I. V.; Sham, Y.; Dudek, A. Z.; van Eijk, L. I.; Wu, G.; Slungaard, A.; Mortari, F.; Griffioen, A. W.; Mayo, K. H. Platelet factor 4 and interleukin-8 CXC chemokine heterodimer formation modulates function at the quaternary structural level. *J. Biol. Chem.* **2005**, *280*, 4948–4958.
(41) Cornell, W.; Abseher, R.; Nilges, M.; Case, D. A. Continuum solvent molecular dynamics study of flexibility in interleukin-8. *J. Mol. Graph. Model.* **2001**, *19*, 136–145.
(42) Chuntharapai, A.; Kim, K. J. Regulation of the expression of IL-8 receptor A/B by IL-8: possible functions of each receptor. *J. Immunol.* **1995**, *155*, 2587–2594.
(43) Lee, J.; Horuk, R.; Rice, G. C.; Bennett, G. L.; Camerato, T.; Wood, W. I. Characterization of two high affinity human interleukin-8 receptors. *J. Biol. Chem.* **1992**, *267*, 16283–16287.
(44) Murphy, P. M.; Tiffany, H. L. Cloning of complementary DNA encoding a functional human interleukin-8 receptor. *Science* **1991**, *253*, 1280–1283.
(45) Holmes, W. E.; Lee, J.; Kuang, W. J.; Rice, G. C.; Wood, W. I. Structure and functional expression of a human interleukin-8 receptor. *Science* **1991**, *253*, 1278–1280.
(46) Ahuja, S. K.; Murphy, P. M. The CXC chemokines growth-regulated oncogene (GRO) alpha, GRObeta, GROgamma, neutrophil-activating peptide-2, and epithelial cell-derived neutrophil-activating peptide-78 are potent agonists for the type B, but not the type A, human interleukin-8 receptor. *J. Biol. Chem.* **1996**, *271*, 20545–20550.
(47) Zhao-hai, L.; Zi-xuan, W.; Horuk, R.; Hesselgesser, J.; Yan-chun, L.; Hadley, T. J.; Peiper, S. C. The promiscuous chemokine binding profile of the duffy antigen/receptor for chemokines is primarily localized to sequences in the amino-terminal domain. *J. Biol. Chem.* **1995**, *270*, 26239–26245.
(48) Hébert, C. A.; Vitangcol, R. V.; Baker, J. B. Scanning mutagenesis of interleukin-8 identifies a cluster of residues required for receptor binding. *J. Biol. Chem.* **1991**, *266*, 18989–18994.
(49) LaRosa, G. J.; Thomas, K. M.; Kaufmann, M. E.; Mark, R.; White, M.; Taylor, L.; Gray, G.; Witt, D.; Navarro, J. Amino terminus of the interleukin-8 receptor is a major determinant of receptor subtype specificity. *J. Biol. Chem.* **1992**, *267*, 25402–25406.
(50) Rajagopalan, L.; Rajarathnam, K. Ligand selectivity and affinity of chemokine receptor CXCR1: Role of N-terminal domain. *J. Biol. Chem.* **2004**, *279*, 30000–30008.
(51) Gayle, R. B. d.; Sleath, P. R.; Srinivason, S.; Birks, C. W.; Weerawarna, K. S.; Cerretti, D. P.; Kozlosky, C. J.; Nelson, N.; Vanden Bos, T.; Beckmann, M. P. Importance of the amino terminus of the interleukin-8 receptor in ligand interactions. *J. Biol. Chem.* **1993**, *268*, 7283–7289.
(52) Attwood, M. R.; Conway, E. A.; Dunsdon, R. M.; Greening, J. R.; Handa, B. K.; Jones, P. S.; Jordan, S. C.; Keech, E.; Wilson, F. X. Peptide based inhibitors of interleukin-8: structural simplification and enhanced potency. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 429–432.
(53) Leong, S. R.; Kabakoff, R. C.; Hebert, C. A. Complete mutagenesis of the extracellular domain of interleukin-8 (IL- 8) type A receptor identifies charged residues mediating IL-8 binding and signal transduction. *J. Biol. Chem.* **1994**, *269*, 19343–19348.
(54) Clubb, R. T.; Omichinski, J. G.; Clore, G. M.; Gronenborn, A. M. Mapping the binding surface of interleukin-8 complexed with an

N-terminal fragment of the Type 1 human interleukin-8 receptor. *FEBS Lett.* **1994**, *338*, 93–97.

(55) Ravindran, A.; Joseph, P. R. B.; Rajarathnam, K. Structural basis for differential binding of the Interleukin-8 monomer and dimer to the CXCR1 N-Domain: Role of coupled interactions and dynamics. *Biochemistry* **2009**, *48*, 8795–8805.

(56) Luo, Z.; Butcher, D. J.; Huang, Z. Molecular modeling of interleukin-8 receptor beta and analysis of the receptor-ligand interaction. *Protein Eng.* **1997**, *10*, 1039–1045.

(57) Kuschert, G. S.; Hoogewerf, A. J.; Proudfoot, A. E.; Chung, C. W.; Cooke, R. M.; Hubbard, R. E.; Wells, T. N.; Sanderson, P. N. Identification of a glycosaminoglycan binding surface on human interleukin-8. *Biochemistry* **1998**, *37*, 11193–11201.

(58) Kuschert, G. S.; Coulin, F.; Power, C. A.; Proudfoot, A. E.; Hubbard, R. E.; Hoogewerf, A. J.; Wells, T. N. Glycosaminoglycans interact selectively with chemokines and modulate receptor binding and cellular responses. *Biochemistry* **1999**, *38*, 12959–12968.

(59) Spillmann, D.; Witt, D.; Lindahl, U. Defining the interleukin-8-binding domain of heparan sulfate. *J. Biol. Chem.* **1998**, *273*, 15487–15493.

(60) Steven, R. L.; Caroline, A. H.; Henry, B. L.; Jun, L.; Steven, S.; Laura, E. D.; Beth, L. G.-C.; Robert, M. IL-8 single-chain homodimers and heterodimers: Interactions with the chemokine receptors CXCR1, CXCR2, and DARC. *Protein Sci.* **1997**, *6*, 609–617.

(61) Rajarathnam, K.; Prado, G. N.; Fernando, H.; Clark-Lewis, I.; Navarro, J. Probing receptor binding activity of interleukin-8 dimer using a disulfide trap. *Biochemistry* **2006**, *45*, 7882–7888.

(62) Fernando, H.; Chin, C.; Rosgen, J.; Rajarathnam, K. Dimer dissociation is essential for interleukin-8 (IL-8) binding to CXCR1 receptor. *J. Biol. Chem.* **2004**, *279*, 36175–36178.

(63) Rajarathnam, K.; Sykes, B. D.; Kay, C. M.; Dewald, B.; Geiser, T.; Baggiolini, M.; Clark-Lewis, I. Neutrophil activation by monomeric interleukin-8. *Science* **1994**, *264*, 90–92.

(64) Krieger, E.; Geretti, E.; Brandner, B.; Goger, B.; Wells, T. N.; Kungl, A. J. A structural and dynamic model for the interaction of interleukin-8 and glycosaminoglycans: support from isothermal fluorescence titrations. *Proteins* **2004**, *54*, 768–775.

(65) Lortat-Jacob, H.; Grosdidier, A.; Imberty, A. Structural diversity of heparan sulfate binding domains in chemokines. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1229–1234.

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

(67) de Paz, J. L.; Moseman, E. A.; Noti, C.; Polito, L.; von Andrian, U. H.; Seeberger, P. H. Profiling heparin-chemokine interactions using synthetic tools. *ACS Chem. Biol.* **2007**, *2*, 735–744.

(68) Hay, D. W. P.; Sarau, H. M. Interleukin-8 receptor antagonists in pulmonary diseases. *Curr. Opin. Pharmacol.* **2001**, *1*, 242–247.

(69) Li, F.; Gordon, J. R. IL-8(3−73)K11R is a high affinity agonist of the neutrophil CXCR1 and CXCR2. *Biochem. Biophys. Res. Commun.* **2001**, *286*, 595–600.

(70) Moser, B.; Dewald, B.; Barella, L.; Schumacher, C.; Baggiolini, M.; Clark-Lewis, I. Interleukin-8 antagonists generated by N-terminal modification. *J. Biol. Chem.* **1993**, *268*, 7125–7128.

(71) Li, F.; Zhang, X.; Gordon, J. R. CXCL8(3−73)K11R/G31P antagonizes ligand binding to the neutrophil CXCR1 and CXCR2 receptors and cellular responses to CXCL8/IL-8. *Biochem. Biophys. Res. Commun.* **2002**, *293*, 939–944.

(72) Attwood, M. R.; Borkakoti, N.; Bottomley, G. A.; Conway, E. A.; Cowan, I.; Fallowfield, A. G.; Handa, B. K.; Jones, P. S.; Keech, E.; Kirtland, S. J.; Williams, G.; Wilson, F. X. Identification and characterisation of an inhibitor of interleukin-8: a receptor based approach. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1869–1874.

(73) Busch-Petersen, J. Small molecule antagonists of the CXCR2 and CXCR1 chemokine receptors as therapeutic agents for the treatment of inflammatory diseases. *Curr. Top. Med. Chem.* **2006**, *6*, 1345–1352.

(74) Bedke, J.; Nelson, P. J.; Kiss, E.; Muenchmeier, N.; Rek, A.; Behnes, C.-L.; Gretz, N.; Kungl, A. J.; Gröne, H.-J. A novel CXCL8 protein-based antagonist in acute experimental renal allograft damage. *Mol. Immunol.* **2010**, *47*, 1047–1057.

(75) Rek, A.; Potzinger, H.; Wabitsch, V.; Geretti, E.; Piccinini, M. A.; Falsone, S. F.; Grosse Kracht, S.; Teixeira, M.; Adage, T.; Kungl, A. J. Turning the glycan switch to change CXCL8 into a potent anti-inflammatory biological drug. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, submitted for publication.

(76) Freeman, C.; Liu, L.; Banwell, M. G.; Brown, K. J.; Bezos, A.; Ferro, V.; Parish, C. R. Use of sulfated linked cyclitols as heparan sulfate mimetics to probe the heparin/heparan sulfate binding specificity of proteins. *J. Biol. Chem.* **2005**, *280*, 8842–8849.

(77) Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402.

(78) Boeckmann, B.; Bairoch, A.; Apweiler, R.; Blatter, M.-C.; Estreicher, A.; Gasteiger, E.; Martin, M. J.; Michoud, K.; O'Donovan, C.; Phan, I.; Pilbout, S.; Schneider, M. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **2003**, *31*, 365–370.

(79) Landau, M.; Mayrose, I.; Rosenberg, Y.; Glaser, F.; Martz, E.; Pupko, T.; Ben-Tal, N. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.* **2005**, *33*, W299–W302.

(80) Rambaut, A. *FigTree, a graphical viewer of phylogenetic trees*, 1.2.2; Institute of Evolutionary Biology: Edinburgh, 2006.

(81) Suhre, K.; Sanejouand, Y.-H. ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Res.* **2004**, *32*, W610–W614.

(82) Maple, J. R. Derivation of class II force fields: V. Quantum force field for amides, peptides, and related compounds. *J. Comput. Chem.* **1998**, *19*, 430–458.

(83) Dealwis, C.; Fernandez, E. J.; Thompson, D. A.; Simon, R. J.; Siani, M. A.; Lolis, E. Crystal structure of chemically synthesized [N33A] stromal cell-derived factor 1α, a potent ligand for the HIV-1 "fusin" coreceptor. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6941–6946.

(84) Lubkowski, J.; Bujacz, G.; Boqué, L.; Domaille, P. J.; Handel, T. M.; Alexander, W. The structure of MCP-1 in two crystal forms provides a rare example of variable quaternary interactions. *Nat. Struct. Biol.* **1997**, *4*, 64–69.

(85) Hoover, D. M.; Shaw, J.; Gryczynski, Z.; Proudfoot, A. E. I.; Wells, T.; Lubkowski, J. The crystal structure of Met-Rantes: comparison with native Rantes and AOP-Rantes. *Protein Pept. Lett.* **2000**, *7*, 73–82.

(86) Gandhi, N. S.; Coombe, D. R.; Mancera, R. L. Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 1. Molecular modeling studies. *Biochemistry* **2008**, *47*, 4851–4862.

(87) Krissinel, E.; Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst. Sect. D* **2004**, *60*, 2256–2268.

(88) Krissinel, E.; Henrick, K. Protein structure comparison service SSM at European Bioinformatics Institute. http://www.ebi.ac.uk/msd-srv/ssm (November 15).

(89) Showalter, S. A.; Brüschweiler, R. Quantitative molecular ensemble interpretation of NMR dipolar couplings without restraints. *J. Am. Chem. Soc.* **2007**, *129*, 4158–4159.

(90) *OMEGA C++ Toolkit*, 2.3.2; OpenEye Scientific Software, Inc.: Sante Fe, NM, 2008.

(91) Thomas, A. H. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490–519.

(92) Thomas, A. H. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comput. Chem.* **1996**, *17*, 520–552.

(93) Thomas, A. H. Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. *J. Comput. Chem.* **1996**, *17*, 553–586.

(94) Thomas, A. H.; Robert, B. N. Merck molecular force field. IV. conformational energies and geometries for MMFF94. *J. Comput. Chem.* **1996**, *17*, 587–615.
(95) Thomas, A. H. Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. *J. Comput. Chem.* **1996**, *17*, 616–641.
(96) Ferro, D. R.; Provasoli, A.; Ragazzi, M.; Casu, B.; Torri, G.; Bossennec, V.; Perly, B.; Sinay, P.; Petitou, M.; Choay, J. Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences. *Carbohydr. Res.* **1990**, *195*, 157–167.
(97) *FRED (Fast Rigid Exhaustive Docking)*, 2.2.4; Openeye Scientific Software: Santa Fe, NM, 2008.
(98) Esther, K.; Jordi, R.; Pascal, M.; Didier, R. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **2004**, *57*, 225–242.
(99) Esko, J. D.; Linhardt, R. J., Proteins that bind sulfated glycosaminoglycans. In *Essentials of glycobiology*, II ed.; Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., Etzler, M. E., Eds.; CSHL Press: New York, 2009.
(100) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* **1997**, *11*, 425–445.
(101) Gehlhaar, D. K.; Verkhivker, G. M.; Rejto, P. A.; Sherman, C. J.; Fogel, D. R.; Fogel, L. J.; Freer, S. T. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem. Biol.* **1995**, *2*, 317–324.
(102) Stahl, M.; Rarey, M. Detailed analysis of scoring functions for virtual screening. *J. Med. Chem.* **2001**, *44*, 1035–1042.
(103) Mark, R. M.; Harold, R. A.; Anthony, N.; Grant, J. A.; Frank, K. B. Gaussian docking functions. *Biopolymers* **2003**, *68*, 76–90.
(104) Garrett, M. M.; David, S. G.; Robert, S. H.; Ruth, H.; William, E. H.; Richard, K. B.; Arthur, J. O. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
(105) Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. A semiempirical free energy force field with charge-based desolvation *J. Comput. Chem.* **2007**, *28*, 1145–1152.
(106) Alonso, H.; Bliznyuk, A. A.; Gready, J. E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* **2006**, *26*, 531–568.
(107) Viktor, H.; Robert, A.; Asim, O.; Bentley, S.; Adrian, R.; Carlos, S. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712–725.
(108) Case, D. A.; Cheatham Iii, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
(109) Wang, J.; Wolf, R.; Caldwell, J.; Kollman, P.; Case, D. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–74.
(110) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* **2008**, *29*, 622–655.
(111) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
(112) Pigache, A.; Cieplak, P.; Dupradeau, F. Y., Automatic and highly reproducible RESP and ESP charge derivation: application to the development of programs RED and X RED. In *227th ACS National Meeting*, Anaheim, CA, March 28−April 1, 2004.
(113) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, J. T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision C. 02; Gaussian, Inc.: Wallingford, CT, 2004.
(114) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
(115) Nesmelova, I. V.; Sham, Y.; Gao, J.; Mayo, K. H. CXC and CC chemokines form mixed heterodimers: Association free energies from molecular dynamics simulations and experimental coorelations. *J. Biol. Chem.* **2008**, *283*, 24155–24166.
(116) Tsui, V.; Case, D. A. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* **2001**, *56*, 257–291.
(117) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.
(118) Sanner, M. F.; Olson, A. J.; Spehner, J.-C. Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers* **1996**, *38*, 305–320.
(119) Nguyen, D. T.; Case, D. A. On finding stationary states on large-molecule potential energy surfaces. *J. Phys. Chem.* **1985**, *89*, 4020–4026.
(120) Eric, F. P.; Thomas, D. G.; Conrad, C. H.; Gregory, S. C.; Daniel, M. G.; Elaine, C. M.; Thomas, E. F. UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.
(121) Chen, M. E.; Cang, H. X.; Nymeyer, H. *NOC*, 3.0.; Institute of Molecular Biology, Florida State University: FL, 2007.
(122) Kanagarajadurai, K.; Sowdhamini, R. Sequence and structural analyses of interleukin-8-like chemokine superfamily. *In Silico Biol.* **2008**, *8*, 307–330.
(123) Kagiampakis, I.; Jin, H.; Kim, S.; Vannucci, M.; LiWang, P. J.; Tsai, J. Conservation of unfavorable sequence motifs that contribute to the chemokine quaternary state. *Biochemistry* **2008**, *47*, 10637–10648.
(124) Laguri, C.; Arenzana-Seisdedos, F.; Lortat-Jacob, H. Relationships between glycosaminoglycan and receptor binding sites in chemokines--the CXCL12 example. *Carbohydr. Res.* **2008**, *343*, 2018–2023.
(125) Brändén, C.-I.; Tooze, J. *Introduction to protein structure*, II ed.; Garland Pub.: New York, 1999; p 410.
(126) Goldenberg, O.; Erez, E.; Nimrod, G.; Ben-Tal, N. The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures. *Nucleic Acids Res.* **2009**, *37*, D323–D327.
(127) Inoue, Y.; Endo, M.; Haruta, C.; Taniuchi, T.; Moritomo, T.; Nakanishi, T. Molecular cloning and sequencing of the silver chimaera (Chimaera phantasma) interleukin-8 cDNA. *Fish Shellfish Immunol.* **2003**, *15*, 269–274.
(128) Lee, E.-Y.; Park, H.-H.; Kim, Y.-T.; Chung, J.-K.; Choi, T.-J. Cloning and sequence analysis of the interleukin-8 gene from flounder (Paralichthys olivaceous). *Gene* **2001**, *274*, 237–243.
(129) Covello, J. M.; Bird, S.; Morrison, R. N.; Battaglene, S. C.; Secombes, C. J.; Nowak, B. F. Cloning and expression analysis of three striped trumpeter (Latris lineata) pro-inflammatory cytokines, TNF-α, IL-1β and IL-8, in response to infection by the ectoparasitic, Chondracanthus goldsmidi. *Fish Shellfish Immunol.* **2009**, *26*, 773–786.
(130) Sepulcre, M. P.; Sarropoulou, E.; Kotoulas, G.; Meseguer, J.; Mulero, V. Vibrio anguillarum evades the immune response of the bony fish sea bass (Dicentrarchus labrax L.) through the inhibition of leukocyte respiratory burst and down-regulation of apoptotic caspases. *Mol. Immunol.* **2007**, *44*, 3751–3757.

(131) Wu, Y. F.; Shien, J. H.; Yin, H. H.; Chiow, S. H.; Lee, L. H. Structural and functional homology among chicken, duck, goose, turkey and pigeon interleukin-8 proteins. *Vet. Immunol. Immunopathol.* **2008**, *125*, 205–215.

(132) Liu, Y.; Yang, S.; Lin, A.; Cavalli-Sforza, L.; Su, B. Molecular evolution of CXCR1, a G protein-coupled receptor involved in signal transduction of neutrophils. *J. Mol. Evol.* **2005**, *61*, 691–696.

(133) Shields, D. C. Gene conversion among chemokine receptors. *Gene* **2000**, *246*, 239–245.

(134) Zlotnik, A.; Yoshie, O.; Nomiyama, H. The chemokine and chemokine receptor superfamilies and their molecular evolution. *Genome Biol.* **2006**, *7*, 243.

(135) DeVries, M. E.; Kelvin, A. A.; Xu, L.; Ran, L.; Robinson, J.; Kelvin, D. J. Defining the origins and evolution of the Chemokine/Chemokine receptor system. *J. Immunol.* **2006**, *176*, 401–415.

(136) Suzuki, H.; Prado, G. N.; Wilkinson, N.; Navarro, J. The N terminus of interleukin-8 (IL-8) receptor confers high affinity binding to human IL-8. *J. Biol. Chem.* **1994**, *269*, 18263–18266.

(137) Reeves, E. P.; Williamson, M.; Byrne, B.; Bergin, D. A.; Smith, S. G. J.; Greally, P.; O'Kennedy, R.; O'Neill, S. J.; McElvaney, N. G. IL-8 dictates glycosaminoglycan binding and stability of IL-18 in cystic fibrosis. *J. Immunol.* **2010**, *184*, 1642–1652.

(138) Caldwell, E. E. O.; Nadkarni, V. D.; Fromm, J. R.; Linhardt, R. J.; Weiler, J. M. Importance of specific amino acids in protein binding sites for heparin and heparan sulfate. *Int. J. Biochem. Cell Biol.* **1996**, *28*, 203–216.

(139) Fromm, J. R.; Hileman, R. E.; Caldwell, E. E. O.; Weiler, J. M.; Linhardt, R. J. Pattern and spacing of basic amino acids in heparin binding sites. *Arch. Biochem. Biophys.* **1997**, *343*, 92–100.

(140) Cardin, A. D.; Weintraub, H. J. Molecular modeling of protein-glycosaminoglycan interactions. *Arterioscler. Thromb. Vasc. Biol.* **1989**, *9*, 21–32.

(141) Ronald, E. H.; Jonathan, R. F.; John, M. W.; Robert, J. L. Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. *BioEssays* **1998**, *20*, 156–167.

(142) Margalit, H.; Fischer, N.; Ben-Sasson, S. A. Comparative analysis of structurally defined heparin binding sequences reveals a distinct spatial distribution of basic residues. *J. Biol. Chem.* **1993**, *268*, 19228–19231.

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

(144) Williams, G.; Borkakoti, N.; Bottomley, G. A.; Cowan, I.; Fallowfield, A. G.; Jones, P. S.; Kirtland, S. J.; Price, G. J.; Price, L. Mutagenesis studies of interleukin-8. Identification of a second epitope involved in receptor binding. *J. Biol. Chem.* **1996**, *271*, 9579–9586.

(145) Clark-Lewis, I.; Schumacher, C.; Baggiolini, M.; Moser, B. Structure-activity relationships of interleukin-8 determined using chemically synthesized analogs. Critical role of NH2-terminal residues and evidence for uncoupling of neutrophil chemotaxis, exocytosis, and receptor binding activities. *J. Biol. Chem.* **1991**, *266*, 23128–23134.

(146) Schug, A.; Herges, T.; Wenzel, W. Reproducible protein folding with the stochastic tunneling method. *Phys. Rev. Lett.* **2003**, *91*, 158102.

(147) Herges, T.; Wenzel, W. An all-atom force field for tertiary structure prediction of helical proteins. *Biophys. J.* **2004**, *87*, 3100–3109.

(148) Meliciani, I.; Klenin, K.; Strunk, T.; Schmitz, K.; Wenzel, W. Probing hot spots on protein-protein interfaces with all-atom free-energy simulation. *J. Chem. Phys.* **2009**, *131*, 034114.

(149) Swanson, J. M.; Henchman, R. H.; McCammon, J. A. Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophys. J.* **2004**, *86*, 67–74.

(150) Benedix, A.; Becker, C. M.; de Groot, B. L.; Caflisch, A.; Bockmann, R. A. Predicting free energy changes using structural ensembles. *Nat. Methods* **2009**, *6*, 3–4.

(151) Yap, K. L.; Ames, J. B.; Swindells, M. B.; Ikura, M. Vector geometry mapping. A method to characterize the conformation of helix-loop-helix calcium-binding proteins. *Methods Mol. Biol.* **2002**, *173*, 317–324.

(152) Clore, G. M.; Gronenborn, A. M. Comparison of the solution nuclear magnetic resonance and crystal structures of interleukin-8. Possible implications for the mechanism of receptor binding. *J. Mol. Biol.* **1991**, *217*, 611–620.

(153) Rajarathnam, K.; Clark-Lewis, I.; Dewald, B.; Baggiolini, M.; Sykes, B. D. 1H NMR evidence that Glu-38 interacts with the N-terminal functional domain in interleukin-8. *FEBS Lett.* **1996**, *399*, 43–46.

(154) Rajarathnam, K.; Clark-Lewis, I.; Sykes, B. D. 1H NMR studies of interleukin 8 analogs: characterization of the domains essential for function. *Biochemistry* **1994**, *33*, 6623–6630.

(155) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(156) Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, *21*, 1049–1074.

(157) García, A. E.; Sanbonmatsu, K. Y. α-Helical stabilization by side chain shielding of backbone hydrogen bonds. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2782–2787.

(158) Ono, S.; Nakajima, N.; Higo, J.; Nakamura, H. Peptide free-energy profile is strongly dependent on the force field: Comparison of C96 and AMBER95. *J. Comput. Chem.* **2000**, *21*, 748–762.

(159) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950–1958.

(160) Wickstrom, L.; Okur, A.; Simmerling, C. Evaluating the performance of the ff99SB force field based on NMR scalar coupling data. *Biophys. J.* **2009**, *97*, 853–856.

(161) Best, R. B.; Buchete, N.-V.; Hummer, G. Are current molecular dynamics force fields too helical?. *Biophys. J.* **2008**, *95* L07–L09.

(162) Showalter, S. A.; Johnson, E.; Rance, M.; Brüschweiler, R. Toward quantitative interpretation of methyl side-chain dynamics from NMR by molecular dynamics simulations. *J. Am. Chem. Soc.* **2007**, *129*, 14146–14147.

(163) Wickstrom, L.; Okur, A.; Song, K.; Hornak, V.; Raleigh, D. P.; Simmerling, C. L. The unfolded state of the villin headpiece helical subdomain: Computational studies of the role of locally stabilized structure. *J. Mol. Biol.* **2006**, *360*, 1094–1107.

(164) Wolfgang, K.; Christian, S. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637.

(165) Grasberger, B. L.; Gronenborn, A. M.; Clore, G. M. Analysis of the backbone dynamics of interleukin-8 by 15N relaxation measurements. *J. Mol. Biol.* **1993**, *230*, 364–372.

(166) Woods, R. J.; Tessier, M. B. Computational glycoscience: characterizing the spatial and temporal properties of glycans and glycan-protein complexes. *Curr. Opin. Struct. Biol.* **2010**, *20*, 575–83.

(167) Gandhi, N. S.; Mancera, R. L. Free energy calculations of glycosaminoglycan-protein interactions. *Glycobiology* **2009**, *19*, 1103–1115.

(168) Kerzmann, A.; Fuhrmann, J.; Kohlbacher, O.; Neumann, D. BALLDock/SLICK: a new method for protein-carbohydrate docking. *J. Chem. Inf. Model.* **2008**, *48*, 1616–1625.

(169) Kerzmann, A.; Neumann, D.; Kohlbacher, O. SLICK--scoring and energy functions for protein-carbohydrate interactions. *J. Chem. Inf. Model.* **2006**, *46*, 1635–1642.

(170) Laederach, A.; Reilly, P. J. Specific empirical free energy function for automated docking of carbohydrates to proteins. *J. Comput. Chem.* **2003**, *24*, 1748–1757.

(171) Rao, V. S. R.; Qasba, P. K.; Balaji, P. V.; Chandrasekaran, R. *Conformation of carbohydrates*; Harwood Academic Pub: Amsterdam, 1998; p 409.

(172) Fadda, E.; Woods, R. J. Molecular simulations of carbohydrates and protein-carbohydrate interactions: motivation, issues and prospects. *Drug Discov. Today* **2010**, *15*, 596–609.

(173) Day, R.; Paschek, D.; Garcia, A. E. Microsecond simulations of the folding/unfolding thermodynamics of the Trp-cage miniprotein. *Proteins* **2010**, *78*, 1889–1899.

(174) Hess, B.; van der Vegt, N. F. Hydration thermodynamic properties of amino acid analogues: a systematic comparison of biomolecular force fields and water models. *J. Phys. Chem. B* **2006**, *110*, 17616–17626.

(175) Murphy, J. W.; Cho, Y.; Sachpatzidis, A.; Fan, C.; Hodsdon, M. E.; Lolis, E. Structural and Functional Basis of CXCL12 (Stromal Cell-derived Factor-1{alpha}) Binding to Heparin. *J. Biol. Chem.* **2007**, *282*, 10018–10027.

(176) Ohnishi, Y.; Senda, T.; Nandhagopal, N.; Sugimoto, K.; Shioda, T.; Nagai, Y.; Mitsui, Y. Crystal Structure of Recombinant Native SDF-1α with Additional Mutagenesis Studies: An Attempt at a More Comprehensive Interpretation of Accumulated Structure-Activity Relationship Data. *J. Interferon Cytokine Res.* **2000**, *20*, 691–700.

(177) Malik, Z. A.; Tack, B. F. Structure of human MIP-3[alpha] chemokine. *Acta Cryst. Sect. F* **2006**, *62*, 631–634.

(178) Barinka, C.; Prahl, A.; Lubkowski, J. Structure of human monocyte chemoattractant protein 4 (MCP-4/CCL13). *Acta Cryst. Sect. D* **2008**, *64*, 273–278.

(179) Wilken, J.; Hoover, D.; Thompson, D. A.; Barlow, P. N.; McSparron, H.; Picard, L.; Wlodawer, A.; Lubkowski, J.; Kent, S. B. H. Total chemical synthesis and high-resolution crystal structure of the potent anti-HIV protein AOP-RANTES. *Chem. Biol.* **1999**, *6*, 43–51.

(180) Shaw, J. P.; Johnson, Z.; Borlat, F.; Zwahlen, C.; Kungl, A.; Roulin, K.; Harrenga, A.; Wells, T. N. C.; Proudfoot, A. E. I. The X-Ray Structure of RANTES: Heparin-Derived Disaccharides Allows the Rational Design of Chemokine Inhibitors. *Structure* **2004**, *12*, 2081–2093.

J Chem Inf Model, 2011, 51, 335-358

Table S1. Accession numbers for CXCL-8 from various species and percentage identity with respect to human CXCL-8.

Accession number	Species	% identity
P10145	*Homo sapiens* (Human)	-
P67814	*Macaca nemestrina* (Pig-tailed macaque)	95
P67813	*Macaca mulatta* (Rhesus macaque)	95
P46653	*Cercocebus torquatus atys* (Red-crowned mangabey)	94
P10145-2	*Homo sapiens* (Human)	100
B6E125	*Cercopithecus sabaeus* (Green monkey)	92
P36925	*Ovis aries* (Sheep)	81
B2BJC6	*Ovis canadensis* (Bighorn sheep)	81
Q0PXQ7	*Cervus elaphus* (Red deer)	80
P19874	*Oryctolagus cuniculus* (Rabbit)	81
B8PZN9	*Marmota monax* (Woodchuck)	80
Q58GK7	*Bubalus bubalis* (Domestic water buffalo)	79
P79255	*Bos taurus* (Bovine)	79
B1PSL8	*Bos indicus* (Zebu)	79
P41324	*Canis familiaris* (Dog)	77
A4PIZ8	*Mustela putorius furo* (European domestic ferret)	76
P26894, Q68CM2	*Sus scrofa* (Pig)	77
Q7YRB5	*Tursiops truncatus* (Atlantic bottle-nosed dolphin)	75
Q9XSX5	*Felis catus* (Cat)	74
O62812	*Equus caballus* (Horse)	72
P49113	*Cavia porcellus* (Guinea pig)	70
Q102R3	*Dasypus novemcinctus* (Nine-banded armadillo)	72
Q5G285	*Bubalus bubalis* (Domestic water buffalo)	79
A3DTN9	*Columba livia* (Domestic pigeon)	53
Q0WZB5	*Anas platyrhynchos* (Domestic duck)	53
P08317	*Gallus gallus* (Chicken)	51
C4PAM2	*Gallus sonneratii* (Gray junglefowl)	51
C4PAK0	*Gallus lafayettei* (Ceylon junglefowl)	51
Q2WG73	*Anser cygnoides* (Chinese goose)	52
A3DTN8	*Anser anser* (domestic goose)	52
C3U5B6	*Trionyx sinensis* (Chinese softshell turtle)	49
A2VBZ2	*Meleagris gallopavo* (Common turkey)	50
Q2ACC2	*Anas platyrhynchos* (Domestic duck)	55
Q8UW91	*Triakis scyllium* (Leopard shark)	40
Q8AXP4	*Chimaera phantasma* (silver chimaera)	31
Q90Y59	*Paralichthys olivaceus* (Japanese flounder)	39
C4NF88	*Latris lineata* (striped trumpeter)	39
A6PZ53	*Dicentrarchus labrax* (European sea bass)	39

J Chem Inf Model, 2011, 51, 335-358

Table S2. MM/PB(GB)SA energy component analysis of the interactions of the CXCL-8 dimer with the receptor peptide averaged over 50 ns[a].

	Complex		Receptor		Ligand		Δ	
	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
ELE	-4441.18	91.11	-3137.09	89.47	254.85	17.15	-1558.95	106.13
VDW	-626.36	24.34	-546.67	20.3	-1.73	6.25	-77.97	11.82
INT	3654.6	39.26	3274.35	37.33	380.25	11.72	0	0
GAS	-1412.94	99.83	-409.4	88.91	633.38	20.75	-1636.91	114.49
PBSUR	54.6	1.35	50.33	0.78	14.55	0.52	-10.28	1.19
PBCAL	-2879.82	94.36	-3513.27	71.33	-932.34	15.13	1565.79	107.64
PBSOL	-2825.22	93.47	-3462.94	71.05	-917.79	15.28	1555.51	106.65
PBELE	-7321	35.81	-6650.36	33.43	-677.49	5.83	6.84	11.52
PBTOT	-4238.16	41.88	-3872.34	40.68	-284.41	12.12	-81.4	11.7
GBSUR	71.31	1.79	65.63	1.03	18.11	0.69	-12.43	1.58
GBCAL	-2874.36	91.07	-3488.48	74.89	-950.55	15.42	1564.67	106.91
GBSOL	-2803.05	89.95	-3422.85	74.54	-932.44	15.61	1552.23	105.6
GBELE	-7315.54	28.6	-6625.56	27.41	-695.7	5.2	5.72	7.97
GBTOT	-4215.98	38.81	-3832.25	36.98	-299.06	12.15	-84.68	10.99
$T\Delta S$[b]							-41.43	8.23
$\Delta G_{binding\text{-}PBSA}$[b]							-39.97	
$\Delta G_{binding\text{-}GBSA}$[b]							-43.25	

a Average over snapshots taken from a 30 ns MD simulation trajectory. b Entropy calculations were based on normal modes analysis using 60 snapshots.

ELE, non-bonded electrostatic energy; VDW, non-bonded van der Waals energy; INT, bond, angle, dihedral energies; GAS, ELE+VDW+INT; PBSUR, hydrophobic contribution to solvation free energy for PB calculations; PBCAL, reaction field energy calculated by PB; PBSOL=PBSUR+PBCAL; PBELE=PBCAL+ELE; PBTOTAL=PBSOL+GAS; GBSUR, hydrophobic contributions to solvation free energy for GB calculations; GBCAL, reaction field energy calculated by GB; GBSOL=GBSUR+GBCAL; GBELE=GBCAL+ELE; GBTOTAL=GBSOL+GAS; TΔS, T(temperature); ΔS (sum of rotational, translational and vibrational entropies); $\Delta G_{binding}$, total binding energy of the system. All energies are in kcal/mol.

J Chem Inf Model, 2011, 51, 335-358

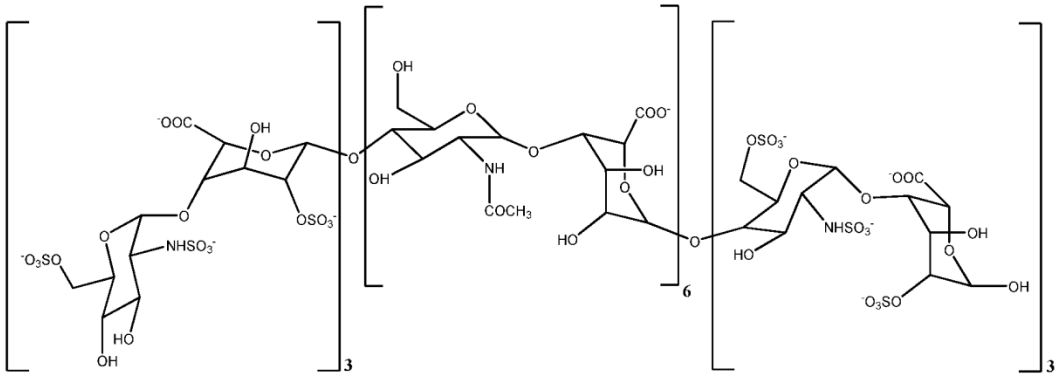
Table S3. MM/PB(GB)SA energy component analysis of the interactions of the CXCL-8 dimer with a 24-mer heparin fragment averaged over 50 ns[a].

	Complex		Receptor		Ligand		Δ	
	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
ELE	-3013.43	170.86	-2686.21	75.41	7055.72	81.09	-7382.93	155.51
VDW	-651.73	22.39	-521.26	18.74	-15.86	10.14	-114.6	9.33
INT	2943.24	40.8	2351.58	35.52	591.67	21.22	0	0
GAS	-721.91	167.95	-855.9	78.43	7631.53	81.45	-7497.53	158.83
PBSUR	63.94	1.02	49.03	0.68	30.53	0.23	-15.62	0.78
PBCAL	-7100.62	568.98	-3691.11	65.41	-10631.2	576.83	7221.68	389.61
PBSOL	-7036.68	568.88	-3642.08	65.44	-10600.67	576.85	7206.06	389.46
PBELE	-10114.04	531.35	-6377.32	22.48	-3575.48	574.29	-161.24	365.28
PBTOT	-7758.59	532.76	-4497.98	37.08	-2969.14	575.21	-291.47	365.6
GBSUR	83.71	1.36	63.91	0.91	39.33	0.31	-19.53	1.03
GBCAL	-6013.45	155.77	-3663.24	67.24	-9720.19	74.51	7369.98	147.27
GBSOL	-5929.73	155.25	-3599.33	67.28	-9680.86	74.61	7350.45	146.66
GBELE	-9026.87	32.36	-6349.45	20.41	-2664.48	16.88	-12.94	15.44
GBTOT	-6651.64	41.56	-4455.23	36.02	-2049.33	19.69	-147.08	16.48
$T\Delta S$[b]							-140.6	28.43
$\Delta G_{binding\text{-}PBSA}$[b]							-150.87	
$\Delta G_{binding\text{-}GBSA}$[b]							-6.48	

a Average over snapshots taken from a 30 ns MD simulation trajectory. b Entropy calculations were based on normal modes analysis using 60 snapshots.

ELE, non-bonded electrostatic energy; VDW, non-bonded van der Waals energy; INT, bond, angle, dihedral energies; GAS, ELE+VDW+INT; PBSUR, hydrophobic contribution to solvation free energy for PB calculations; PBCAL, reaction field energy calculated by PB; PBSOL=PBSUR+PBCAL; PBELE=PBCAL+ELE; PBTOTAL=PBSOL+GAS; GBSUR, hydrophobic contributions to solvation free energy for GB calculations; GBCAL, reaction field energy calculated by GB; GBSOL=GBSUR+GBCAL; GBELE=GBCAL+ELE; GBTOTAL=GBSOL+GAS; T⊿S, T(temperature); ⊿S (sum of rotational, translational and vibrational entropies); ⊿G binding, total binding energy of the system. All energies are in kcal/mol.

J Chem Inf Model, 2011, 51, 335-358



$(GlcNS6S\text{-}IdoA2S)_3\text{-}(GlcNAc\text{-}GlcA)_6\text{-}(GlcNS6S\text{-}IdoA2S)_3$

Figure S1. Chemical structure of the 24-mer heparin fragment. Dimeric CXCL-8 is believed to interact with this 24-mer heparin fragment in an horseshoe manner wherein the six tri-sulphated saccharides (sulphated domain consisting of GlcNS6S-IdoA2S disaccharide repeats) binds to each monomer and a flexible linker consisting of 12 sugar residues (the N-acetylated domain formed by GlcNAc-GlcA disaccharide repeats) bridges the two sulphated domains.

J Chem Inf Model, 2011, 51, 335-358

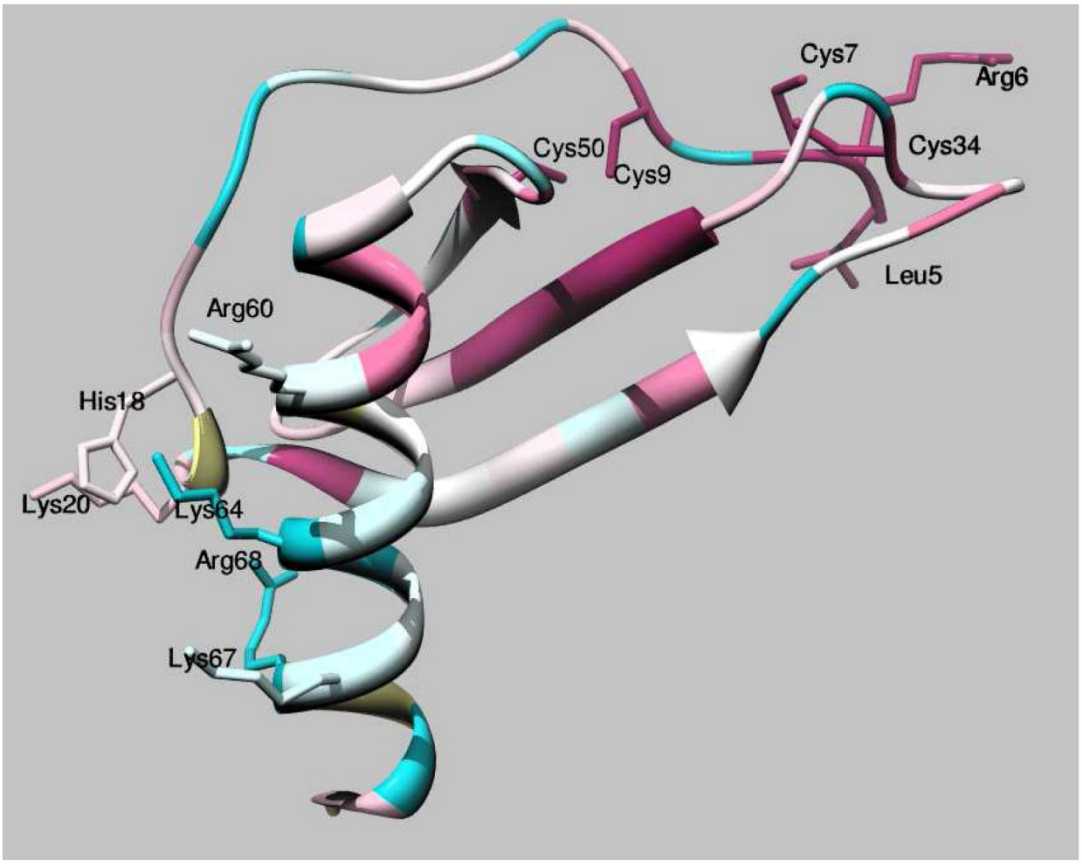


Figure S2. Amino acid residue conservation in the structure of CXCL-8 (PBD code 3IL8) obtained using default Consurf parameters. The structure is coloured according to the Consurf color scheme, from turquoise to white to burgandy to represent conservation scores from variable through to conserved, respectively.

J Chem Inf Model, 2011, 51, 335-358

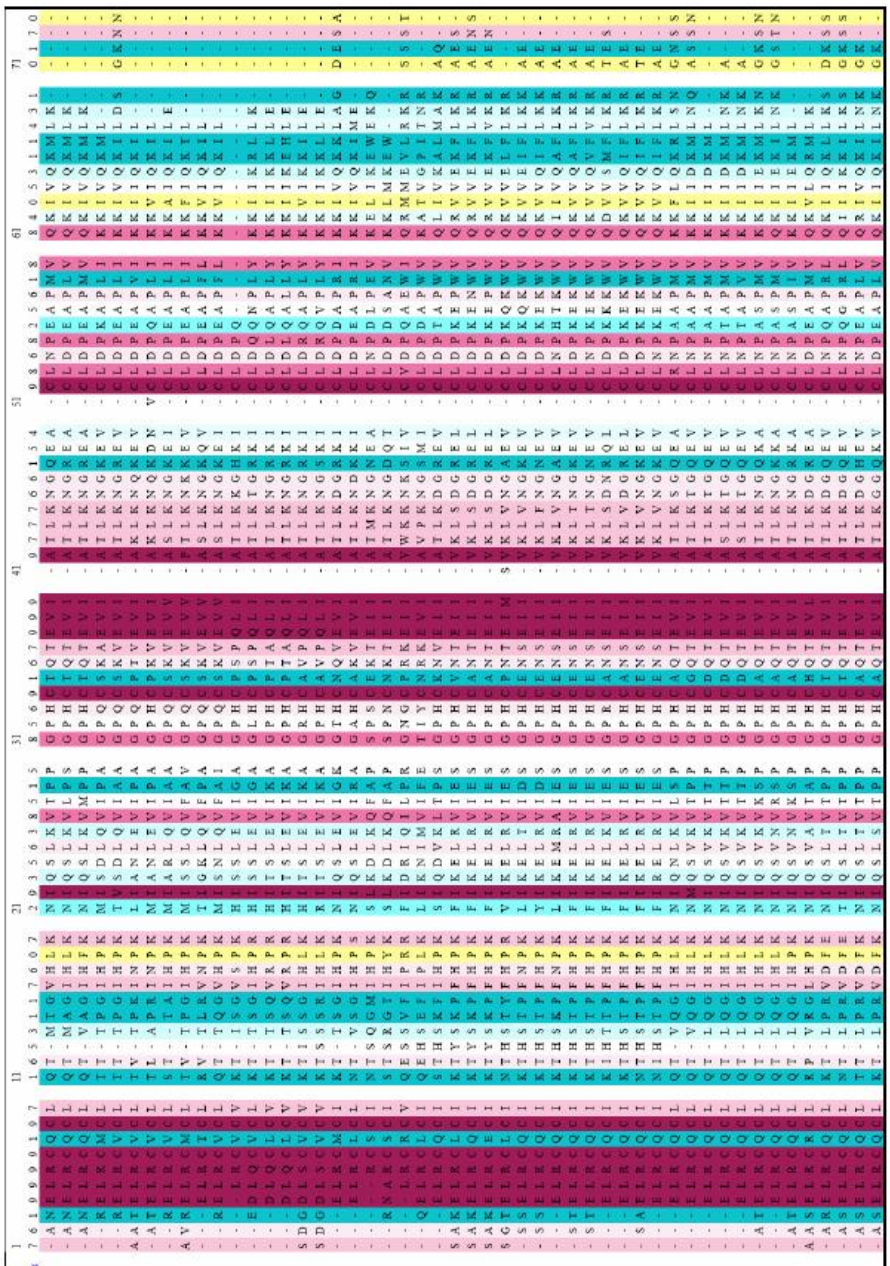


Figure S3. Multiple sequence alignment of CXCL-8 homologous sequences obtained using Consurf DB. The alignment is coloured according to the Consurf colour scheme, from turquoise to white to burgundy, to represent conservation scores from variable through to conserved, respectively. The sequences are referred to by their SWISS-PROT accession numbers.

J Chem Inf Model, 2011, 51, 335-358

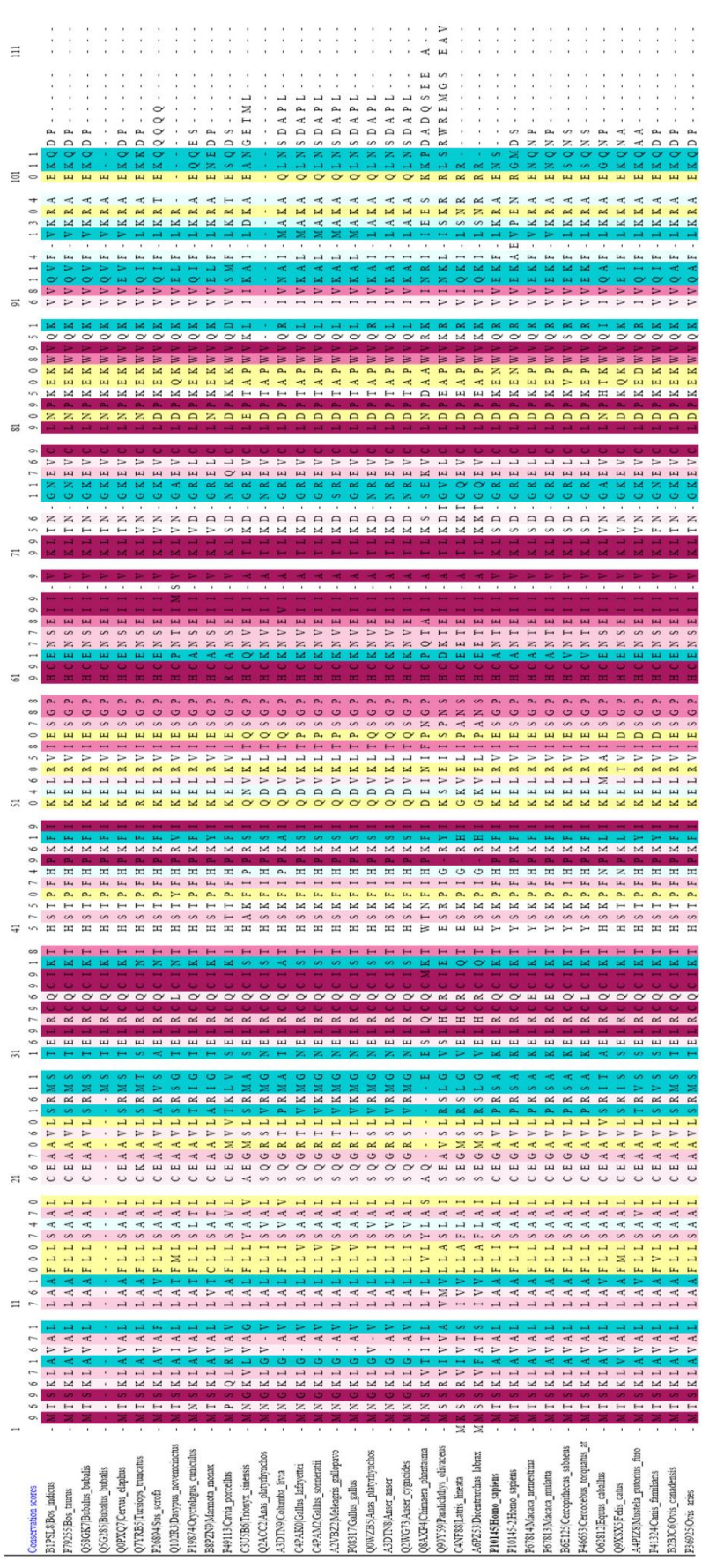


Figure S4. Multiple sequence alignment of CXCL-8 protein sequences collected from SWISS-PROT using a BLAST search as mentioned in Table 1. The alignment is coloured according to the Consurf color scheme, from turquoise to white to burgundy, to represent conservation scores from variable through to conserved, respectively. The sequences are referred to by their SWISS-PROT accession numbers.

J Chem Inf Model, 2011, 51, 335-358

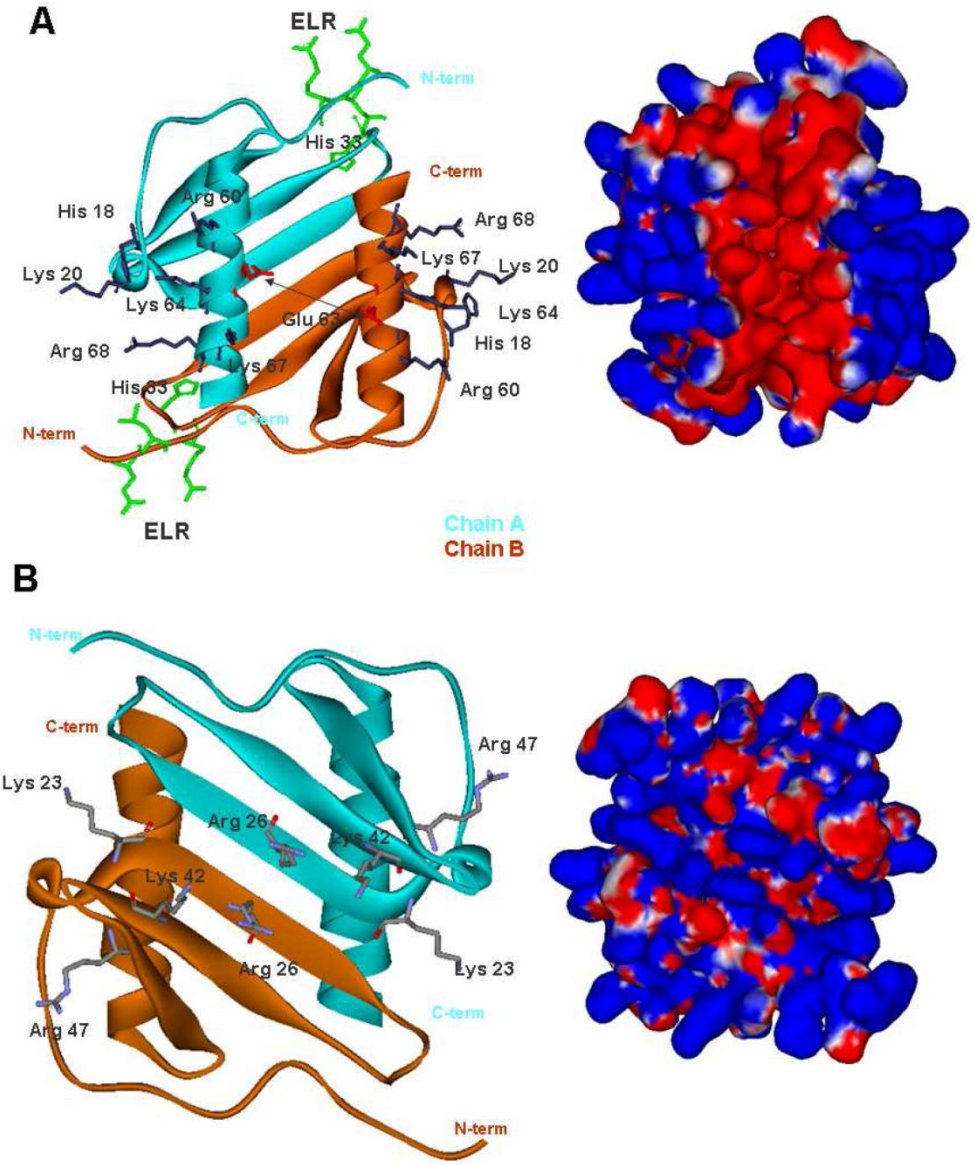


Figure S5. Ribbon diagrams highlighting those residues in the CXCL-8 dimer (taken from the X-ray structure, PDB code 3IL8) involved in the binding to receptors and GAGs. The electrostatic potential surface of CXCL-8 is also shown. Blue colour represents positively charged surfaces and red indicates negatively charged surfaces. (A) The residues of CXCL-8 involved in GAG and receptor binding are shown on the C-terminal helix. The adjacent figure shows the electrostatic potential surface wherein the basic residues alongside the alpha helices form the positively charged region. (B) CXCL-8 is oriented to show the presence of basic residues on the beta sheet surface. The adjacent figure shows the electrostatic potential surface of these beta sheets.

J Chem Inf Model, 2011, 51, 335-358

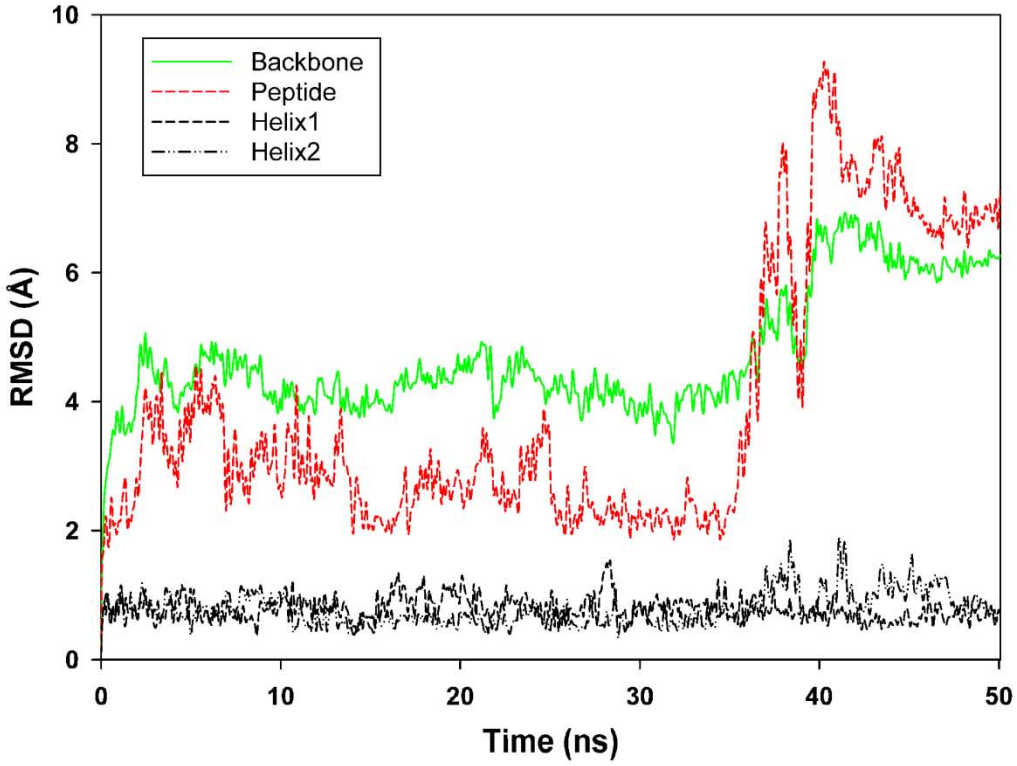


Figure S6. Time evolution of the average backbone RMSD of the whole CXCL-8 dimer, the two C-terminal helices of CXCL-8 and the receptor peptide. The flexibility in the loop of the receptor peptide makes the largest contribution to the overall RMSD of the CXCL-8-receptor peptide complex.

J Chem Inf Model, 2011, 51, 335-358

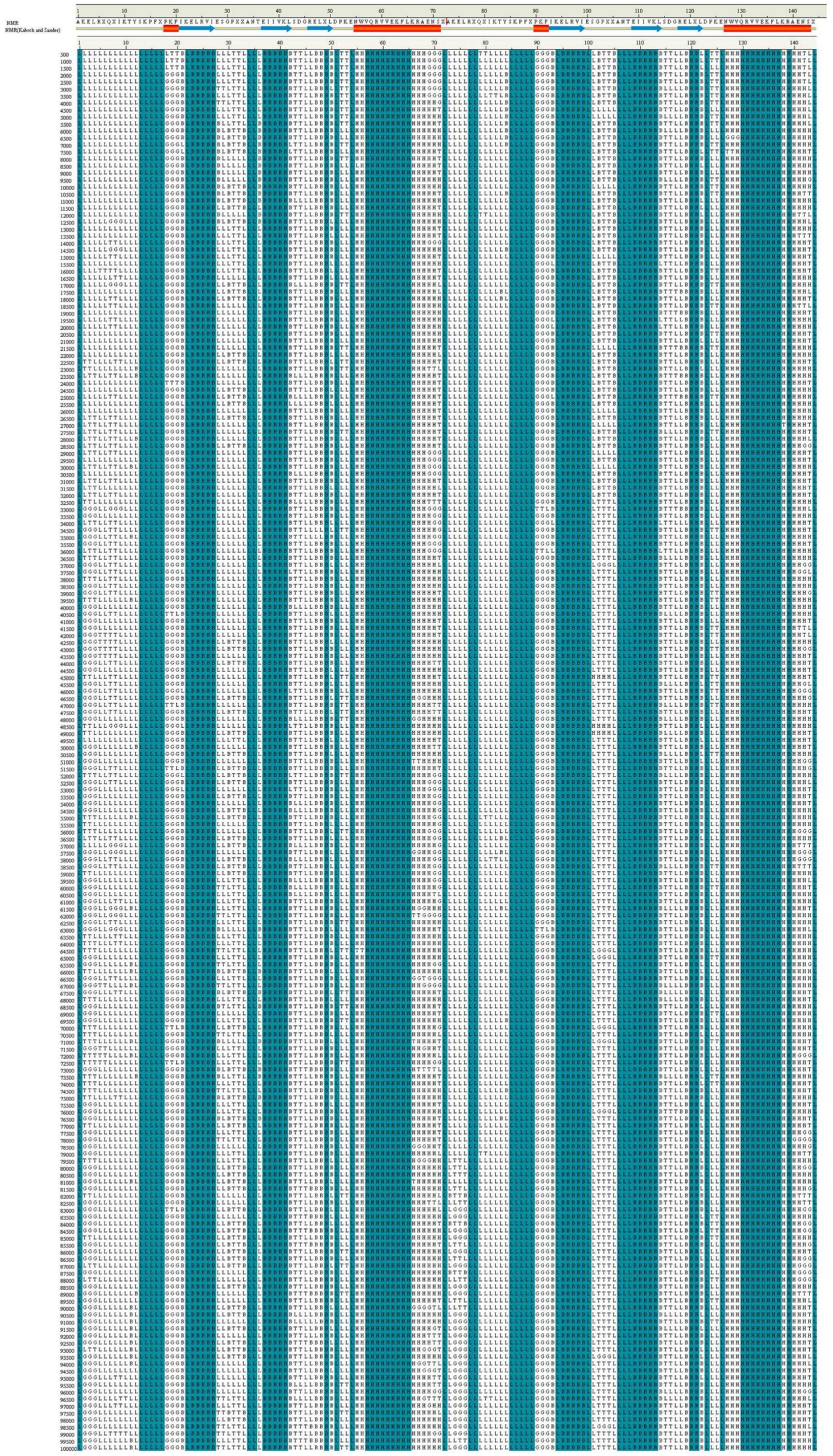


Figure S7. Secondary structure "dynamic conservation' analysis using the DSSP method in the simulation starting from the NMR structure. The first line indicates residue numbering and the secondary structure representation of the initial CXCL-8 dimer structure. X represents the capping residues. Every residue in every snapshot is indicated by "H" for an α-helix, "G" for a 3_{10}-helix, "I" for a pi-helix, "T" for a turn, "L" for a loop, "b" for a parallel β-sheet and "B" for an anti-parallel β-sheet. The blue background indicates portions of secondary structure conserved in a particular region. The numbers on the left refer to snapshots taken at picosecond intervals. This image was prepared using Discovery Studio 2.5. Monomer 1 ranges from residues 1 to 72 and monomer 2 ranges from residues 73 to 144.

J Chem Inf Model, 2011, 51, 335-358

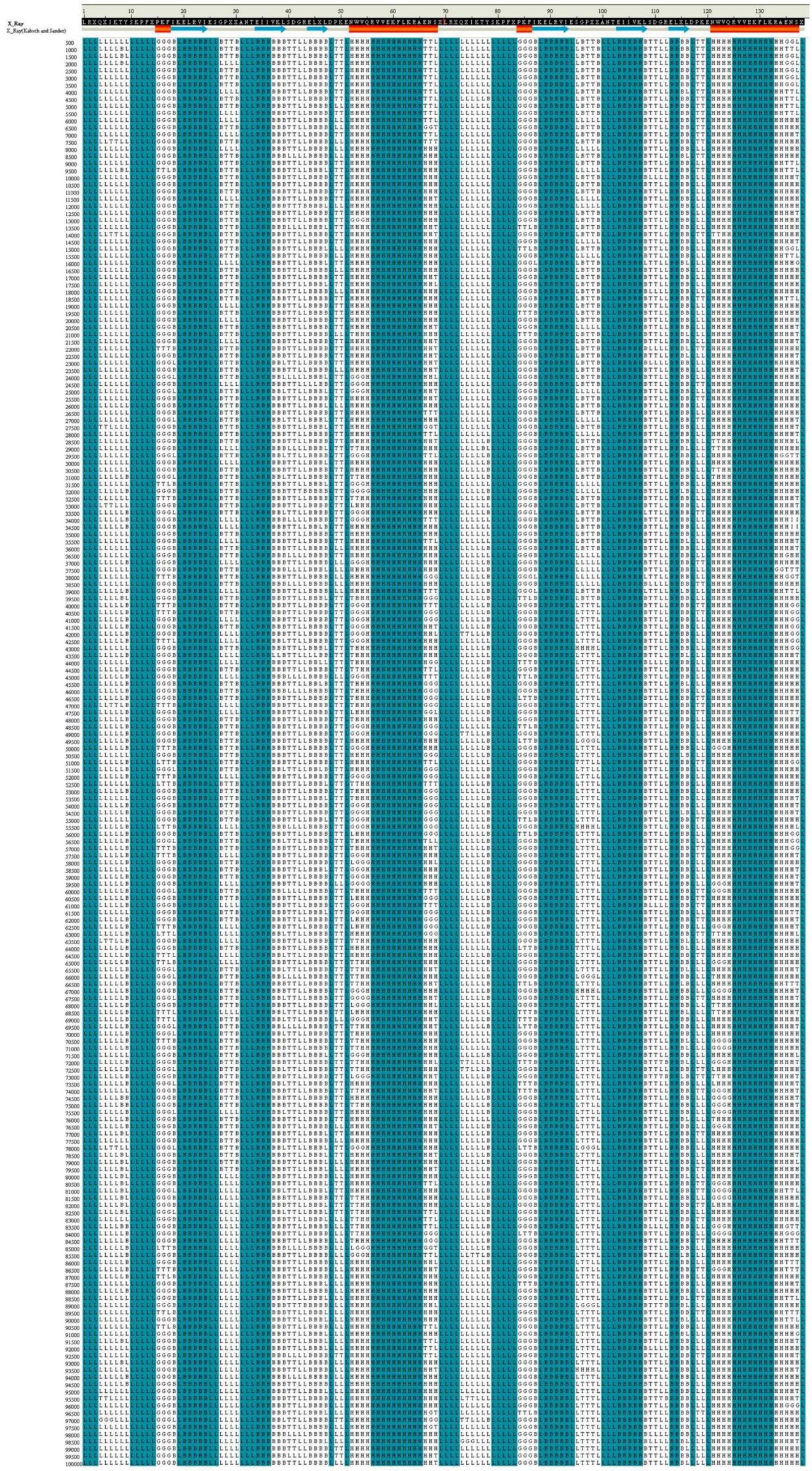
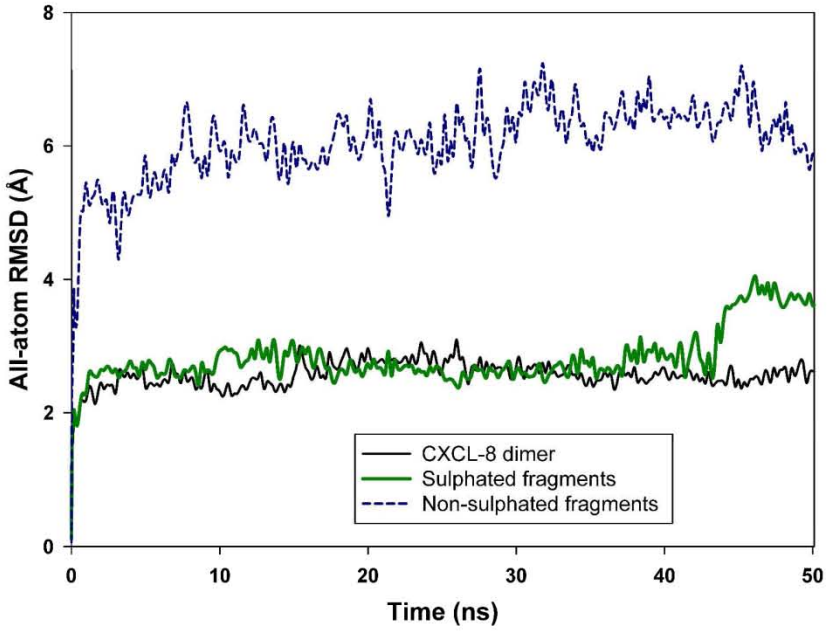


Figure S8. Secondary structure "dynamic conservation' analysis using the DSSP method in the simulation starting from the X-ray structure. The first line indicates residue numbering and the secondary structure representation of initial CXCL-8 dimer structure. X represents the capping residues. Every residue in every snapshot is indicated by "H" for an α-helix, "G" for a 3_{10}-helix, "I" for a pi-helix, "T" for a turn, "L" for a loop, "b" for a parallel β-sheet and "B" for an anti-parallel β-sheet. The blue background indicates portions of secondary structure conserved in a particular region. The numbers on the left refer to snapshots taken at picosecond intervals. This image was prepared using Discovery Studio 2.5. Monomer 1 ranges from residues 1 to 69 and monomer 2 ranges from residues 70 to 138.

J Chem Inf Model, 2011, 51, 335-358

A



B

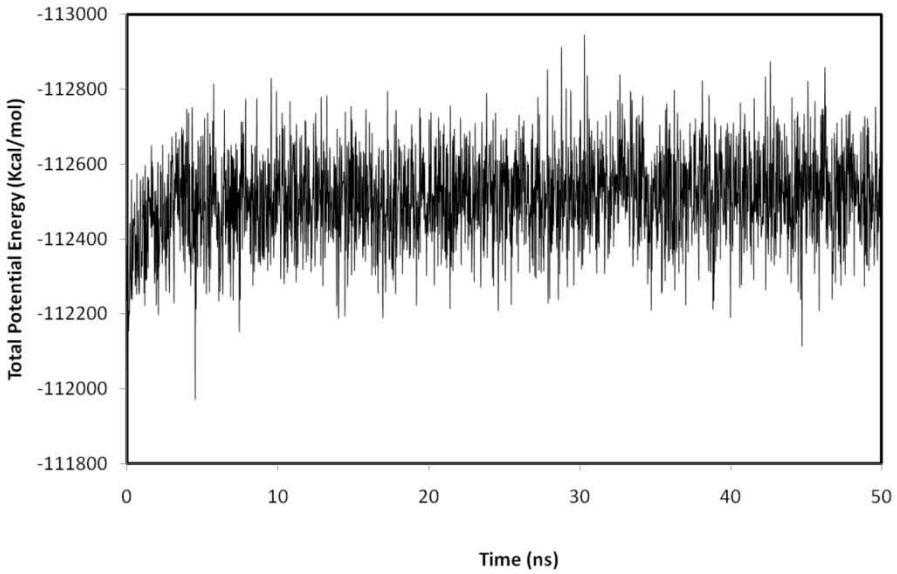


Figure S9. (A) Time evolution of the average all-atom RMSD of the CXCL-8 dimer, the two sulphated and two non-sulphated regions of the 24-mer heparin fragment. The non-sulphated region of heparin is more flexible compared to the sulphated regions, which are bound to the protein in the CXCL-8-heparin complex. (B) The total potential energy of the system (CXCL-8- 24-merheparin oligosaccharide complex) converged at the end of simulation.

J Chem Inf Model, 2011, 51, 335-358

MATERIALS AND METHODS

Consurf analysis

Given a structure and an alignment of homologous sequences, the ConSurf server[1] maps phylogenetic information onto protein tertiary structure to identify putative functional regions. To map evolutionary conservation scores onto protein structure, a MUSCLE alignment of CXCL-8 sequences and the X-ray crystal structure of CXCL-8 (PDB code 3IL8) were submitted to the ConSurf server. The sequence homologues of human CXCL-8 in SWISS-PROT[2] were collected using PSI-BLAST[3] and the selected homologues were aligned using MUSCLE[4] to produce a multiple sequence alignment (MSA). The Rate4Site algorithm[5] was subsequently used to construct a phylogenetic tree using the neighbour joining algorithm and to calculate evolutionary conservation scores. An empirical Bayesian approach was used to calculate the evolutionary rate of each amino acid position of the MSA, taking into account the stochastic nature of the evolutionary process. Amino acid evolution is traced using the JTT substitution model.[6] The conservation codes were projected on the 3D structure of CXCL-8 and colour coded, where 1 corresponds to maximal variability and 9 to maximal conservation. While a pre-calculated analysis for this structure can be found at the ConSurf-DB site[7], it consists of all interleukin sequences including C-C and CXC chemokines collected from the SWISS-PROT database using the PSI-Blast algorithm.

The MM/PBSA method was used to calculate free energies of binding. This method combines explicit solvation simulations with Poisson–Boltzmann calculations and non-polar solvation free energy calculations to estimate the free energy of binding.[8, 9] Free energies of binding are defined as

$$\Delta G_{binding} = \Delta G_{gas} + \Delta G_{sol\text{-}cmplx} - [\Delta G_{sol\text{-}prot} + \Delta G_{sol\text{-}lig}] \quad (1)$$

where ΔG_{gas} is the interaction energy between protein and ligand in the gas phase, which is calculated using a molecular mechanics approach. $\Delta G_{sol\text{-}prot}$, $\Delta G_{sol\text{-}lig}$ and $\Delta G_{sol\text{-}cmplx}$ are the solvation free energies of the protein, ligand and ligand-protein complex, respectively, which are estimated using a continuum Poisson-Boltzmann/surface area approach (or the generalised Born/surface approach in the case of GBSA calculations). In MM/PBSA, the free energy of binding is calculated as

$$\Delta G_{binding} = \Delta E_{MM} + \Delta G_{PBSA} - T\Delta S \quad (2)$$

where ΔE_{MM} is the difference in the average molecular mechanics energy, which is calculated as

$$\Delta E_{MM} = \Delta E_{int} + \Delta E_{vdw} + \Delta E_{elec}, \quad (3)$$

where $\Delta E_{int} = \Delta E_{bond} + \Delta E_{angle} + \Delta E_{tors}$.

ΔE_{int} corresponds to the sum of the differences in average internal bond stretching, bond bending and torsional angle energies. ΔE_{vdw} is the difference in average van der Waals energy, whilst ΔE_{elec} is the difference in average electrostatic energy.

ΔG_{PBSA} is the free energy of solvation, given by:

$$\Delta G_{PBSA} = \Delta G_{PB} + \Delta G_{SA}. \quad (4)$$

where ΔG_{PB} is the electrostatic component of the free energy of solvation calculated by solving the Poisson–Boltzmann equation or the generalised Born equation in the case of GB calculations.[10] ΔG_{SA} is the non-polar contribution to the free energy of solvation calculated from the solvent-accessible surface area (SASA).[11] This term is computed with the equation $\Delta G_{SA} = \gamma SA + \beta$, where SA is the solvent-accessible surface area calculated by the MSMS program,[12] and γ and β are parameterised constants.

The final term, $T\Delta S$, is the change in vibrational, translational and rotational entropies upon formation of the ligand-protein complex, and is approximated by performing normal modes calculations.

J Chem Inf Model, 2011, 51, 335-358

REFERENCES

1. Landau, M.; Mayrose, I.; Rosenberg, Y.; Glaser, F.; Martz, E.; Pupko, T.; Ben-Tal, N., ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.* **2005,** 33, (Web Server issue), W299-W302.
2. Boeckmann, B.; Bairoch, A.; Apweiler, R.; Blatter, M.-C.; Estreicher, A.; Gasteiger, E.; Martin, M. J.; Michoud, K.; O'Donovan, C.; Phan, I.; Pilbout, S.; Schneider, M., The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **2003,** 31, (1), 365-370.
3. Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **1997,** 25, (17), 3389-3402.
4. Edgar, R. C., MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004,** 32, (5), 1792-1797.
5. Pupko, T.; Bell, R. E.; Mayrose, I.; Glaser, F.; Ben-Tal, N., Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. *Bioinformatics* **2002,** 18, (suppl_1), S71-S77.
6. Jones, D. T.; Taylor, W. R.; Thornton, J. M., The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **1992,** 8, (3), 275-282.
7. Goldenberg, O.; Erez, E.; Nimrod, G.; Ben-Tal, N., The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures. *Nucleic Acids Res.* **2009,** 37, (Database issue), D323-D327.
8. Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A., Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* **1998,** 120, (37), 9401-9409.
9. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W., Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000,** 33, (12), 889-897.
10. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T., Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990,** 112, (16), 6127-6129.
11. Sitkoff, D.; Sharp, K. A.; Honig, B., Accurate calculation of hydration free energies using macroscopic solvent models. *J Phys Chem* **1994,** 98, (7), 1978-1988.
12. Sanner, M. F.; Olson, A. J.; Spehner, J.-C., Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers* **1996,** 38, (3), 305-320.

5

5 Computational Analyses of the Catalytic and Heparin Binding Sites and their Interactions with Glycosaminoglycans in Glycoside Hydrolase Family 79 Endo-β-D-Glucuronidase (Heparanase)

Computational analyses of the catalytic and heparin binding sites and their interactions with glycosaminoglycans in glycoside hydrolase family 79 Endo-β-D-glucuronidase (heparanase)

Neha S. Gandhi[2,3], Craig Freeman[5], Christopher Parish[5] and Ricardo L. Mancera[1,2,3,4]

[2]Curtin Health Innovation Research Institute, Western Australian Biomedical Research Institute, [3]School of Biomedical Sciences and [4]School of Pharmacy, Curtin University, GPO Box U1987, Perth, WA 6845, Australia. [5]Cancer and Vascular Biology Group, Department of Immunology, The John Curtin School of Medical Research, The Australian National University, P. O. Box 334, Canberra, ACT 2601, Australia.

Received on April 23, 2011; revised on July 4, 2011; accepted on July 6, 2011

Mammalian heparanase is an endo-β-glucuronidase associated with cell invasion in cancer metastasis, angiogenesis, and inflammation. Heparanase cleaves heparan sulphate proteoglycans in the extracellular matrix and basement membrane, releasing heparin/heparan sulphate oligosaccharides of appreciable size. This in turn causes the release of growth factors, which accelerate tumour growth and metastasis. Heparanase has two glycosaminoglycan binding domains; however, no three-dimensional structure information is available for human heparanase that can provide insight into how the two domains interact to degrade heparin fragments. We have constructed a new homology model of heparanase that takes into account the most recent structural and bioinformatics data available. Heparin analogues and glycosaminoglycan mimetics were computationally docked into the active site with energetically-stable ring conformations and their interaction energies were compared. The resulting docked structures were used to propose a model for substrate and conformer selectivity based on the dimensions of the active site. The docking of substrates and inhibitors indicate the existence of a large binding site extending at least two saccharide units beyond the cleavage site (towards the non-reducing end) and at least three saccharides towards the reducing end (towards heparin-binding site 2). The docking of substrates suggests that heparanase recognises the N-sulphated and O-sulphated glucosamine at subsite +1 and glucuronic acid at the cleavage site, whereas in the absence of 6-O-sulphation in glucosamine, glucuronic acid is docked at subsite +2. These findings will help to focus the rational design of heparanase-inhibiting molecules for anti-cancer drug development by targeting the two heparin/heparan sulphate recognition domains.

Keywords: Heparin, heparanase, molecular dynamics, glycol-split, glycosidase.

Introduction

Mammalian heparanase (Heparanase 1, HPSE or Hpa1) is an endo-β-D-glucuronidase that hydrolyses/cleaves HSPGs (heparan sulphate proteoglycans) at sites of injury or inflammation (Dempsey, L.A., Plummer, T.B., et al. 2000, Parish, C.R., Freeman, C., et al. 2001). This facilitates the structural alteration of the ECM and disrupts the basement membrane to allow extravasation of inflammatory cells and liberation of growth factors and chemokines that induce proliferation and migration of endothelial cells and fibroblasts (Vlodavsky, I. and Friedmann, Y. 2001). Heparanase activity may therefore play a decisive role in disease-related processes such as cell invasion, angiogenesis, and cancer metastasis (Barash, U., Cohen-Kaplan, V., et al. 2010b, Elkin, M., Ilan, N., et al. 2001, Nakajima, M., Irimura, T., et al. 1988, Parish, C.R., Freeman, C., et al. 2001, Vlodavsky, I., Elkin, M., et al. 2008).

[1] Curtin Health Innovation Research Institute, Western Australian Biomedical Research Institute, Curtin University, GPO Box U1987, Perth, WA 6845, Australia. Email: R.Mancera@curtin.edu.au, Phone: +61 8 9266 1017. Fax: +61 8 9266 2342.

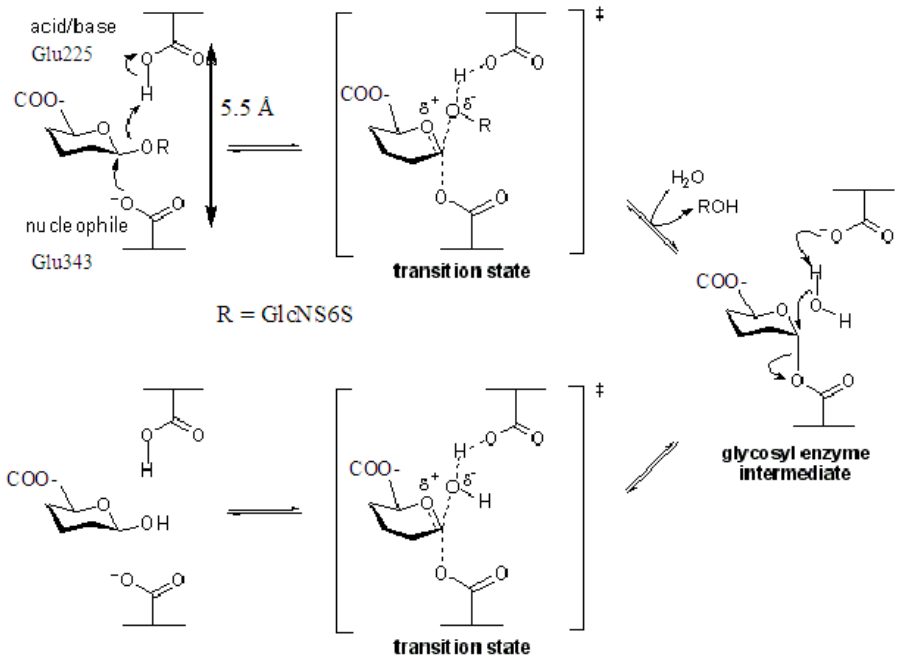
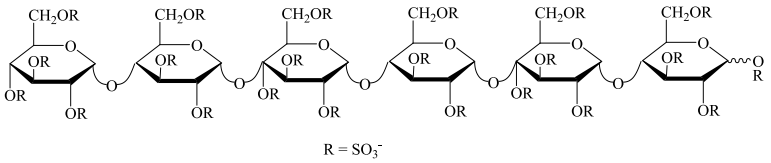
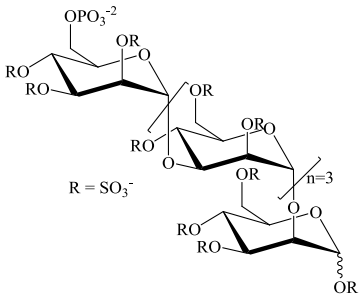
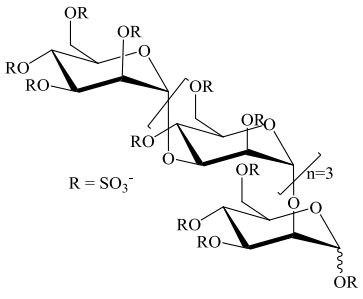


Fig. 1. Molecular representation of the mechanism of action of heparanase based on the retention mechanism of a β-glycosidase. Retaining glycosidases involve two carboxylates: one acts as a nucleophile and the other as an acid/base, with the distance between the nucleophile and acid/base being less than 5.5 Å. These enzymes operate through a two-step mechanism, with each step resulting in inversion, for a net retention of stereochemistry. During the first step, the nucleophile attacks the anomeric center, resulting in the formation of a glycosyl enzyme intermediate, with acidic assistance provided by the acidic carboxylate. In the second step, the now deprotonated acidic carboxylate acts as a base and assists a nucleophilic water to hydrolyze the glycosyl enzyme intermediate, giving the final hydrolyzed product.

Human heparanase cDNA encodes the pre pro-enzyme, which consists of 543 amino acid residues (65 kDa) with six *N*-glycosylation sites (Simizu, S., Ishida, K., et al. 2004b), five cysteine residues and a NH_2 terminal signal peptide (Met1-Ala35). Signal peptidase cleavage of the pre pro-enzyme at Ala35 generates the latent (pro-heparanase) 65 kDa heparanase. Proteolytic processing of the pro-enzyme removes a 6 kDa linker peptide (Ser110-Gln 157), yielding an 8 kDa domain at the N-terminus (Gln36-Glu109) and a non-covalently linked 50 kDa catalytic domain with a lysine residue at the N-terminus (Lys158-Ile543) (Fairbanks, M.B., Mildner, A.M., et al. 1999). These two polypeptides form a heterodimer that is the mature and active form of heparanase (Fairbanks, M.B., Mildner, A.M., et al. 1999). Heparanase acts with a β-retaining hydrolytic mechanism (Figure 1) and belongs to family 79, clan A glycosyl hydrolases, containing an $(\alpha/\beta)_8$ TIM-barrel fold (Henrissat, B. and Bairoch, A. 1993, Henrissat, B. and Bairoch, A. 1996, Hulett, M.D., Hornby, J.R., et al. 2000, Parish, C.R., Freeman, C., et al. 2001). Site-directed mutagenesis and sequence alignments have indicated that catalysis involves two conserved acidic residues (Figure 1), a proton donor at Glu225 and a nucleophile at Glu343 (Hulett, M.D., Hornby, J.R., et al. 2000).

Three heparin/HS binding sites have been identified: HBD-1 (domain 1, residues Lys158-Asn162), HBD-2 (domain 2, residues Gln270-Lys280) and HBD-3 (domain 3, residues Lys411–Arg432) (Levy-Adam, F., Abboud-Jarrous, G., et al. 2005). Attention has focused on HBD-1 and HBD-2 because of the presence of the GAG consensus motifs XBBXBX and XBBBXXBX (where B is basic and X is neutral and hydrophobic amino acid). The amino acids present in HBD-1 [KKFKNSTYSRSSVD(C); termed KKDC] directly interact with heparin/HS, resulting in the inhibition of both heparanase uptake and enzymatic activity, most likely due to competition of HS with the substrate. Notably, the KKDC sequence is in close proximity to catalytic residues Glu225 and Glu343 in the active site of the enzyme, thus making it an important target in the development of heparanase inhibitors. The above mode of action probably represents a HS-dependent non-enzymatic function of heparanase in its simplest form; however, evidence gathered in recent years indicates that heparanase also elicits HS-independent signalling (Fux, L., Ilan, N., et al. 2009).

Table I. The structures of human heparanase sensitive (enzymatically cleaved) and inhibitory oligosaccharides

Substrates/Inhibitors	*Oligosaccharide units*
1	GlcA-GlcNAc (heparosan)
2	GlcA-GlcNAc6S
3	GlcA-GlcNS
4	GlcA-GlcNS6S
5	GlcA-GlcNS3S
6	GlcA2S-GlcNS
7	IdoA2S-GlcNS
8	GlcNS6S-GlcA-GlcNS
9	AGA*IA
10	gs-AGA*IA
11	MHS 
12	PI-88 
13	

ΔHexA = unsaturated uronic acid, GlcA = β-D-glucuronic acid, GlcN = α-D-glucosamine, 2S = 2-*O*-sulphate, 6S = 6-*O*-sulphate, NS = 2-*N*-sulphate, AGA*IA = GlcNAc6S-GlcA-GlcNS,3,6S-IdoA2S-GlcNS6S, gs = glycol split, MHS = maltohexaose sulphate, PI-88 = phosphomannopentaose polysulfate, AGA*IA taken from X-ray structure (PDB:2GD4).

Heparanase cleaves both heparin and HS presumably at sites of low sulphation, releasing saccharide products of appreciable size (5–7 kDa) and biological potencies. Large efforts have been made to elucidate the substrate specificity and recognition mechanism of heparanase (Bame, K.J. 2001). Studies of the enzyme expressed in mouse and human melanoma cell lines have suggested that heparanase and heparin/HS interacting protein (HIP) recognise the same GAG sequence (Marchetti, D., Liu, S., et al. 1997) and may cleave the chain either within or adjacent to the octasaccharide that contains the highly sulphated anti-thrombin III (AT) binding sequence (Thunberg, L., Backstrom, G., et al. 1982). It has been observed that if the 2-*O*-sulphate group on the IdoA residue adjacent to the heparanase cleavage site is removed, the octasaccharide is no longer a substrate, suggesting that 2-*O*-sulphation on the IdoA or GlcA is essential for recognition by heparanase, but not *N*-sulphation (Pikas, D.S., Li, J.P., et al. 1998). Substrate specificity of recombinant heparanase was later investigated using tetra- and hexa-heparin/HS fragments (Okada, Y., Yamada, S., et al. 2002). It was found that (1) the minimum heparanase recognition site is a trisaccharide GlcN-GlcA-GlcN, wherein the GlcA is in the sulphated region; (2) the IdoA2S residue located two residues away from the target GlcA in the octasaccharide is not indispensable for heparanase cleavage; and (3) heparanase cleaves the glycosidic linkage between GlcA and GlcNS, where GlcNS is either 6-*O*-sulphated or 3,6-*O*-disulphated. A recent study revealed that heparanase cleaves repeating disaccharides with certain sulphation types (Peterson, S.B. and Liu, J. 2010). In addition to the cleavage of GlcA-GlcNS6S (or -3S), heparanase cleaves the linkage between GlcA and GlcN with a GlcA2S residue in close vicinity, but not IdoA2S. Interestingly, a polysaccharide with repeating IdoA2S-GlcNS units is not a substrate of heparanase, but rather inhibits the activity of this enzyme (Peterson, S.B. and Liu, J. 2010).

Different classes of molecules, including chemically modified natural products, small molecule inhibitors and antibodies, have been described as heparanase inhibitors (Ferro, V., Hammond, E., et al. 2004, Miao, H.Q., Liu, H., et al. 2006, Nasser, N.J., Avivi, A., et al. 2007). A strong affinity inhibitor of heparanase is heparin (Nakajima, M., Irimura, T., et al. 1984), wherein inhibition is best achieved by a heparin sequence containing 16 or more sugar units and having sulphate groups at both the *N* and *O* positions (Vlodavsky, I., Mohsen, M., et al. 1994). The HS mimetic PI-88 is a potent inhibitor of tumour growth and metastasis, and its anti-cancer activities are attributed are attributed to its inhibition of heparanase as well as antagonism of the interactions of angiogenic growth factors with their receptors (Ferro, V., Dredge, K., et al. 2007, Gandhi, N.S. and Mancera, R.L. 2010, Kudchadkar, R., Gonzalez, R., et al. 2008, Parish, C.R., Freeman, C., et al. 1999). Currently molecules are being developed that lack the anti-coagulant activity typical of heparin but that are potent inhibitors of the enzymatic activity of heparanase (Casu, B., Vlodavsky, I., et al. 2008, Irimura, T., Nakajima, M., et al. 1986, Naggi, A. 2005). Examples of such molecules are glycol split heparins (chemically-modified heparins through the perodate oxidation of non-sulphated uronic acid residues) (Fux, L., Ilan, N., et al. 2009, Naggi, A., Casu, B., et al. 2005). In the absence of the three-dimensional (3D) structure of heparanase, structure and/or ligand based approaches have been implemented in the rational design of new high-affinity inhibitors (Gozalbes, R., Mosulen, S., et al. 2009, Ishida, K., Hirai, G., et al. 2004, Ishida, K., Wierzba, M.K., et al. 2004, Jalali-Heravi, M., Asadollahi-Baboli, M., et al. 2008, McKenzie, E.A. 2007, Zhou, Z., Bates, M., et al. 2006).

This paper reports the molecular modelling study of various substrates and inhibitors such as glycol-split and PI-88 derivatives (Table I). A deeper understanding of the nature of heparanase-heparin interactions will provide crucial information for the future design of novel molecules which might act as anti-angiogenesis inhibitors to down-regulate the actions of heparanase.

Results and discussions

ConSeq analysis

A ConSeq analysis using the 543-residue long human heparanase as the query sequence was carried out. The ConSeq results (Figure 2) reveal that 53 residues are predicted to be functionally important and 37 residues to have a structural role. ConSeq assigned conservation grades above average (colour grades 6–9) in the catalytic domain, whereas the heparin binding domains are not predicted to be conserved.

Among all heparanases, heparanase-2 (which is unable to degrade its substrate) possesses only ~40% homology to heparanase-1 (Levy-Adam, F., Feld, S., et al. 2010, McKenzie, E., Tyson, K., et al. 2000). Two alternatively spliced variants of human heparanase-1 identified as having 485 and 169 amino acids are closely related to wild type active enzyme heparanase-1. The 485 isoform results from the skipping of exon 5 and, as a

consequence, lacks 174 base pairs encoding for 58 amino acids, including the active site proton donor (Glu225). Consequently, this alternatively spliced variant of heparanase-1 escapes proteolytic cleavage and is devoid of HS degradation activity (Nasser, N.J., Avivi, A., et al. 2007, Sato, M., Amemiya, K., et al. 2008). The 169 amino acid spliced form of human heparanase, termed T5, is a truncated, enzymatically inactive protein formed when 144 base pairs of intron 5 are joined with exon 4 (Barash, U., Cohen-Kaplan, V., et al. 2010a). The similarity tree (Figure 3) shows that Spalax is situated on a branch separate from the mouse and rat heparanases, as reported previously (Nasser, N.J., Nevo, E., et al. 2005), and rodents are situated in a cluster separate from the other mammals (human and bovine). The Florida lancelet *B. floridae* exhibits additional regions in the N-terminus although the core region is homologous to family 79 of glycosyl hydrolases. All the plant sequences are aligned on one side of the cladogram. These plant sequences comprise an additional conserved motif, e.g. "WWPpdKCdYgtC", where the lowercase indicates semi-conserved residues as compared to other species.

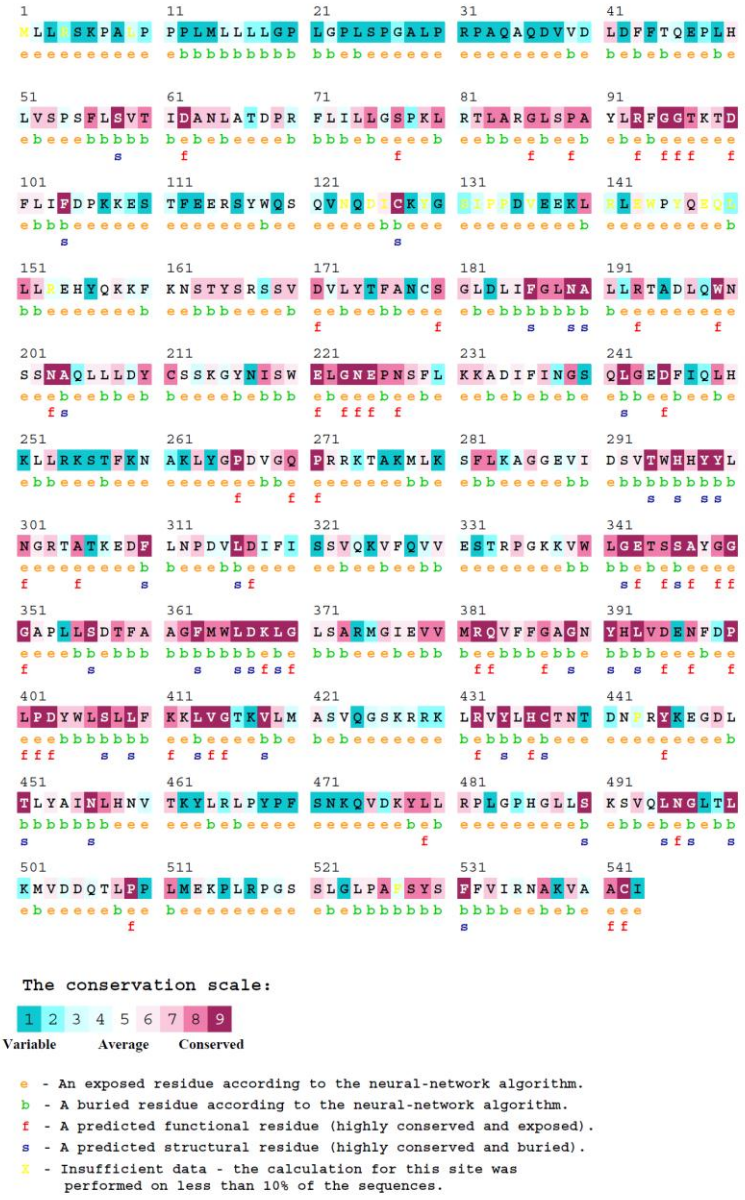


Fig. 2. ConSeq predictions on human heparanase (SWISS-PROT: Q9Y251), using nonredundant homologs obtained from the CLEAN_UNIPROT (filtered UNIPROT database in Consurf). The sequence of the query protein is displayed with the evolutionary rates color coded in each site (see legend). The residues of the query sequence are numbered starting from 1. The residues in the catalytic sites are predicted to be structurally and functionally important, "s" and "f", respectively, whereas the heparin-binding domains are not conserved but exposed to the surface.

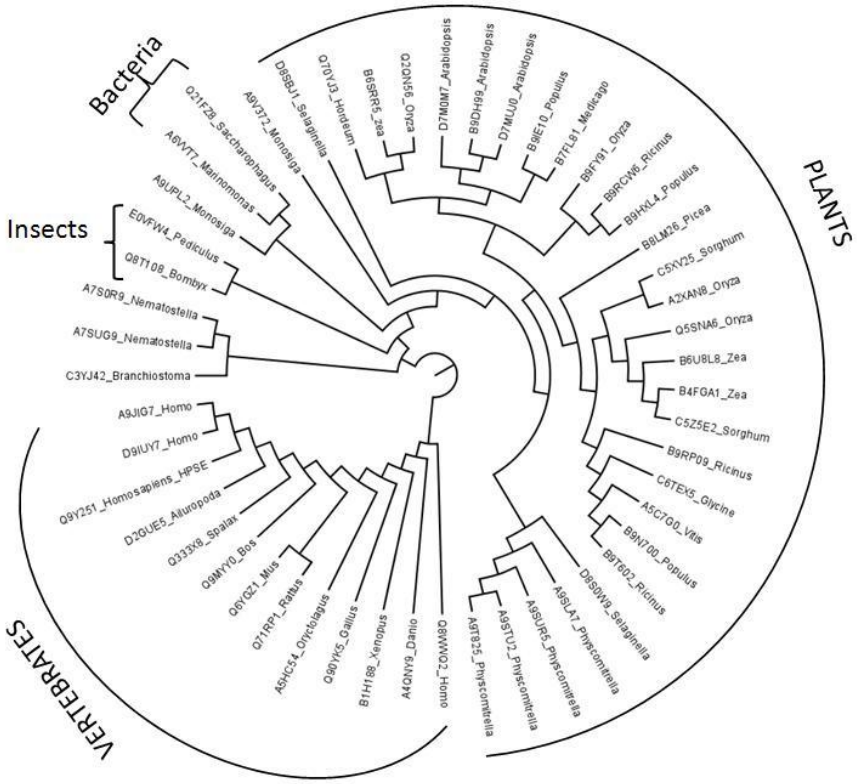


Fig 5. Polar phylogenetic trees of heparanase sequences derived from a multiple sequence alignment using MAFFT. The phylogenetic tree produced by ConSeq was plotted using FigTree 1.2.2.(Rambaut, A. 2006a)

Homology modelling

The secondary structure of the amino acid sequence predicted by the PSIPRED and NPSA servers are shown in Figures S1 and S2, respectively. The secondary structure prediction reveals the presence of approximately 25% helices, 12% strands and 63% coils at various positions in the protein structure, shown along with level of confidence of the predictions. However, some discrepancies were observed between the predictions from PSIPRED and those from NPSA. PSIPRED predicted the presence of beta sheets in the region Lys161-Thr164 and Leu183-Gly188, and predicted the presence of an α-helix in the region Tyr216-Leu222. On the other hand, PSIPRED, PHD and Predator predicted the presence of a strand in region Glu378-Gln383, whereas DSC and MLRC overestimated the extent of the α-helix preceding this region. These differences in secondary structure do not affect the modelling of HBD-1, HBD-2 and the catalytic residues.

Analyses of the heparanase sequence by various 3D protein structure modelling web-servers, such as Phyre (Kelley, L.A. and Sternberg, M.J. 2009), HHPred (Soding, J., Biegert, A., et al. 2005), I-TASSER (Roy, A., Kucukural, A., et al. 2010, Zhang, Y. 2008, Zhang, Y. 2009) and Fugue (Shi, J., Blundell, T.L., et al. 2001) resulted in structural template matches with a sequence identity <17%. The templates, along with their sequence identity from HHPred, are reported in Table II. The templates are all retaining enzymes with very similar $(\beta/\alpha)_8$ folds and are superimposable over the main chain atoms. However, most of the enzymes differ in their respective catalytic sites, depending on the ability of a glycoside hydrolase to cleave a substrate at the end (exo-acting) or within the middle of a chain (endo-acting). The active sites of exo-acting glycosyl hydrolases is pocket shaped, whereas endo-acting enzymes generally have an open, extended substrate binding cleft that can interact at random positions along the saccharide chains (Davies, G. and Henrissat, B. 1995). Hence, the selection of the templates on the basis of the type (endo- or exo-acting) of glycoside hydrolases is critical for the rational design of inhibitors of heparanase.

Table II. Structures and mechanism of reaction in various families of glycosyl hydrolase (clan A) families homologous to human heparanase.

PDB codes	***Enzyme***	***Organism***	***Substrate***	***Mechanism***	***Ref***	***Seq. Identity (%)***
2C7F 2C8N	Family 51 arabinofuranosidase, Araf51	Clostridium thermocellum	α-1,3-arabinofuranosyl from xylose	α-retaining, exo-acting	(Taylor, E.J., Smith, N.L., et al. 2006)	12
1QW9 1QW8 1PZ2 1PZ3	Family 51 arabinofuranosidase	Geobacillus stearothermophilus	α-1,3-arabinofuranosyl from xylose	α-retaining, exo-acting	(Hovel, K., Shallom, D., et al. 2003)	12
2VRQ 2VRK	Family 51 arabinofuranosidase	Thermobacillus xylanilyticus	Branched pentasaccharide and α-1,3-arabinofuranosyl from xylose	α-retaining, exo-acting	(Paes, G., Skov, L.K., et al. 2008)	12
3II1 3FW6	Family 44, Cellulase CelM2	Uncultured bacterium	β-1,3/4-glucan and β-1,4-xylan	β-retaining, endo-acting	(Nam, K.H., Lee, W.H., et al. 2010)	16
3IK2	Family 44, Endoglucanase	Clostridium acetobutylicum	cellotetraose, cellopentaose and cellohexaose	β-retaining, endo-acting	(Warner, C.D., Hoy, J.A., et al. 2010)	14
2E4T 2E0P 2E07 2EJ1 2EEX 2EQD	Family 44, Endoglucanase Cel44A	Clostridium thermocellum	cellooligosaccharide, xylan, lichenan, and various celluloses	β-retaining, endo-acting	(Kitago, Y., Karita, S., et al. 2007)	16
1W91 2BS9 2BFG	Family 39, β-D-Xylosidase	Geobacillus stearothermophilus	xylooligosaccharides	β-retaining, exo-acting	(Czjzek, M., Ben David, A., et al. 2005)	15
1UHV 1PX8	Family 39, β-D-Xylosidase	Thermoanaerobacterium saccharolyticum	xylooligosaccharides	β-retaining, exo-acting	(Yang, J.K., Yoon, H.J., et al. 2004)	15

The top hit templates (PDB codes 2C7F, 1QW9, 2VRQ, etc.) (Hovel, K., Shallom, D., et al. 2003, Paes, G., Skov, L.K., et al. 2008, Taylor, E.J., Smith, N.L., et al. 2006) belong to the exo-acting arabinofuranosidase family 51, which preferentially hydrolyses the α(1→3) linkage of arabinose-containing branched substrates and releases arabinofuranosyl moieties. The crystal structures of branched substrate complexes with arabinofuranosidase reveal the presence of extended substrate binding sites that appear to be induced by substrate binding. Hence, caution should be exercised while using members of this family as templates for the homology modelling of heparanase. The homology modelling of heparanase previously reported (Ishida, K., Hirai, G., et al. 2004, Zhou, Z., Bates, M., et al. 2006) was carried out using the structure of family 10 1,4,β-xylanase (PDB code 1BG4) (Schmidt, A., Schlacher, A., et al. 1998) which has a shallow active site. Similarly to heparanase, xylanase (PDB code 1BG4) is β-retaining and endo-acting; however, the two

enzymes differ in their substrates. Aldopentose serves as substrate for xylanase whereas hexapyranose is a substrate for heparanase. A current PSI-BLAST search shows β-D-xylosidase (PDB codes 1UHV, 2BS9, etc.) (Czjzek, M., Ben David, A., et al. 2005, Yang, J.K., Yoon, H.J., et al. 2004) to be a closer match; however, this enzyme hydrolyses the substrate at the end. Furthermore, Modbase considers family 30 human β-glucoside (PDB code 1OGS)(Dvir, H., Harel, M., et al. 2003) as the best template but is only able to provide a model of heparanase in the amino acid range 187-466 due to the lack of structural folds in other regions. On other hand, members of family 44 (clan not classified), which have composite domains of the glycosyl hydrolase families 5, 30, 39 and 51, are endo-acting and act on substrates containing (β-D-glucopyranosyl-(1→4))$_n$-β-D-glucose motifs (where n=1-5). Besides family 79, β-glucuronidase is classified under the glycosyl hydrolase family 2, which also catalyses the hydrolysis of β-D-glucuronic acid residues from the non-reducing end of HS. However, human β-glucuronidases in families 2 and 79 lack appreciable sequence and structural similarity (PDB codes 1BHG and 3HN3) (Jain, S., Drendel, W.B., et al. 1996). In particular, all threading programs fail to correctly align the HBD-1. Heparanase is endo-acting and cleaves pyranose substrates with a 1,4-β linkage. Consequently, endoglucanases (PDB codes 3IK2 and 2E4T) (Kitago, Y., Karita, S., et al. 2007, Warner, C.D., Hoy, J.A., et al. 2010) were selected as the most appropriate templates to model the heparanase sequence. The sequence alignment as obtained using HHPred is shown in Figure 4.

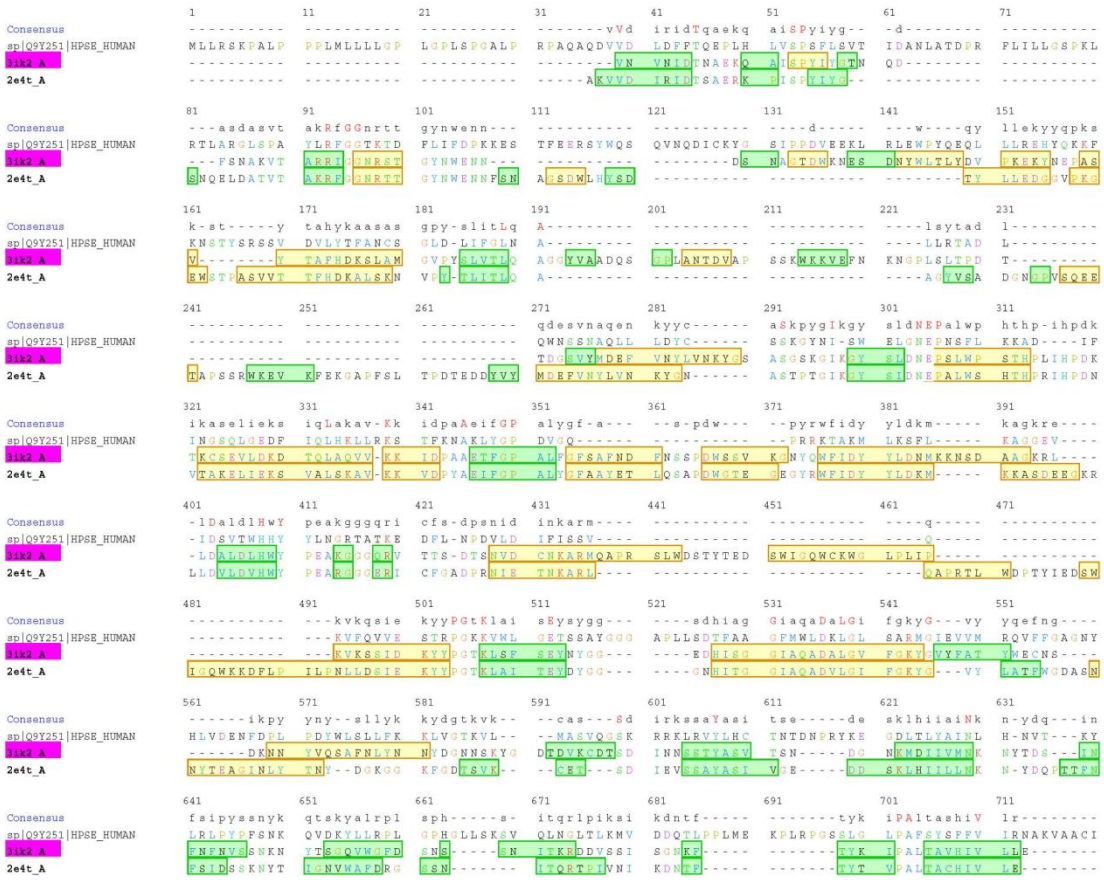


Fig. 4. Alignment of the sequence of heparanase with the sequences of two template proteins of family 44 endo-acting glucanases. The lighter rectangles show the conserved α-helices and the darker regions show the presence of β-strands. Overall, the alignment is consistent with the consensus secondary structure prediction.

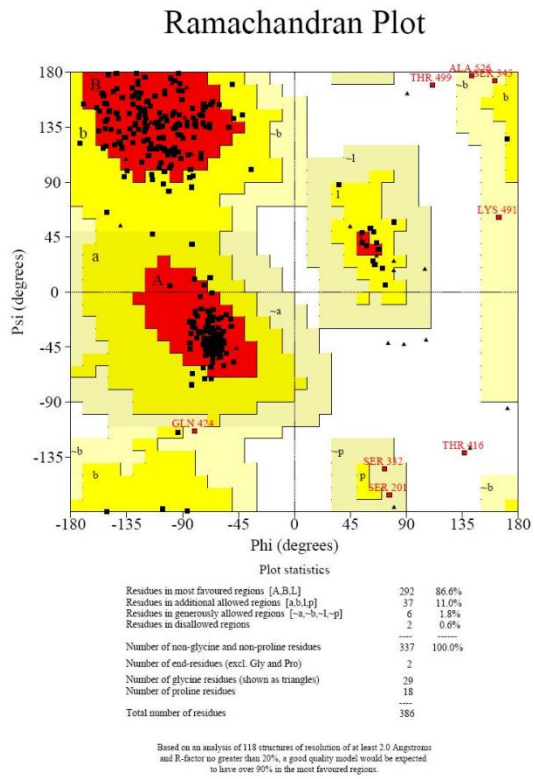


Fig. 5. Ramachandran plot of the derived homology model of heparanase (residues 125–543). It can be seen that 86.6, 11.0 and 1.8% of the residues are located within the most favorable, additionally allowed and generously allowed regions of the Ramachandran plot, respectively, with only two residues found in the disallowed region. Given the low sequence identity and abundant presence of coils, the model can be described as being of good quality.

Generally speaking, the alignment in the catalytic regions is consistent with the secondary structure prediction (Figure 4); however, sequence-structure alignment with the majority of the templates mentioned in Table II using any of the 3D recognition method fail to assign conserved secondary structures in the region from Asp125 to Gln157. As a consequence, secondary structure restraints were applied while producing a homology model with MODELLER. As a further evaluation of the homology-modelled structures of heparanase (residues 157-543), the model with the lowest probability density function (PDF) energy was selected from a set of 20 models and its backbone ψ and φ dihedral angles were analysed using a Ramachandran plot. In this model, 86.6%, 11.0% and 1.8% of the residues are located within the most favourable, additionally allowed and generously allowed regions, respectively, with only two residues being found in the disallowed region (Figure 5). The C-terminal residues were the ones in the generously allowed and disallowed regions. The heparin binding sites and the catalytic residues were found within the most favourable region of the Ramachandran plot. The homology model was also analysed using VERIFY 3D, which assessed and confirmed the quality of the predicted structure of heparanase with acceptable 3D-1D average scores. The scores were added and plotted for residues 157 to 543 (Figure 6). Residues at the C-terminal with values ≤ 0.1 are predicted with less accuracy and are likely to have less stable conformations. It can be seen that most of the residues, including the heparin binding domains and catalytic sites have values in the range 0.2-0.78. These assessments suggest that the modelled structure is of reasonable quality, given the low sequence homology between the templates and the target, as shown in Figure 4.

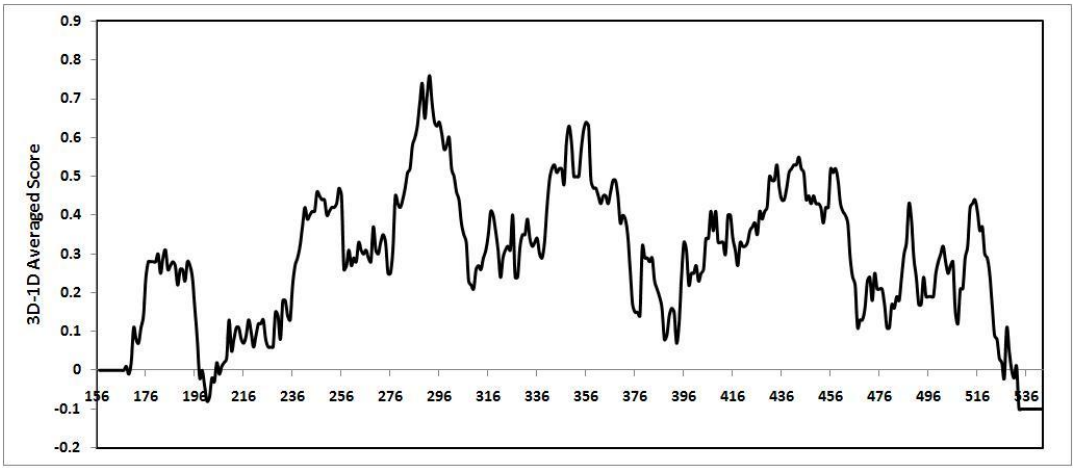


Fig. 6. Evaluation of the model of heparanase with VERIFY 3D. Very few residues cross the reference line (negative values), suggesting that the model is of good quality.

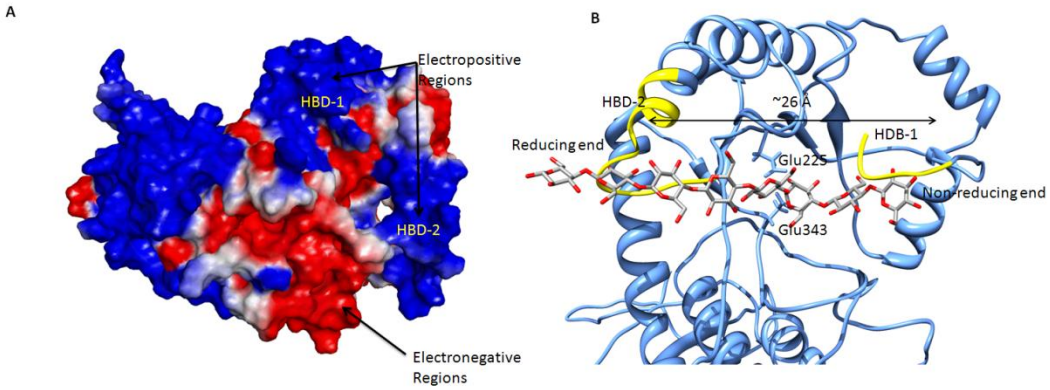


Fig. 7. (**A**) Electrostatic potential surface of the model of heparanase, showing the heparin-binding sites HBD-1 and HBD-2. The molecular surfaces are colored according to electrostatic potential. Positive potential is shown in blue and negative potential is in red. The figure is rotated by 90° to map the residues in the binding site. (**B**) Superimposition of the binding sites of cellooctaose (in stick) in Cel44A (PDB code 2EQD) onto the model of heparanase (represented in blue ribbons). The distance between the two domains (HBD-1 and HBD-2 shown in yellow) is large enough to allow binding of oligosaccharides with at least six pyranose rings.

Table III. Differences in residues between heparanase and *C. thermocellum* Cel44A based on a tertiary structure alignment.

Notes	**Residue in heparanase model**	**Residue in *C. thermocellum***
Residues in the catalytic site	Asn224	Asn185
	Glu225	Gln186
	Pro226	(mutant)
	Gly269	Pro187
	His296	Tyr233
	His297	His283
	Tyr298	Trp284
	Glu343	Tyr285
	Gln383	Glu359
		Trp392
Residues in HBD-1	Lys158	Ala60
	Lys159	Asp63
	Phe160	-
	Asn162	Trp 64
Residues in HDB-2	Gln270	Gly234
	Pro271	Phe235
	Arg272	-
	Arg273	-
	Lys274	-
	Thr275	Glu239
	Ala276	-
	Lys277	Trp256
	Met278	Phe257
	Leu279	Ile258
	Lys280	Asp259

Analysis of the molecular surface and electrostatic potential of modelled heparanase reveals the presence of electropositive regions (due to positively charged residues) in HBD-1 and HBD-2 (Figure 7A). The distance between the centroid of residues in HBD-1 and HBD-2 is approximately 26 Å, which is important for the binding of heparin oligosaccharides larger than a hexasaccharide. The crystal structure of the Cel44A structure bound with substrate (cellooctaose) from *C. thermocellum* (PDB code: 2EQD) was used to compare residues with the model of heparanase. This structure was chosen in order to get a general impression of how heparanase might require an oligosaccharide to bridge the heparin binding domains and cleave an octasaccharide (Figure 7B) that contains the highly sulphated anti-thrombin III (AT) binding sequence. Table III shows a list of the residue differences in the catalytic and heparin binding domains. The pyranose rings in cellooctaose are surrounded primarily by tryptophan residues, whereas other residues dominate this site in heparanase (Table III). Superimposition of a complex of cellulase with its substrate shows that the acidic groups of Glu225 and Glu343 are very close to the cleavage point between the β-1,4 glucose units; however, there are steric clashes observed with cellooctaose and HBD-1. HBD-1 has a coiled-coil structure at the terminal and may undergo further conformational changes during MD simulations. These steric clashes with HBD-1 can be avoided by replacing the β-1,4 linkage with an α-linkage in the ligand.

PSFloger (Protein Structure and Function Loger) (Mukherjee, S., Roy, A., et al. 2010) was used to identify the structural and functional analogues of heparanase. This webserver takes the ligand(s) in analogue structures and transfers them to the model structure to compare the local structure and sequence similarity in the binding site. It also annotates the function of the target protein (ligand-binding sites, Gene-Ontology terms, enzyme classifications) by global and local matches of the structure with known protein function libraries. Table S2 in the Supporting Information shows the binding site residues identified on the heparanase model. Members of family 30 glycoside hydrolase, which is responsible for hydrolysing the β-glucoside from the glycolipid glucosylceramide, show high resemblance in the catalytic binding site to heparanase (PDB codes 2VT0 and 2V3F). Although these analogue structures are co-crystallised with sulphate ions, none map to the heparin binding domains in the heparanase model upon superimposition. In fact, three sulphate ions map to the C-terminal β-barrel domain. These results suggest that the heparin binding domains (in particular HBD-1) are highly variable compared to other members of clan A, which might explain variations in specificity of different binding substrates.

For catalysis to take place the proton donor must be protonated and the nucleophilic attack and the stabilisation of the intermediate are mediated by a protein residue with a free lone pair or a negative charge (nucleophile). Consequently, the pK_a of the proton donor (Glu 225) should be at least 5.0 and the pK_a of the catalytic nucleophile (Glu 343) should be at least 1.5 units lower than the pK_a of the proton donor (Nielsen, J.E. and McCammon, J.A. 2003a, Nielsen, J.E. and McCammon, J.A. 2003b). Subsequent benchmarking of the two approaches (H++ and PROPKA) to predict pK_a values with the set of templates listed in Table II show the PROPKA approach to be superior, with pK_a values of titrable groups in the active site matching those mentioned in the literature. PROPKA gave equally good predictions with a set of xylanases reported previously (Davies, M.N., Toseland, C.P., et al. 2006, Kongsted, J., Ryde, U., et al. 2007). In all the glycosyl hydrolases, including the heparanase model reported here, the pK_a of the proton donor (Glu 225) was higher than 5.0 and hence it was protonated, while the pK_a of the catalytic nucleophile (Glu 343) was lower than the pK_a value of the proton donor. Table S3 (Supporting Information) lists residues chosen based on initial pK_a predictions using PROPKA with the energy-minimised homology model of heparanase.

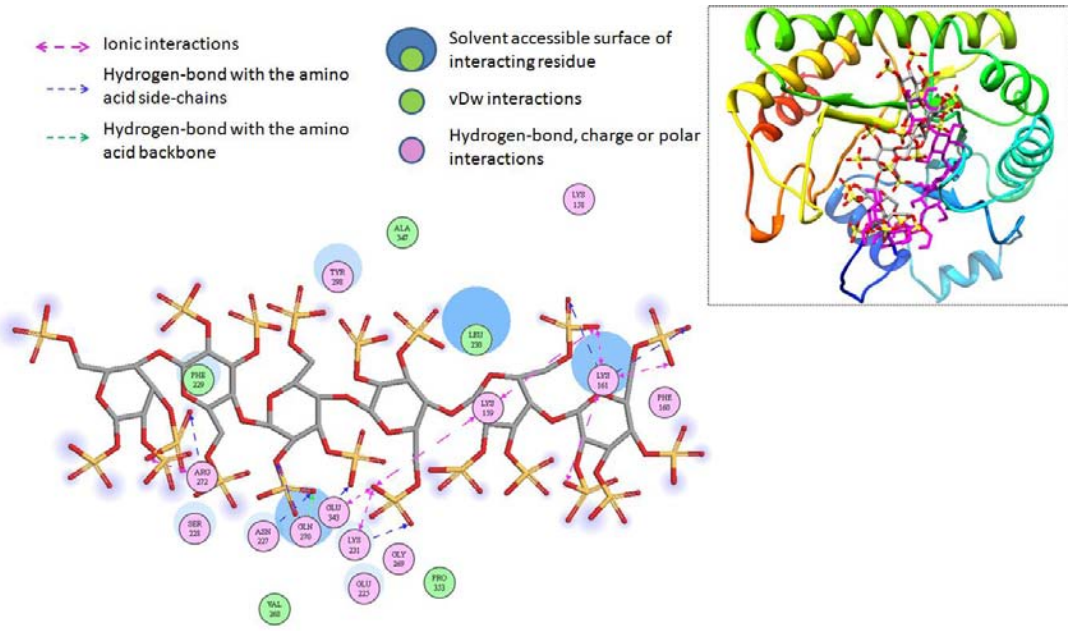


Fig. 8. Docking predictions for the MHS with the catalytic binding sites and heparin-binding domains HBD-1 and HBD-2 in heparanase. The superimposition of maltohexaose (in magenta) with MHS is shown in the inset, wherein maltohexaose adopts a "U-shaped kinked structure" and MHS has a different binding mode. The heparanase 128–430 domains are represented as ribbons.

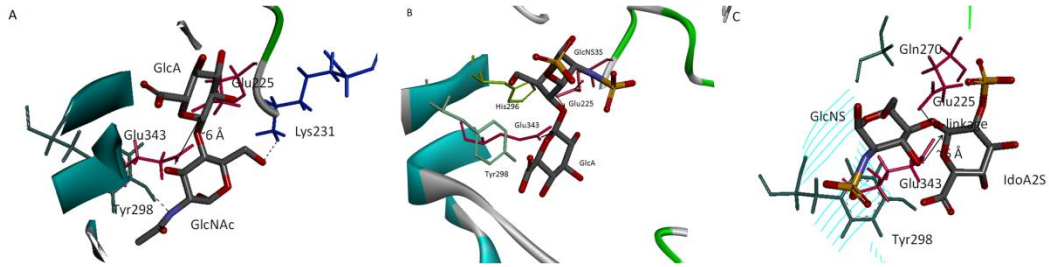


Fig. 9. Conformations of docked substrates (compounds 1 and 5) and inhibitor (compound 7) in the heparanase-binding site. Hydrogens are not shown for clarity. (A) Docking of compound 1 shows that the distance between the anomeric carbon of GlcA and the side chain of nucleophile Glu343 is greater than 6.0 Å, rendering compound 1 resistant to catalytic digestion. The hydroxyl groups of GlcNAc form hydrogen bonds with the side chains of Tyr298 and Lys231. (B) The top-ranked pose of compound 5 reveals that the distance between the oxygen of the glycosidic linkage and the protonated Glu225 is 3.8 Å and that the distance between the anomeric carbon and Glu343 is 3.9 Å. This indicates that heparanase can cleave this substrate. (C) Docking of inhibitor 7 shows that IdoA2S in the skew-boat conformation fits in the cleavage site with GlcNS being placed at subsite +1; however, the distance between the glycosidic bond oxygen and Glu225 was measured to be >5 Å.

Molecular dynamics simulations

The homology model of heparanase was further refined using MD simulations, which also allowed an assessment of conformational changes in the HDB-1. After energy minimisation and equilibration, a 1.0 ns production run was performed and trajectory data was collected for analyses. The time-dependent profile of the potential energy was plotted (Figure not shown) wherein the potential energy decreases gradually from the start and then stabilises after 500 ps. The RMSD of backbone atoms (C, Cα and N) between the initial modelled structure and the final structure at the end of the MD simulation was ~1.5 Å. The Lys residues in the 158-160 range adopt a 310-helical conformation after 800ps, which matches the secondary structure prediction of the NPSA server (see above). Even after the MD refinement of heparanase model alone, steric clashes are still observed upon superimposition of the cellooctaose substrate discussed above. The persistence of this steric hindrance may cause a considerable deterioration in the accuracy of ligand-protein docking approaches. While a longer simulation may be necessary to sample the conformational change in the modelled enzyme, such an extended simulation without restraints might result in distortions of the catalytic binding site and changes to the rotamersic preferences of the nucleophile and proton donor residues located in the loop regions. Such problems may be rectified by local refinement of ligand-binding regions (in this case, the catalytic and heparin binding regions) in protein models using remotely related templates identified by threading or using induced-fit docking.

Binding site analysis and molecular docking

To dock substrates into the modelled structure of heparanase, FINDSITE was used to identify the putative ligand binding pose and conserved anchor region geometries based on binding-site similarity across groups of weakly homologous template structures identified by threading. The results of FINDSITE are provided as data files in the Supporting Information.[2] The top-ranked binding sites included the catalytic residues and heparin binding domains HBD-1 and HBD-2. In most of the cases, the ligands were hexapyranoses with β-(1,4)-linkages and belong to the anti-protonation class of retaining enzymes. Superimposition of heparanase with the holo-enzymes indicates that family 79 enzymes such as heparanase are anti-protonators. In addition, the two catalytic residues are separated by a distance of 4.87 Å, which is consistent with hydrolysis occurring *via* a retention mechanism.

[2] The file findsitepocket.rtf contains information on binding pockets identified by the spatial clustering of template-bound ligands aligned to the target crystal structure and ranked by the number of binding ligands. The file findsitefunction.rtf contains function annotation (GO ontology) of heparanase using the templates selected by structure similarity. The file findsitealignments.rtf has template structures that have been selected from the template library by structure alignments to the native heparanase model.

The results of FINDSITE also suggest the presence of β-D-glucose rings at subsite -1 (cleavage occurs between subsites −1 and +1, with non-reducing ends and reducing ends being labelled with −n and +n signs, respectively (Davies, G.J., Wilson, K.S., et al. 1997)). To start docking, the last non-reducing ring of each substrate described in Table I was superimposed with the average position of the pyranosyl rings at the cleavage site (near subsite -1) of the substrates derived from the results of FINDSITE.

Apart from hexapyranose rings and furanose rings, core structures containing ring systems such as piperidin, morpholin, pyridine and indole rings were mapped near the cleavage site, according the FINDSITE analysis. Ligands such as cellooligosaccharides and their sulphur (thio) substituted glycosidic linkage, and GlcNAc covered the rest of the subsites from -4 to +4. Thio-linked oligosaccharides, wherein at least one glycosidic bond oxygen atom is replaced with a sulphur atom (PDB codes 1E5J (Fort, S., Varrot, A., et al. 2001), 1H5V (Varrot, A., Schulein, M., et al. 2001), 2O9R (Isorna, P., Polaina, J., et al. 2007) and 3CUI), are well established substrate mimics and competitive inhibitors of glycosidases (Driguez, H. 2001), and one such potential inhibitor of heparanse, the *S*-linked trisaccharide, is already known (Cao, H. and Yu, B. 2005).

The mapping of molecules into the various subsites of heparanase might help to establish if such molecules are good substrates of heparanase and, if so, whether the introduction of side chain substituents such as sulphone or phosphate groups on core structures and/or changes in the substituents at the glycosidic linkage (β v/s α) can be employed to search for carbohydrate-based inhibitors of GAG-heparanase interactions. For example, one of the ligands bridging HBD-1 and HBD-2 (site 18 of FINDSITE analyses, PDB code 3K8L (Koropatkin, N.M. and Smith, T.J. 2010)) is maltohexaose/maltoheptaose (α-1,4 linked glucose), whereas its polysulphated analogue **11** (MHS; maltohexaose sulphate with at least 3 sulphate groups per internal sugar and up to four sulphates in the terminal sugar residues) is already known to be a good inhibitor of heparanase (IC_{50} = 5 μg/mL compared to heparin with IC50 = 2 μg/mL) (Parish, C.R., Freeman, C., et al. 1999). MHS was docked onto the heparanase catalytic binding site and heparin binding domains and its various interactions are shown in Figure 8. Residues Lys159 and Lys161 of HBD-1 and Arg272 of HBD-2 make electrostatic interactions with the sulphates, whereas Gln270 is involved in hydrogen bonding with one of the sulphates of MHS. The majority of docked poses showed similar binding with Tyr298 through hydrogen bonding at a distance of 3.0 Å. Several interactions are made with the loop, which consists of residues Asn227-Leu230. Phe229 and Leu230 are involved in hydrophobic interactions, the side-chain of Asn227 makes hydrogen bonds with the sulphates and the side chain and backbone of Lys231 make electrostatic interactions and hydrogen bonds with the side chains of MHS. In general, maltohexaose adopts a "U-shaped kinked structure"; however, superimposition of the docked MHS and maltohexaose (PDB: 3K8L) indicates the occurrence of a different binding mode. This binding mode of MHS blocks the catalytic site as well as bridges both HBD-1 and HBD-2. This validates our hypothesis that the addition of sulphated substituents to the core structure can be used in the rational design of new GAG mimetics.

Manual docking of compound **1** has been reported previously (Ishida, K., Hirai, G., et al. 2004). The disaccharide unit in compound **1** in the absence of sulphation is known to be resistant to catalytic digestion (Okada, Y., Yamada, S., et al. 2002, Peterson, S.B. and Liu, J. 2010). The docking simulations reported here revealed that cluster 1, with the majority of the binding poses, has GlcA in subsite +1 while GlcNAc is docked near cleavage site (Figure 9A), whereas an inverse binding mode was reported in a previous study (Ishida, K., Hirai, G., et al. 2004). As a consequence, the distance between the anomeric carbon of GlcA and the side-chain of nucleophile Glu343 is greater than 6.0 Å, which might render compound resistant to catalytic digestion. The hydroxyl groups of GlcNAc form hydrogen bonds with the side chains of Tyr298 and Lys231. The second largest cluster with 14 poses has GlcA placed in subsite -1 and the glycosidic linkage is located further away from the catalytic residues.

Similarly, 6-*O*-sulpho heparosan (**2**) and *N*-sulpho heparosan (**3**) are resistant to cleavage by heparanase (Peterson, S.B. and Liu, J. 2010, Pikas, D.S., Li, J.P., et al. 1998). Clustering of the docked poses of compound **2** indicates that the catalytic residues are at a distance of 5.8 Å from the glycosidic linkage and, hence, the substrates cannot be cleaved. The GlcA residue is not exactly placed at the subsite -1. The 6-*O*-sulphate of compound **2** is involved in hydrogen bonding with Gln270, whereas the carboxylic acid of GlcA interacts with the charged side chain of Lys159 and with Tyr298. The adjacent hydroxyl groups (at C2 and C4) of GlcA are within hydrogen bonding distance from Glu343 and Glu225. The average binding energies for compound **3** in most of the clusters were

positive. The GlcA residue is docked in subsite +2 in the vicinity of Gln270, while the *N*-sulphate makes electrostatic interactions and a hydrogen bond with Lys159 and Tyr298, respectively. The NH of GlcNS forms a hydrogen bond with Glu 343.

Both *N*- and *O*-sulphations (6-*O*-sulphation or 3-*O*-sulphation) are required for cleavage by heparanase, as shown by experimental data for substrates **4** and **5**. In the top-ranked pose of compound **4**, the distance between the oxygen of the glycosidic linkage and the protonated Glu225 is 4.0 Å and the distance between the anomeric carbon and Glu343 is 4.2 Å**.** The *N*-sulphate of compound **4** interacts with the charged side chain of Lys231 while the *O*-sulphate interacts with the side chain of Gln270. The hydroxyl group at position 3 of GlcA of compound **4** is within hydrogen bonding distance of the side chain of Tyr296. Docking of compound **5** revealed the formation of a hydrogen bond between the carboxylic acid of GlcA and Tyr298, whereas the 6-*O*-sulphate and the pyranose oxygen form hydrogen bonds with Gln270. In top ranked pose of compound **5**, the distance between the oxygen of the glycosidic linkage and the protonated Glu225 is 3.8 Å and the distance between the anomeric carbon and Glu343 is 3.9Å (Figure 9B). The most populated clusters upon docking compounds **4** and **5** show that the GlcA residue is in the plane of Glu225 and that the C1-O5 bond is 'anti' to the catalytic residue. In the case of substituted GlcA (compound **6**), the docking simulations indicate that GlcA2S is located towards site 2, whereas GlcNS is between the cleavage site and subsite -1. These findings are consistent with experimental data showing that heparanase does not cleave the linkage of GlcA2S-GlcNS directly, but rather cleaves the linkage of the neighbouring GlcA-GlcNS.(Peterson, S.B. and Liu, J. 2010) Electrostatic interactions between the 2-*O*-sulphate of GlcA2S of substrate **6** and the side chain of Lys 159 and between the *N*-sulphate of GlcNS and His 296 are also observed. The hydrophobic Phe 160 is in close proximity to the glycosidic linkage of substrate **6**. The hydroxyl groups in GlcNS form hydrogen bonds with the side chain of Glu225. Hydrogen bonds are also observed between the amide of GlcNS and Gly269 and between the sulphate and backbone of Tyr298.

Sulphated iduronic acid (IdoA2S) adopts either the $^{1}C_{4}$ chair conformation or the skew-boat $^{2}S_{0}$ conformation (Mulloy, B., Forster, M.J., et al. 1993). Therefore, docking simulations of compound **7** were performed with IdoA2S using both conformations. The IdoA2S in the skew-boat conformation was found to fit at the cleavage site with GlcNS being placed at subsite +1; however, the distance between the glycosidic bond oxygen and Glu225 was measured to be > 5 Å. In the chair conformation, the IdoA2S was found to sit at subsite +1. The 2-*O*-sulphate of IdoA2S interacts through electrostatic interactions and hydrogen bonds with the side chains of Arg272 and Gln270, respectively. The amide group makes a hydrogen bond with the side chain of Tyr298. An electrostatic interaction between Lys159 and the *N*-sulphate group is observed, whereas the hydroxyl groups of GlcNS form hydrogen bonds with Glu343 and Lys231. The computed free energy of interaction of this disaccharide is lower with the chair conformation of IdoA2S compared to that of the boat conformation (Table IV). The computed free energies of interaction and the cluster analyses show that the chair conformation of IdoA2S might be the preferred form adopted in the heparanase binding site. The α-linkage containing disaccharide is further away from catalytic residues and hence might be resistant to cleavage (Figure 9C); however, the docking simulations performed are not able to explain the mechanism of action of compound **7** as an inhibitor on the basis of the interactions of sulphates alone. The sequence preceding and proceeding this disaccharide unit might also influence its behaviour as an inhibitor or substrate. However, taken together, the results from FINDSITE discussed above and the docking of maltohexaose and compound **7** suggest that it is possible that the conformational change in the linkage (α in the IdoA2S disaccharide versus β in the GlcA disaccharide) might be one of the factors that determine the behaviour of these molecules as inhibitors.

The docking simulations of substrates listed in Table IV suggest that heparanase recognises the *N*- and *O*-sulphated glucosamine at subsite +1 and interacts with GlcA at the cleavage site (near to subsite -1). In the absence of 6-*O*-sulphation in glucosamine, GlcA is found at subsite +2, which is in agreement with experiments (Peterson, S.B. and Liu, J. 2010).

The top ranked cluster of **8** shows that GlcA is docked at the cleavage site i.e. the distance between oxygen of the glycosidic linkage and protonated Glu225 is 3.4 Å and the distance between anomeric carbon and Glu343 is 4.36 Å. As for free energy of binding, heparanase showed the highest activity against the trisaccharide (Table IV). The 6-*O*-sulphate and *N*-sulphate of GlcNS6S at subsite +1 forms hydrogen bond with the backbone of Tyr298 and with side-chain of Gln270, respectively at a distance of 3.5 Å. No ionic interactions with the

COO^- group of GlcA were observed. The *N*-sulphate of GlcNS6S at subsite +1 forms hydrogen bond with the side-chain of Tyr298 and the amide forms hydrogen bond with the backbone oxygen of Gly389. Although 6-*O*-sulphate at non-reducing end is in close proximity of HBD-1, we did not observe any ionic interactions at 3.5 Å. Docking simulations of compound **8** reveal that, while the trisaccharide is essential for recognition by heparanase, the binding site is large enough to accommodate another sugar residue such as ΔHexUA at subsites -2 or +2. Consequently a tetrasaccharide containing the above sequence would be an ideal substrate.

The optimal length of a heparin chain required to bridge HBD-1 and HBD-2 for the effective inhibition is known to be an octadecasaccharide (Simizu, S., Ishida, K., et al. 2004a, Vlodavsky, I., Ilan, N., et al. 2007), whereas heparin tetra and hexasaccharides as well as *N*-acetylated heparins (*N*-acetylation higher than 50%) are poor inhibitors of heparanase (Bar-Ner, M., Eldor, A., et al. 1987, Naggi, A., Casu, B., et al. 2005). Nonetheless, the pentasaccharide AGA*IA (compound **9**), corresponding to the heparin/HS sequence that interacts with the active site in anti-thrombin III (AT), is susceptible to cleavage by heparanase (Bisio, A., Mantegazza, A., et al. 2007). 2D NMR determinations revealed the presence of a strong interaction of this pentasaccharide with HBD-1 (Levy-Adam, F., Abboud-Jarrous, G., et al. 2005). The docking simulations reported here revealed that the GlcA is positioned at subsite -1 and the 6-*O* sulphate of GlcNS6S preceding GlcA make electrostatic interactions with Lys158 and Lys161, whereas the *N*-sulphate of GlcNS6S interacts with the side chain of Lys231. The 3-*O*-sulphate of GlcNS3S6S is not directly involved in electrostatic interactions whereas the 6-*O*-sulphate makes a hydrogen bond with Tyr298. The 6-*O*-sulphate of GlcNS6S at the reducing end makes electrostatic interactions with the side chains of Arg272 and Gln270 in HBD-2. In some of the predicted binding modes, the carboxylic acid of GlcA makes an electrostatic interaction with Lys 161. The catalytic residues were found to be at a distance of ~4.5 Å from the anomeric carbon and glycosidic linkage and, hence, make the oligosaccharide susceptible to cleavage. The 2-*O*-sulphate of IdoA2S was not involved in any interaction, which explains the experimental observation that the desulphation of the IdoA residue has very little effect on the inhibitory activity of heparin (Naggi, A., Casu, B., et al. 2005).

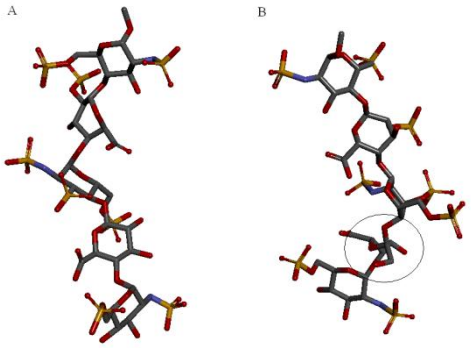


Fig. 10. Conformations of docked pentasaccharides in the binding site of heparanase: (A) AGA*IA, wherein the sulfates alternate in each disaccharide and (B) the glycol-split derivative of AGA*IA. The glycol-split residue is highlighted with a circle. Glycol-splitting introduces conformational changes in the linkage between GlcA and GlcNS3S6S and the A*IA trisaccharide, such that the sulfates reside on the same side of the GAG chain. These sulfates form close interactions with the enzyme.

The sequence AGA*IA (compound **9**) has anti-coagulant activity. Consequently, heparanase inhibitors need to be non-anticoagulant heparin mimics, with anti-metastatic and anti-angiogenic activity and without favouring the release and activation of growth factors. Examples of such molecules are acetylated heparins in their glycol-split forms (which involves splitting of C2-C3 bonds of non-sulphated uronic acid). The *N*-acetylated, glycol-split heparin such as ^{100}NA,RO.H inhibit heparanase activity at nanomolar concentrations and have no anti-coagulant activity (Naggi, A., Casu, B., et al. 2005, Nasser, N.J., Avivi, A., et al. 2007). In this study, a glycol-split was introduced in the AGA*IA sequence (henceforth referred to as gs-AGA*IA). In principle, the sugar units A*IA of compound **10** should occupy similar positions in the catalytic binding site as the natural pentasaccharide. Docking simulations revealed the presence of conformations that can be grouped into two clusters and are different from that of AGA*IA bound to heparanase. The difference in both ligand conformations lies in the interaction of the terminal end residue. The *N*-sulphate in one conformation interacts with HDB-1 whereas in the other it interacts with HBD-2. Hence, this suggests that at least one *N*-sulphate group is required to inhibit heparanase activity as effectively as the corresponding *N*-acetylated glycol split derivatives (Naggi, A., Casu, B., et al. 2005, Nasser, N.J., Avivi, A., et al. 2007). Furthermore, in both clusters the 2-*O*-sulphates of IdoA2S form hydrogen bonds with the side

chain of Ser228 and with the backbones of Phe229 and Ala230. The glycol-split GlcA modifies the glycosidic linkage between unsulphated GlcA and GlcNS3S6S. This conformational change is propagated to the remaining three sugar residues in the pentasaccharide. The *N*-sulphate, 2-*O*-sulphate and 6-*O*-sulphate are alternately arranged on each side of the AGA*IA pentasaccharide (Figure 10A) or the *N*-sulphated heparin structure, while glycol-splitting introduces a "kink" and generates novel structures where the *O*-sulphate groups of adjacent disaccharide units reside on the same side of the GAG chain and are involved in the binding to the enzyme (Figure 10B). These binding modes are in agreement with the previously proposed hypothesis of binding of the glycol-split oligosaccharides to the active site of heparanase (Naggi, A., Casu, B., et al. 2005, Nasser, N.J., Avivi, A., et al. 2007).

PI-88 (compound **12**) is a mixture of highly sulphated oligosaccharides, ranging from disaccharides to hexasaccharides, with the majority (60%) being pentasaccharides, and which inhibits angiogenesis and metastasis by virtue of its inhibition of heparanase (Yu, G., Gunay, N.S., et al. 2002). Furthermore, structure-activity relationship (SAR) studies have indicated that tetrasaccharides exhibit comparable activity to that of pentasaccharides in cell-based and *ex vivo* assays of angiogenesis (Johnstone, K.D., Karoli, T., et al. 2010). The binding modes of pentasaccharides of PI-88 (compound **12**) and its analogue (compound **13**, which is hydrophobic aromatic in nature) were investigated. Docking of PI-88 and compound **13** indicates that the pentasaccharide is able to bridge HBD-1 and HBD-2. The docked poses can be clustered into two binding modes. In the majority of binding modes, the sulphates from both molecules make electrostatic interactions with the side chains of Lys159, Lys161 and Arg272, whereas hydrogen bonds were made with the side chains of Gln270, Tyr298 and Ser228. No interactions were observed with the phosphate group of PI-88. The representative docked poses of compound **13** from both the clusters are shown in Figure S3 in the Supporting Information. The CH_2-Ph group fits in the hydrophobic pocket formed by Ala352 and neighbouring residues Gly350 and Pro353. On the other hand, in the second cluster, the CH_2-Ph ring is oriented towards Phe229. Docking simulations cannot reliably rank this class of ligands as the differences in the predicted binding affinities of PI-88 analogues are not very large (Johnstone, K.D., Karoli, T., et al. 2010). Hence a representative ligand (compound **13)** was used to predict the binding mode of similar classes of compounds.

Unlike α-L-iduronic acid, which may have $^1C_4$, $^4C_1$ and $^2S_0$ conformations depending on the number of sulphate groups, interactions with metal ions and the position of the sugar in the oligosaccharide, β-D-GlcA exists predominantly in the $^4C_1$ conformation (Nielsen, J.E. and McCammon, J.A. 2003a). The docking of various substrates consisting of GlcA and the superimposition of GlcA based on cellooligosaccharides suggest that the sugar ring is in the $^4C_1$ conformation, the position of the proton donor Glu225 is within hydrogen-bonding distance of the glycosidic bond oxygen and residue Glu343 is within contact distance to facilitate the nucleophilic attack on the anomeric carbon in subsite -1. Based on the docking of substrates and inhibitors, the structural features of the active site of heparanase are proposed in Figure S4 in the Supporting Information. Besides catalytic residues, which act as proton donor and acceptor, and the heparin binding domains, which consist of basic residues, the active site has a hydrophobic/aromatic patch formed by residues Phe129 and Leu230, and another similar patch formed by residues Phe160, Ala347, Ala352, Val384 and Ala388. Various hydrophobic groups could target these hydrophobic regions in the rational design of heparanase inhibitors. The docking of substrates and inhibitors indicates the presence of a large binding site extending over at least two saccharide units beyond the cleavage site (towards the non-reducing end) and at least three saccharides towards the reducing end (towards HBD-2).

In the work reported in this paper, the inter-glycosidic torsion angles of disaccharide substrates bound to heparanase were measured and compared with the corresponding torsion angles determined from experimental NMR and X-ray crystallographic determinations as well as MD simulations (Table S4). The φ angles for molecules **2**, **3** and **6** and the ψ angles for substrate **6** are inconsistent with the data from related structures, suggesting that AutoDock has a marked tendency to distort inter-glycosidic torsion angles away from normally expected values. MD simulations of free molecules such as cellulose, chitin, mannan, xylan, and hyaluronan in the presence of explicit solvent have been used to correct the glycosidic torsions around β(1→4) (Almond, A. and Sheehan, J.K. 2003), which is the linkage found in most of the hexapyranoses studied here. However, MD simulations with unrestrained β-D-glucuronic acid demonstrated that the force fields are unable to reproduce the

experimentally observed prevalence of the $^4C_1$ chair conformation (Sattelle, B.M. and Almond, A. 2010). In such scenario, it seems appropriate to select docked poses on the basis of favourable torsion angles in a cluster (Tables IV and S4) at the expense of those with the best free energies (and the rank of the cluster). As seen from experimental data for GlcA oligosaccharides (Table S4), the glycosidic torsions vary with the chain length. Furthermore, sulphation on glucosamine can change the torsion angles of the disaccharides. Therefore, the conformations of molecules **2** and **3** are reasonable. The docked poses of substrates **6** and **7** do not follow the expected conformational preferences and hence the predicted interactions of these molecules should not be taken into consideration for rational drug design.

The docking results reported here also suggest the presence of an induced-fit mechanism in the binding of substrates to the catalytic and HBDs of heparanase. The existence of such induced fit in the enzyme has been proposed in cel5 endoglucanase (PDB code 1CEN) from *Clostridium thermocellum* (Dominguez, R., Souchon, H., et al. 1996) and family 51 arabinofuranosidase (PDB codes 2VRQ, and 2VRK) from *Thermobacillus xylanilyticus* (Paes, G., Skov, L.K., et al. 2008). Furthermore, the catalytic site consisting of the 224NEP226 motif is a loop which might impart some flexibility and impact substrate binding. In addition, increased flexibility is conferred near the HBD-1 domain by cleavage of the inter-chain disulphide bond Cys127-Cys179.

The His residue forms part of the catalytic triad in clan A family, while the second residue is a strictly conserved Glu (proton donor), but variations (either Ser, Thr, Glu, Asp or water) are observed in the third residue of the triad (Debeche, T., Bliard, C., et al. 2002, Nielsen, J.E. and McCammon, J.A. 2003b). The H296XY298 sequence (where X is any residue) in heparanases is also present in endoglucanases or cellulose-binding proteins and in α-L-arabinofuranosidases and xylanases, but structural superimposition with various clan A members did not reveal the presence of any conserved or semi-conserved residues that can form a triad in heparanase. Furthermore, docking simulations with substrates and analyses with FINDSITE with various ligands indicate that Tyr298 is involved in direct interactions with the substrates, whereas His296, although conserved, is not involved in any direct hydrogen bonding with the substrate. There are Ser, Thr, Asp and Glu residues in the vicinity of the catalytic binding site but at a distance not sufficient to form hydrogen bonds with the His or Glu residues of the classic triad. Therefore, the Glu225-His296 dyad is suitably positioned to lower the pK_a of the acid/base residue and has a common catalytic mechanism that involves two conserved acidic residues, a putative proton donor at Glu225 and a nucleophile at Glu343.

Previous molecular modelling studies have identified residues Lys139, Val170, Asp171, Thr175, Arg307, Trp340 and Tyr264 as important for the binding of small molecule inhibitors of heparanase (Zhou, Z., Bates, M., et al. 2006). However, no mutagenesis data is available to validate these predictions. No interactions of these residues with carbohydrate-based ligands were observed in this docking study. NMR and SPR (surface plasmon resonance) methods identified residues Glu225, Glu343, Lys158, Lys161, Arg272 and Thr275 as important for the binding of poly-sulphonated ligands such as suramin (Mosulen, S., Orti, L., et al. 2011). Molecular modelling studies found Phe 229 in the binding site to be also important for activity of KI-105 (Ishida, K., Wierzba, M.K., et al. 2004). The docking simulations reported here agree with these experimental observations (Ishida, K., Wierzba, M.K., et al. 2004, Mosulen, S., Orti, L., et al. 2011).

Conclusions

Homology modelling together with the Conseq and FINDSITE analyses lends insight into the structure/function relationships of heparanase. In contrast to the apparent structural similarities between families 5, 10, 30, 39 and 51 and heparanase, structural links between the substrate binding clefts of family 44 and endoglucuronidase were established. The binding of β-glucopyranosides substrates to glycosyl hydrolase 79 heparanase has been investigated using ligand-protein docking simulations to increase our understanding of substrate specificity of heparanase, wherein the substrate includes a GlcA and a GlcNS unit carrying *O*-sulphation. Molecular docking and pK_a prediction calculations show that Glu225 acts as the proton donor in human heparanase and that the enzyme is anti-protonator. The distance between Glu225 and the glycosidic linkage of the substrate may determine the specificity according to the substrate docking model. We have proposed substrate recognition sites on heparanase based on docking studies. Various GAG-based inhibitors were docked to determine which amino acids in the protein interact with these sulphated sugar molecules. Apart from electrostatic interactions with the heparin binding domains, hydrophobic

interactions also contribute to the increase in binding affinity of several GAG-based inhibitors. Several interactions of ligands such as PI-88 analogues, which contain hydrophobic linkers, can be observed in the active site formed by residues Phe129 and Leu230. Consequently, future mutagenesis studies on residues 227-230 in the catalytic active site will aid the rational design of heparanase inhibitors. These docking ligand-protein complex models can interpret the substrate specificity of heparanase, providing a rationale for the design of polysaccharides (e.g. IdoA2S at subsite -1) that may act as inhibitors of the enzymatic activity of heparanase. Predicted heparin/enzyme complexes show that the interactions of the heparin binding domains in combination with the catalytic domain can be targeted for the design of inhibitors. These findings also emphasize the challenges in the use of molecular modelling methods for the development of new inhibitory non-anti-coagulant molecules against the enzymatic functions of heparanase.

Materials and methods

ConSeq analysis

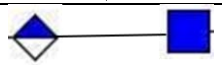
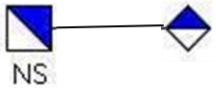
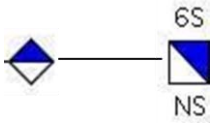
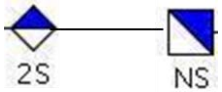
Given a sequence and an alignment of homologous sequences, the ConSeq (Berezin, C., Glaser, F., et al. 2004) server maps functionally important residues that take part in ligand binding and protein–protein interactions onto each amino acid in the primary sequence. ConSeq takes into account the evolutionary conservation and solvent accessibility profiles to indicate residues that have potential structural or functional importance. A PSI-BLAST (Altschul, S.F., Madden, T.L., et al. 1997) search was performed against the CLEAN_UNIPROT (filtered UNIPROT database in Consurf) (Goldenberg, O., Erez, E., et al. 2009) to search for non-redundant homologues. The query sequences correspond to the heparanase human sequence (Q9Y251). To map evolutionary conservation scores onto query sequence, a multiple sequence alignment (MSA) from clan A hydrolases (Table S1 in the Supporting Information) was calculated using the MAFFT (Katoh, K. and Toh, H. 2008) alignment algorithm implemented in the new ConSurf server (Ashkenazy, H., Erez, E., et al. 2010). The sequences, species, SWISS-PROT (Boeckmann, B., Bairoch, A., et al. 2003) accession numbers and sequence identity with reference to human heparanase are listed in Table S1.

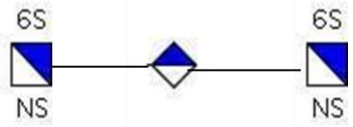
The Rate4Site algorithm (Pupko, T., Bell, R.E., et al. 2002) was subsequently used to construct a phylogenetic tree using the neighbour joining algorithm and to calculate evolutionary conservation scores. An empirical Bayesian approach was used to calculate the evolutionary rate of each amino acid position of the MSA, taking into account the stochastic nature of the evolutionary process. Amino acid evolution is traced using the JTT substitution model (Jones, D.T., Taylor, W.R., et al. 1992). The conservation codes were projected on the heparanase sequence and MSA, and colour coded, where 1 corresponds to maximal variability and 9 to maximal conservation. The phylogenetic tree produced by ConSeq was plotted using FigTree 1.2.2 (Rambaut, A. 2006b).

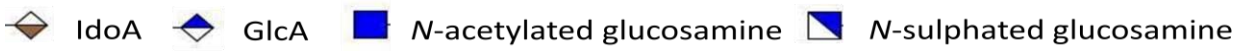
Homology modelling

The secondary structure was predicted using PSIPRED. The secondary structure consensus prediction program (NPSA; Network protein sequence analysis) (Combet, C., Blanchet, C., et al. 2000), which combines the DSC (Discrimination of protein Secondary structure Class) (King, R.D. and Sternberg, M.J. 1996), MLRC (Guermeur, Y., Geourjon, C., et al. 1999), PHD (Profile network from HeiDelberg) (Rost, B. and Sander, C. 1993, Rost, B. and Sander, C. 1994) and Predator (Frishman, D. and Argos, P. 1996) methods, was used to generate a secondary structure consensus and validate the prediction. As stated earlier, heparanase consists of 543 amino acids that can be structured into three domains, encompassing the 157 N-terminal regions, the catalytic domain and a C-terminal domain Val418-Ile543. The N-terminal domain is crucial for expression and activity (Hulett, M.D., Freeman, C., et al. 1999), while the hydrophobic C-terminal domain, consisting of disulphide bridge Cys437-Cys542, is critical for heparanase activity and secretion (Simizu, S., Suzuki, T., et al. 2007), as well as mediating non-enzymatic functions of heparanase, facilitating Akt phosphorylation, cell proliferation and tumour xenograft progression (Fux, L., Feibish, N., et al. 2009, Lai, N.S., Simizu, S., et al. 2008). Recently, the region encompassing residues 158-417, which includes the GAG binding site, was found suitable for drug discovery purposes in the search for novel inhibitors of heparanase (Mosulen, S., Orti, L., et al. 2011). We focused on regions from 158 to 543 for the purpose of binding site refinement and docking studies.

Table IV. Summary of binding modes and free energies of binding of various heparin disaccharides. The architecture of the substrate/inhibitor binding sites (subsites) adjacent to the cleavage site is shown. The subsites are labelled with negative numbers for subsites to the right (non-reducing end) and positive numbers to the left (reducing end) of the catalytic site.

Substrate /inhibitor	*Possible binding mode (subsites) based on docking* -3 -2 -1 +1 +2 +3 cleavage	*Cluster No.*	*Number of docked poses in the cluster*	*Mean free energy of binding in the cluster (kcal/mol)*	*Free energy (kcal/mol) of the top ranking pose in the cluster*
1		1 6	22 14	-1.10 -1.45	-2.24 -1.87
2		1 2	11 187	-1.07 -0.81	-1.52 -1.35
3		7 8	45 19	+1.29 +0.35	-0.15 -0.14
4		1	19	-0.49	-1.19
5		3	20	+0.37	-0.16
6		3 20	124 8	+1.77 +6.00	+0.36 +4.87

7 (boat conformation)		1	19	+0.16	-1.08
7 (chair conformation)		1	34	-0.84	-2.03
8		1	39	-0.08	-3.80



Structure construction, assignment of disulphide bridges, optimisation, and visualisation were performed using the molecular modelling package Discovery Studio 2.5 (Accelrys, Inc.). Ten models were constructed and loops were built using the discrete optimised protein energy (DOPE) (Shen, M.Y. and Sali, A. 2006) loop-modelling protocol in MODELLER (Sali, A. and Blundell, T.L. 1993). Essential hydrogen and charges were added to the structure. After coarse energy minimisation, the model with the lowest probability density function (PDF) energy was assessed for its overall geometric and stereochemical quality using a Ramachandran Plot and the Verify3D server (Bowie, J.U., Luthy, R., et al. 1991, Luthy, R., Bowie, J.U., et al. 1992).

At physiological pH or above, heparanase is inactive, does not cleave HS but binds to the ECM or cell-surface HSPGs, thereby facilitating inflammatory leukocyte adhesion (Gilat, D., Hershkoviz, R., et al. 1995, Ihrcke, N.S., Parker, W., et al. 1998). Heparanase is active in the pH range 4.0 to 7.5, with maximal activity at pH 5.5 to 5.8 (Freeman, C. and Parish, C.R. 1998, Gilat, D., Hershkoviz, R., et al. 1995, Graham, L.D. and Underwood, P.A. 1996), cleaving HSPGs. Consequently, all ionisable residues were protonated depending on their pK_a using the H++ (Gordon, J.C., Myers, J.B., et al. 2005) web-server, based on finite difference Poisson-Boltzmann calculations, and PropKa 3.0 (Olsson, M., Søndergaard, C., et al. 2011). The model from PropKa was further evaluated for flexibility by molecular dynamics (MD) simulations, while the reliability of the model was evaluated by docking and binding analyses.

Electrostatic potential surface calculations

Electrostatic potential surface calculations were done using the DelPhi program (Sharp, K.A., Nicholls, A., et al. 1998) implemented in DS Modelling 2.1 (Accelrys, inc.) using the default Delphi atomic partial charges and radii, with a protein interior dielectric constant of 1, a solvent dielectric constant of 80 and an ionic strength of 0.145 M.

Molecular dynamics simulations

All energy minimisations and MD simulations were performed with AMBER 9.0 (Case, D.A., Cheatham Iii, T.E., et al. 2005). The Parm99SB (Viktor, H., Robert, A., et al. 2006) force-field in AMBER 9.0 was used for the simulations of heparanase. A cubic box of TIP3P water molecules (Jorgensen, W.L., Chandrasekhar, J., et al. 1983) was added to solvate the protein, keeping a minimum distance of 12.0 Å between each face of the box and the protein. The number of water molecules added to the heparanase structure was 19632. Net charges in the protein were neutralised by adding an appropriate number of chloride ions.

In each simulation initial unfavourable contacts with the solvent were removed by energy minimisation after performing 10 steps of steepest descents followed by 990 steps of conjugate gradients, keeping the protein rigid. The system was energy minimised again as before allowing the side chains and then the entire structure to relax. A force constant of 25 kcal/mol $Å^2$ was used to restrain the atoms of the system that were not being relaxed. Three stages of MD equilibration were performed to relax water and hydrogen atoms (200 ps), side chains (200 ps) and the whole system (500 ps). Finally, a 1.0 ns production simulation was performed. A timestep of 2.0 fs was used in all simulations and the coordinates were saved every 10 ps. During all simulations the particle mesh Ewald (PME) method (Tom, D., Darrin, Y., et al. 1993) was used to compute long range electrostatic interactions, using a 1.0 Å grid spacing and a fourth-order spline for interpolation. The non-bonded cutoff was set to 10.0 Å and the SHAKE algorithm (Ryckaert, J.-P., Ciccotti, G., et al. 1977) was used to constrain all bonds involving hydrogen atoms. All simulations were carried out in the isobaric-isothermal (NPT) ensemble at a pressure of 1 atm. The temperature and pressure were kept constant using the weak-coupling algorithm with coupling constants τ_T and τ_P of 0.1 and 1 ps, respectively (Berendsen barostat) (Berendsen, H.J.C., Postma, J.P.M., et al. 1984). Periodic boundary conditions were applied throughout.

Binding site prediction

Ligand-binding residues were identified in the target structure using FINDSITE (Brylinski, M. and Skolnick, J. 2008, Brylinski, M. and Skolnick, J. 2009, Brylinski, M. and Skolnick, J. 2010, Skolnick, J. and Brylinski, M. 2009), a structure/evolution-based method for ligand binding site prediction and molecular function inference. FINDSITE detects common ligand-

binding sites in a set of evolutionarily related proteins with <35% sequence identity to the target. As the targets for local refinement, we used the best of top five binding sites predicted within 6 Å from the geometrical centre of a bound ligand in the reference crystal structure. FINDSITE also provides information on the chemical identity of molecules that likely occupy the predicted binding sites by performing virtual screening against the ZINC, NCI and other databases. This feature was not used in our work.

Molecular docking

The structures of substrates and inhibitors used in this work are listed in Table I. The backbone of each molecule was constructed using the building facility offered at the GLYCAM web-server (2005-2011), which provides the final coordinates in PDB format. The coordinates of the AGA*IA pentasaccharide were taken from the X-ray structure of the complex of anti-thrombin and Factor Xa (PDB code 2GD4) (Johnson, D.J., Li, W., et al. 2006). Structural modifications were carried out using Discovery Studio 3.0 (Accelrys, Inc.). We (Coombe, D.R., Stevenson, S.M., et al. 2008, Gandhi, N.S. and Mancera, R.L. 2011) and others have shown that AutoDock can be used to perform docking of carbohydrates to their protein targets and the scoring function can reasonably rank the relative binding affinities of carbohydrate ligands/substrates (Bitomsky, W. and Wade, R.C. 1999, Cantu, D., Nerinckx, W., et al. 2008, Coombe, D.R., Stevenson, S.M., et al. 2008, Coutinho, P.M., Dowd, M.K., et al. 1997, Coutinho, P.M., Dowd, M.K., et al. 1998, Gandhi, N.S., Coombe, D.R., et al. 2008, Hill, A.D. and Reilly, P.J. 2008, Mertz, B., Hill, A.D., et al. 2007, Mulakala, C. and Reilly, P.J. 2002, Takaoka, T., Mori, K., et al. 2007), although much more expensive molecular dynamics approaches have been used successfully for relatively small GAG fragments (Gandhi, N.S. and Mancera, R.L. 2009). The drawback of AutoDock is the need to reach a compromise between accuracy and efficiency. In particular, the number of energy evaluations for each run had to be set to allow some runs to produce docked conformers in a reasonable amount of computer time. This issue becomes especially critical with an increasing number of torsions in the ligand molecule (over 18 for a trisaccharide), since the additional torsions add several more degrees of freedom in the system. Hence, in this work we used two different approaches for docking ligands with at least 12 rotatable bonds. the molecular docking engine FRED (FRED 2008a) was first used to dock ligands in the active site followed by re-docking and scoring of the poses with the Lamarckian genetic algorithm (LGA) (Morris, G.M., Goodsell, D.S., et al. 1998) method in Autodock 4.2 (Morris, G.M., Huey, R., et al. 2009). During both stages, the ligand was allowed full flexibility whereas the protein was held rigid. While most of the docking programs can accommodate flexibility in the protein, this was not practical in our study as there are too many flexible basic residues in the binding site that interact with flexible ligands, such as GAGs. Furthermore, molecules such as heparin have a large number of torsional degrees of freedom, which can lead to the failure of the search method to find binding poses effectively.

The program OMEGA v.2.3.2 (FRED 2008b) was used to convert all molecules described into 3D multi-conformer structures and to add hydrogen atoms and partial charges using the MMFF force-field (Thomas, A.H. 1996a, Thomas, A.H. 1996b, Thomas, A.H. 1996c, Thomas, A.H. and Robert, B.N. 1996). The maximum number of conformers (maxconfs) was set to 50. OMEGA was used because of its ability to enumerate ring conformations and invertible nitrogen atoms. This method is useful in enumerating ring conformations in heparin, particularly sulphated α-L-iduronate (IdoA2S) residues, which may adopt skew boat ($^{2}S_{0}$) and chair ($^{1}C_{4}$) conformations (Ferro, D.R., Provasoli, A., et al. 1990). The multi-conformer database of one or more ligands generated by OMEGA was used as input to the molecular docking engine FRED (FRED 2008a). The active site of the protein was defined on the basis of the superimposed ligands predicted by FINDSITE, as described above. The "inner contour" (a shape complementary to the active site, used during the exhaustive search) was disabled as the substrate-binding sites of endo-hydrolases consist of clefts or sets of juxtaposed surface residues rather than pockets (Esko, J.D. and Linhardt, R.J. 2009). Distance constraints between the catalytic residues and the GlcA substrate were setup using the FRED receptor module. The number of poses (num_poses) to be returned by the exhaustive search was set to 5000, defined as the top scoring poses

selected from the list of all poses and scored by the scoring functions specified by the exhaustive scoring. The number of alternative poses (num_alt_poses) was set to 300, defined as additional poses to the top consensus structure poses. In the optimisation step four scoring functions were used: ChemGauss, ChemScore (Eldridge, M.D., Murray, C.W., et al. 1997), PLP (Gehlhaar, D.K., Verkhivker, G.M., et al. 1995) and ScreenScore (Stahl, M. and Rarey, M. 2001). Binding poses obtained from FRED using the consensus score were extracted and submitted for re-docking.

For re-docking, the global minimum structure and the low-energy structures of the significant –representative conformations of a given oligosaccharide were subjected to docking, but using the Lamarckian Genetic Algorithm (LGA) in AutoDock with a population of 200 individuals, 256 runs and $50x10^6$ energy evaluations, and permitting a maximal translation of 0.1 Å per step, followed again by cluster analysis. To reduce the number of degrees of freedom, the pyranosyl ring at subsite -1 was fixed but full flexibility of all rotatable dihedral angles was allowed. In all simulations, the mutation rate was set to 0.80, the crossover rate was 0.02, the maximal number of generations was 2.7×10^4, elitism was set to 1 and the local search frequency was 0.06. The grid box was defined with a constant grid spacing of 0.37 Å around each ligand molecule using the previously obtained binding pose. The AutoDock 4.0 scoring function (Huey, R., Morris, G.M., et al. 2007) was used to predict the free energies of ligands shorter than tetrasaccharides.

Most of the modelling studies (docking or MD simulations) described to date have been carried out using crystal structures of glycosyl hydrolases in complex with non-sulphated monosaccharide, disaccharide or trisaccharide substrates (Coutinho, P.M., Dowd, M.K., et al. 1997, Mertz, B., Hill, A.D., et al. 2007, Mulakala, C. and Reilly, P.J. 2002). The molecular mechanics based scoring functions fail to predict binding affinities for molecules such as heparin with a large number of degrees of freedom (translational, rotational and torsional). Hence, AutoDock scoring was not carried out for molecules bigger than a trisaccharide.

Visualisation

All the 3D protein and ligand structures were edited and visualised using Discovery Studio 3.0 (Accelrys, Inc.). Consurf results and MD simulation trajectories were visualised using UCSF Chimera (Eric, F.P., Thomas, D.G., et al. 2004).

Supplementary data

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Funding

N.S.G is grateful for the award of an Endeavour International Postgraduate Research Studentship.

Acknowledgment

NSG is grateful for the award of an Endeavour International Postgraduate Research Studentship. We gratefully acknowledge the Western Australian Interactive Virtual Environments Centre (IVEC) and the National Computational Infrastructure (NCI) Facility for access to high-performance computing.

Conflict of interest statement

None of the authors of the above manuscript has declared any conflict of interest.

Abbreviations

HSPG, heparan sulphate proteoglycans; AT, antithrombin III; IdoA2S, sulfated iduronic acid; LGA, Lamarckian genetic algorithm; MHS, maltohexaose sulfate; MSA, multiple sequence alignment; NMR, nuclear magnetic resonance.

References

Almond A, Sheehan JK. 2003. Predicting the molecular shape of polysaccharides from dynamic interactions with water. *Glycobiology*, 13:255-264.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acids Res*, 25:3389-3402.

Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. 2010. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucl Acids Res*, 38 Suppl:W529-W533.

Bame KJ. 2001. Heparanases: endoglycosidases that degrade heparan

sulfate proteoglycans. *Glycobiology*, 11:91R-98R.

Bar-Ner M, Eldor A, Wasserman L, Matzner Y, Cohen IR, Fuks Z, Vlodavsky I. 1987. Inhibition of heparanase-mediated degradation of extracellular matrix heparan sulfate by non-anticoagulant heparin species. *Blood*, 70:551-557.

Barash U, Cohen-Kaplan V, Arvatz G, Gingis-Velitski S, Levy-Adam F, Nativ O, Shemesh R, Ayalon-Sofer M, Ilan N, Vlodavsky I. 2010a. A novel human heparanase splice variant, T5, endowed with protumorigenic characteristics. *FASEB J*, 24:1239-1248.

Barash U, Cohen-Kaplan V, Dowek I, Sanderson RD, Ilan N, Vlodavsky I. 2010b. Proteoglycans in health and disease: new concepts for heparanase function in tumor progression and metastasis. *FEBS J*, 277:3890-3903.

Berendsen HJC, Postma JPM, Gunsteren WFv, DiNola A, Haak JR. 1984. Molecular dynamics with coupling to an external bath. *J Chem Phys*, 81:3684-3690.

Berezin C, Glaser F, Rosenberg J, Paz I, Pupko T, Fariselli P, Casadio R, Ben-Tal N. 2004. ConSeq: the identification of functionally and structurally important residues in protein sequences. *Bioinformatics*, 20:1322-1324.

Bisio A, Mantegazza A, Urso E, Naggi A, Torri G, Viskov C, Casu B. 2007. High-performance liquid chromatographic/mass spectrometric studies on the susceptibility of heparin species to cleavage by heparanase. *Semin Thromb Hemost*, 33:488-495.

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

Boeckmann B, Bairoch A, Apweiler R, Blatter M-C, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I*, et al.* 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucl Acids Res*, 31:365-370.

Bowie JU, Luthy R, Eisenberg D. 1991. A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, 253:164-170.

Brylinski M, Skolnick J. 2008. A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. *Proc Natl Acad Sci U S A*, 105:129-134.

Brylinski M, Skolnick J. 2009. FINDSITE: a threading-based approach to ligand homology modeling. *PLoS Comput Biol*, 5:e1000405.

Brylinski M, Skolnick J. 2010. Comparison of structure-based and threading-based approaches to protein functional annotation. *Proteins*, 78:118-134.

Cantu D, Nerinckx W, Reilly PJ. 2008. Theory and computation show that Asp463 is the catalytic proton donor in human endoplasmic reticulum alpha-(1-->2)-mannosidase I. *Carbohydr Res*, 343:2235-2242.

Cao H, Yu B. 2005. Synthesis of a S-linked heparan sulfate trisaccharide as the substrate mimic of heparanase. *Tetrahedron Lett*, 46:4337-4340.

Case DA, Cheatham Iii TE, Darden T, Gohlke H, Luo R, Merz Jr KM, Onufriev A, Simmerling C, Wang B, Woods RJ. 2005. The Amber biomolecular simulation programs. *J Comput Chem*, 26:1668-1688.

Casu B, Vlodavsky I, Sanderson RD. 2008. Non-anticoagulant heparins and inhibition of cancer. *Pathophysiol Haemost Thromb*, 36:195-203.

Combet C, Blanchet C, Geourjon C, Deleage G. 2000. NPS@: network protein sequence analysis. *Trends Biochem Sci*, 25:147-150.

Coombe DR, Stevenson SM, Kinnear BF, Gandhi NS, Mancera RL, Osmond RI, Kett WC. 2008. Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 2. Biochemical analyses. *Biochemistry*, 47:4863-4875.

Coutinho PM, Dowd MK, Reilly PJ. 1997. Automated docking of monosaccharide substrates and analogues and methyl alpha-acarviosinide in the glucoamylase active site. *Proteins*, 27:235-248.

Coutinho PM, Dowd MK, Reilly PJ. 1998. Automated docking of α-(1,4)- and α-(1,6)-linked llucosyl trisaccharides in the glucoamylase active site. *Ind Eng Chem Res*, 37:2148-2157.

Czjzek M, Ben David A, Bravman T, Shoham G, Henrissat B, Shoham Y. 2005. Enzyme-substrate complex structures of a GH39 beta-xylosidase from Geobacillus stearothermophilus. *J Mol Biol*, 353:838-846.

Davies G, Henrissat B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure*, 3:853-859.

Davies GJ, Wilson KS, Henrissat B. 1997. Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J*, 321 (Pt 2):557-559.

Davies MN, Toseland CP, Moss DS, Flower DR. 2006. Benchmarking pK(a) prediction. *BMC Biochem*, 7:18.

Debeche T, Bliard C, Debeire P, O'Donohue MJ. 2002. Probing the catalytically essential residues of the alpha-L-arabinofuranosidase from Thermobacillus xylanilyticus. *Protein Eng*, 15:21-28.

Dempsey LA, Plummer TB, Coombes SL, Platt JL. 2000. Heparanase expression in invasive trophoblasts and acute vascular damage. *Glycobiology*, 10:467-475.

Dominguez R, Souchon H, Lascombe M, Alzari PM. 1996. The crystal structure of a family 5 endoglucanase mutant in complexed and uncomplexed forms reveals an induced fit activation mechanism. *J Mol Biol*, 257:1042-1051.

Driguez H. 2001. Thiooligosaccharides as tools for structural biology. *Chembiochem*, 2:311-318.

Dvir H, Harel M, McCarthy AA, Toker L, Silman I, Futerman AH, Sussman JL. 2003. X-ray structure of human acid-beta-glucosidase, the defective enzyme in Gaucher disease. *EMBO Rep*, 4:704-709.

Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP. 1997. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J Comput Aided Mol Des*, 11:425-445.

Elkin M, Ilan N, Ishai-Michaeli R, Friedmann Y, Papo O, Pecker I, Vlodavsky I. 2001. Heparanase as mediator of angiogenesis: mode of action. *FASEB J*, 15:1661-1663.

Eric FP, Thomas DG, Conrad CH, Gregory SC, Daniel MG, Elaine CM, Thomas EF. 2004. UCSF Chimera - A visualization system for exploratory research and analysis. *J Comput Chem*, 25:1605-1612.

Esko JD, Linhardt RJ. 2009. Proteins that bind sulfated glycosaminoglycans. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME editors. Essentials of glycobiology. New York: CSHL Press.

Fairbanks MB, Mildner AM, Leone JW, Cavey GS, Mathews WR, Drong RF, Slightom JL, Bienkowski MJ, Smith CW, Bannow CA*, et al.* 1999. Processing of the human heparanase precursor and evidence that the active enzyme is a heterodimer. *J Biol Chem*, 274:29587-29590.

Ferro DR, Provasoli A, Ragazzi M, Casu B, Torri G, Bossennec V, Perly B, Sinay P, Petitou M, Choay J. 1990. Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences. *Carbohydr Res*, 195:157-167.

Ferro V, Dredge K, Liu L, Hammond E, Bytheway I, Li C, Johnstone K, Karoli T, Davis K, Copeman E*, et al.* 2007. PI-88 and novel heparan sulfate mimetics inhibit angiogenesis. *Semin Thromb Hemost*, 33:557-568.

Ferro V, Hammond E, Fairweather JK. 2004. The development of inhibitors of heparanase, a key enzyme involved in tumour metastasis, angiogenesis and inflammation. *Mini Rev Med Chem*, 4:693-702.

Fort S, Varrot A, Schulein M, Cottaz S, Driguez H, Davies GJ. 2001. Mixed-linkage cellooligosaccharides: a new class of glycoside hydrolase inhibitors. *Chembiochem*, 2:319-325.

FRED. 2008a. FRED (Fast Rigid Exhaustive Docking). Santa Fe, NM 87508: Openeye Scientific Software.

FRED. 2008b. OMEGA C++ Toolkit. Sante Fe, NM 87508: OpenEye Scientific Software, Inc.

Freeman C, Parish CR. 1998. Human platelet heparanase: purification, characterization and catalytic activity. *Biochem J*, 330 1341-1350.

Frishman D, Argos P. 1996. Incorporation of non-local interactions in protein secondary structure prediction from the amino acid sequence. *Protein Eng*, 9:133-142.

Fux L, Feibish N, Cohen-Kaplan V, Gingis-Velitski S, Feld S, Geffen C, Vlodavsky I, Ilan N. 2009. Structure-function approach identifies a COOH-terminal domain that mediates heparanase signaling. *Cancer Res*, 69:1758-1767.

Fux L, Ilan N, Sanderson RD, Vlodavsky I. 2009. Heparanase: busy at the cell surface. *Trends Biochem Sci*, 34:511-519.

Gandhi NS, Coombe DR, Mancera RL. 2008. Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 1. Molecular modeling studies. *Biochemistry*, 47:4851-4862.

Gandhi NS, Mancera RL. 2009. Free energy calculations of glycosaminoglycan-protein interactions. *Glycobiology*, 19:1103-1115.

Gandhi NS, Mancera RL. 2010. Heparin/heparan sulphate-based drugs. *Drug Discov Today*, 15:1058-1069.

Gandhi NS, Mancera RL. 2011. Molecular dynamics simulations of CXCL-8 and its interactions with a receptor peptide, heparin fragments, and sulfated linked cyclitols. *J Chem Inf Model*, 51:335-358.

Gehlhaar DK, Verkhivker GM, Rejto PA, Sherman CJ, Fogel DR, Fogel LJ, Freer ST. 1995. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem Biol*, 2:317-324.

Gilat D, Hershkoviz R, Goldkorn I, Cahalon L, Korner G, Vlodavsky I, Lider O. 1995. Molecular behavior adapts to context: heparanase functions as an extracellular matrix-degrading enzyme or as a T cell adhesion molecule, depending on the local pH. *J Exp Med*, 181:1929-1934.

Goldenberg O, Erez E, Nimrod G, Ben-Tal N. 2009. The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures. *Nucl Acids Res*, 37:D323-D327.

Gordon JC, Myers JB, Folta T, Shoja V, Heath LS, Onufriev A. 2005. H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucl Acids Res*, 33:W368-W371.

Gozalbes R, Mosulen S, Carbajo RJ, Pineda-Lucena A. 2009. Development and NMR validation of minimal pharmacophore hypotheses for the generation of fragment libraries enriched in heparanase inhibitors. *J Comput Aided Mol Des*, 23:555-569.

Graham LD, Underwood PA. 1996. Comparison of the heparanase enzymes from mouse melanoma cells, mouse macrophages, and human platelets. *Biochem Mol Biol Int*, 39:563-571.

Guermeur Y, Geourjon C, Gallinari P, Deleage G. 1999. Improved performance in protein secondary structure prediction by inhomogeneous score combination. *Bioinformatics*, 15:413-421.

Henrissat B, Bairoch A. 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*, 293 781-788.

Henrissat B, Bairoch A. 1996. Updating the sequence-based classification of glycosyl hydrolases. *Biochem J*, 316 695-696.

Hill AD, Reilly PJ. 2008. Computational analysis of glycoside hydrolase family 1 specificities. *Biopolymers*, 89:1021-1031.

Hovel K, Shallom D, Niefind K, Belakhov V, Shoham G, Baasov T, Shoham Y, Schomburg D. 2003. Crystal structure and snapshots along the reaction pathway of a family 51 alpha-L-arabinofuranosidase. *EMBO J*, 22:4922-4932.

Huey R, Morris GM, Olson AJ, Goodsell DS. 2007. A semiempirical free energy force field with charge-based desolvation. *J Comput Chem*, 28:1145-1152.

Hulett MD, Freeman C, Hamdorf BJ, Baker RT, Harris MJ, Parish CR. 1999. Cloning of mammalian heparanase, an important enzyme in tumor invasion and metastasis. *Nat Med*, 5:803-809.

Hulett MD, Hornby JR, Ohms SJ, Zuegg J, Freeman C, Gready JE, Parish CR. 2000. Identification of active-site residues of the pro-metastatic endoglycosidase heparanase. *Biochemistry*, 39:15659-15667.

Ihrcke NS, Parker W, Reissner KJ, Platt JL. 1998. Regulation of platelet heparanase during inflammation: role of pH and proteinases. *J Cell Physiol*, 175:255-267.

Irimura T, Nakajima M, Nicolson GL. 1986. Chemically modified heparins as inhibitors of heparan sulfate specific endo-beta-glucuronidase (heparanase) of metastatic melanoma cells. *Biochemistry*, 25:5322-5328.

Ishida K, Hirai G, Murakami K, Teruya T, Simizu S, Sodeoka M, Osada H. 2004. Structure-based design of a selective heparanase inhibitor as an antimetastatic agent. *Mol Cancer Ther*, 3:1069-1077.

Ishida K, Wierzba MK, Teruya T, Simizu S, Osada H. 2004. Novel heparan sulfate mimetic compounds as antitumor agents. *Chem Biol*, 11:367-377.

Isorna P, Polaina J, Latorre-Garcia L, Canada FJ, Gonzalez B, Sanz-Aparicio J. 2007. Crystal structures of Paenibacillus polymyxa beta-glucosidase B complexes reveal the molecular basis of substrate specificity and give new insights into the catalytic machinery of family I glycosidases. *J Mol Biol*, 371:1204-1218.

Jain S, Drendel WB, Chen ZW, Mathews FS, Sly WS, Grubb JH. 1996. Structure of human beta-glucuronidase reveals

candidate lysosomal targeting and active-site motifs. *Nat Struct Biol*, 3:375-381.

Jalali-Heravi M, Asadollahi-Baboli M, Shahbazikhah P. 2008. QSAR study of heparanase inhibitors activity using artificial neural networks and Levenberg-Marquardt algorithm. *Eur J Med Chem*, 43:548-556.

Johnson DJ, Li W, Adams TE, Huntington JA. 2006. Antithrombin-S195A factor Xa-heparin structure reveals the allosteric mechanism of antithrombin activation. *EMBO J*, 25:2029-2037.

Johnstone KD, Karoli T, Liu L, Dredge K, Copeman E, Li CP, Davis K, Hammond E, Bytheway I, Kostewicz E*, et al.* 2010. Synthesis and biological evaluation of polysulfated oligosaccharide glycosides as inhibitors of angiogenesis and tumor growth. *J Med Chem*, 53:1686-1699.

Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*, 8:275-282.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*, 79:926-935.

Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform*, 9:286-298.

Kelley LA, Sternberg MJ. 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc*, 4:363-371.

King RD, Sternberg MJ. 1996. Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. *Protein Sci*, 5:2298-2310.

Kitago Y, Karita S, Watanabe N, Kamiya M, Aizawa T, Sakka K, Tanaka I. 2007. Crystal structure of Cel44A, a glycoside hydrolase family 44 endoglucanase from Clostridium thermocellum. *J Biol Chem*, 282:35703-35711.

Kongsted J, Ryde U, Wydra J, Jensen JH. 2007. Prediction and rationalization of the pH dependence of the activity and stability of family 11 xylanases. *Biochemistry*, 46:13581-13592.

Koropatkin NM, Smith TJ. 2010. SusG: a unique cell-membrane-associated alpha-amylase from a prominent human gut symbiont targets complex starch molecules. *Structure*, 18:200-215.

Kudchadkar R, Gonzalez R, Lewis KD. 2008. PI-88: a novel inhibitor of angiogenesis. *Expert Opin Investig Drugs*, 17:1769-1776.

Lai NS, Simizu S, Morisaki D, Muroi M, Osada H. 2008. Requirement of the conserved, hydrophobic C-terminus region for the activation of heparanase. *Exp Cell Res*, 314:2834-2845.

Levy-Adam F, Abboud-Jarrous G, Guerrini M, Beccati D, Vlodavsky I, Ilan N. 2005. Identification and characterization of heparin/heparan sulfate binding domains of the endoglycosidase heparanase. *J Biol Chem*, 280:20457-20466.

Levy-Adam F, Feld S, Cohen-Kaplan V, Shteingauz A, Gross M, Arvatz G, Naroditsky I, Ilan N, Doweck I, Vlodavsky I. 2010. Heparanase 2 interacts with heparan sulfate with high affinity and inhibits heparanase activity. *J Biol Chem*, 285:28010-28019.

Luthy R, Bowie JU, Eisenberg D. 1992. Assessment of protein models with three-dimensional profiles. *Nature*, 356:83-85.

Marchetti D, Liu S, Spohn WC, Carson DD. 1997. Heparanase and a synthetic peptide of heparan sulfate-interacting protein recognize common sites on cell surface and extracellular matrix heparan sulfate. *J Biol Chem*, 272:15891-15897.

McKenzie E, Tyson K, Stamps A, Smith P, Turner P, Barry R, Hircock M, Patel S, Barry E, Stubberfield C*, et al.* 2000. Cloning and expression profiling of Hpa2, a novel mammalian heparanase family member. *Biochem Biophys Res Commun*, 276:1170-1177.

McKenzie EA. 2007. Heparanase: a target for drug discovery in cancer and inflammation. *Br J Pharmacol*, 151:1-14.

Mertz B, Hill AD, Mulakala C, Reilly PJ. 2007. Automated docking to explore subsite binding by glycoside hydrolase family 6 cellobiohydrolases and endoglucanases. *Biopolymers*, 87:249-260.

Miao HQ, Liu H, Navarro E, Kussie P, Zhu Z. 2006. Development of heparanase inhibitors for anti-cancer therapy. *Curr Med Chem*, 13:2101-2111.

Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. 1998. Automated docking using a Lamarckian genetic algorithm and an empirical binding

free energy function. *J Comput Chem*, 19:1639-1662.

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*, 30:2785-2791.

Mosulen S, Orti L, Bas E, Carbajo RJ, Pineda-Lucena A. 2011. Production of heparanase constructs suitable for nuclear magnetic resonance and drug discovery studies. *Biopolymers*, 95:151-160.

Mukherjee S, Roy A, Zhang Y. 2010. PSFloger: an on-line service system for identifying protein structural and functional analogs. Ann Arbor: University of Michigan.

Mulakala C, Reilly PJ. 2002. Understanding protein structure-function relationships in Family 47 alpha-1,2-mannosidases through computational docking of ligands. *Proteins*, 49:125-134.

Mulloy B, Forster MJ, Jones C, Davies DB. 1993. N.m.r. and molecular-modelling studies of the solution conformation of heparin. *Biochem J*, 293 (Pt 3):849-858.

Naggi A. 2005. Glycol-splitting as a device for modulating inhibition of growth factors and heparanase by heparin and heparin derivatives. In: Hari GG, Robert JL, Charles AH editors. Chemistry and Biology of Heparin and Heparan Sulfate. Amsterdam: Elsevier Science. p. 461-481.

Naggi A, Casu B, Perez M, Torri G, Cassinelli G, Penco S, Pisano C, Giannini G, Ishai-Michaeli R, Vlodavsky I. 2005. Modulation of the heparanase-inhibiting activity of heparin through selective desulfation, graded N-acetylation, and glycol splitting. *J Biol Chem*, 280:12103-12113.

Nakajima M, Irimura T, Di Ferrante N, Nicolson GL. 1984. Metastatic melanoma cell heparanase. Characterization of heparan sulfate degradation fragments produced by B16 melanoma endoglucuronidase. *J Biol Chem*, 259:2283-2290.

Nakajima M, Irimura T, Nicolson GL. 1988. Heparanases and tumor metastasis. *J Cell Biochem*, 36:157-167.

Nam KH, Lee WH, Rhee KH, Hwang KY. 2010. Structural characterization of the bifunctional glucanase-xylanase CelM2 reveals the metal effect and substrate-binding moiety. *Biochem Biophys Res Commun*, 391:1726-1730.

Nasser NJ, Avivi A, Shushy M, Vlodavsky I, Nevo E. 2007. Cloning, expression, and characterization of an alternatively spliced variant of human heparanase. *Biochem Biophys Res Commun*, 354:33-38.

Nasser NJ, Nevo E, Shafat I, Ilan N, Vlodavsky I, Avivi A. 2005. Adaptive evolution of heparanase in hypoxia-tolerant Spalax: gene cloning and identification of a unique splice variant. *Proc Natl Acad Sci U S A*, 102:15161-15166.

Nielsen JE, McCammon JA. 2003a. Calculating pKa values in enzyme active sites. *Protein Sci*, 12:1894-1901.

Nielsen JE, McCammon JA. 2003b. On the evaluation and optimization of protein X-ray structures for pKa calculations. *Protein Sci*, 12:313-326.

Okada Y, Yamada S, Toyoshima M, Dong J, Nakajima M, Sugahara K. 2002. Structural recognition by recombinant human heparanase that plays critical roles in tumor metastasis. Hierarchical sulfate groups with different effects and the essential target disulfated trisaccharide sequence. *J Biol Chem*, 277:42488-42495.

Olsson M, Søndergaard C, Rostkowski M, Jensen J. 2011. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J Chem Theory Comput*, 7:525-537.

Paes G, Skov LK, O'Donohue MJ, Remond C, Kastrup JS, Gajhede M, Mirza O. 2008. The structure of the complex between a branched pentasaccharide and Thermobacillus xylanilyticus GH-51 arabinofuranosidase reveals xylan-binding determinants and induced fit. *Biochemistry*, 47:7441-7451.

Parish CR, Freeman C, Brown KJ, Francis DJ, Cowden WB. 1999. Identification of sulfated oligosaccharide-based inhibitors of tumor growth and metastasis using novel in vitro assays for angiogenesis and heparanase activity. *Cancer Res*, 59:3433-3441.

Parish CR, Freeman C, Hulett MD. 2001. Heparanase: a key enzyme involved in cell invasion. *Biochim Biophys Acta*, 1471:M99-M108.

Peterson SB, Liu J. 2010. Unraveling the specificity of heparanase utilizing synthetic substrates. *J Biol Chem*, 285:14504-14513.

Pikas DS, Li JP, Vlodavsky I, Lindahl U. 1998. Substrate specificity of heparanases

from human hepatoma and platelets. *J Biol Chem*, 273:18770-18777.

Pupko T, Bell RE, Mayrose I, Glaser F, Ben-Tal N. 2002. Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. *Bioinformatics*, 18:S71-77.

Rambaut A. 2006a. FigTree, a graphical viewer of phylogenetic trees. Edinburgh: Institute of Evolutionary Biology. p. FigTree is designed as a graphical viewer of phylogenetic trees and as a program for producing publication-ready figures.

Rambaut A. 2006b. FigTree, a graphical viewer of phylogenetic trees and as a program for producing publication-ready figures (http://tree.bio.ed.ac.uk/software/figtree/).

Rost B, Sander C. 1993. Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol*, 232:584-599.

Rost B, Sander C. 1994. Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins*, 19:55-72.

Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc*, 5:725-738.

Ryckaert J-P, Ciccotti G, Berendsen HJC. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys*, 23:327-341.

Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*, 234:779-815.

Sato M, Amemiya K, Hayakawa S, Munakata H. 2008. Subcellular localization of human heparanase and its alternative splice variant in COS-7 cells. *Cell Biochem Funct*, 26:676-683.

Sattelle BM, Almond A. 2010. Less is more when simulating unsulfated glycosaminoglycan 3D-structure: comparison of GLYCAM06/TIP3P, PM3-CARB1/TIP3P, and SCC-DFTB-D/TIP3P predictions with experiment. *J Comput Chem*, 31:2932-2947.

Schmidt A, Schlacher A, Steiner W, Schwab H, Kratky C. 1998. Structure of the xylanase from Penicillium simplicissimum. *Protein Sci*, 7:2081-2088.

Sharp KA, Nicholls A, Sridharan S. 1998. DelPhi. New York: Dept. of Biochemistry and Molecular Biophysics, Columbia University.

Shen MY, Sali A. 2006. Statistical potential for assessment and prediction of protein structures. *Protein Sci*, 15:2507-2524.

Shi J, Blundell TL, Mizuguchi K. 2001. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J Mol Biol*, 310:243-257.

Simizu S, Ishida K, Osada H. 2004a. Heparanase as a molecular target of cancer chemotherapy. *Cancer Sci*, 95:553-558.

Simizu S, Ishida K, Wierzba MK, Osada H. 2004b. Secretion of heparanase protein is regulated by glycosylation in human tumor cell lines. *J Biol Chem*, 279:2697-2703.

Simizu S, Suzuki T, Muroi M, Lai NS, Takagi S, Dohmae N, Osada H. 2007. Involvement of disulfide bond formation in the activation of heparanase. *Cancer Res*, 67:7841-7849.

Skolnick J, Brylinski M. 2009. FINDSITE: a combined evolution/structure-based approach to protein function prediction. *Brief Bioinform*, 10:378-391.

Soding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucl Acids Res*, 33:W244-W248.

Stahl M, Rarey M. 2001. Detailed analysis of scoring functions for virtual screening. *J Med Chem*, 44:1035-1042.

Takaoka T, Mori K, Okimoto N, Neya S, Hoshino T. 2007. Prediction of the structure of complexes comprised of proteins and glycosaminoglycans using docking simulation and cluster analysis. *J Chem Theory Comput.*, 3:2347-2356.

Taylor EJ, Smith NL, Turkenburg JP, D'Souza S, Gilbert HJ, Davies GJ. 2006. Structural insight into the ligand specificity of a thermostable family 51 arabinofuranosidase, Araf51, from Clostridium thermocellum. *Biochem J*, 395:31-37.

Thomas AH. 1996a. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J Comput Chem*, 17:490-519.

Thomas AH. 1996b. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J Comput Chem*, 17:520-552.

Thomas AH. 1996c. Merck molecular force field. III. Molecular geometries and

vibrational frequencies for MMFF94. *J Comput Chem*, 17:553-586.

Thomas AH, Robert BN. 1996. Merck molecular force field. IV. conformational energies and geometries for MMFF94. *J Comput Chem*, 17:587-615.

Thunberg L, Backstrom G, Wasteson A, Robinson HC, Ogren S, Lindahl U. 1982. Enzymatic depolymerization of heparin-related polysaccharides. Substrate specificities of mouse mastocytoma and human platelet endo-beta-D-glucuronidases. *J Biol Chem*, 257:10278-10282.

Tom D, Darrin Y, Lee P. 1993. Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems. *J Chem Phys*, 98:10089-10092.

Varrot A, Schulein M, Fruchard S, Driguez H, Davies GJ. 2001. Atomic resolution structure of endoglucanase Cel5A in complex with methyl 4,4II,4III,4IV-tetrathio-alpha-cellopentoside highlights the alternative binding modes targeted by substrate mimics. *Acta Crystallogr D Biol Crystallogr*, 57:1739-1742.

Viktor H, Robert A, Asim O, Bentley S, Adrian R, Carlos S. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*, 65:712-725.

Vlodavsky I, Elkin M, Abboud-Jarrous G, Levi-Adam F, Fuks L, Shafat I, Ilan N. 2008. Heparanase: one molecule with multiple functions in cancer progression. *Connect Tissue Res*, 49:207-210.

Vlodavsky I, Friedmann Y. 2001. Molecular properties and involvement of heparanase in cancer metastasis and angiogenesis. *J Clin Invest*, 108:341-347.

Vlodavsky I, Ilan N, Nadir Y, Brenner B, Katz BZ, Naggi A, Torri G, Casu B, Sasisekharan R. 2007. Heparanase, heparin and the coagulation system in cancer progression. *Thromb Res*, 120 Suppl 2:S112-S120.

Vlodavsky I, Mohsen M, Lider O, Svahn CM, Ekre HP, Vigoda M, Ishai-Michaeli R, Peretz T. 1994. Inhibition of tumor metastasis by heparanase inhibiting species of heparin. *Invasion Metastasis*, 14:290-302.

Warner CD, Hoy JA, Shilling TC, Linnen MJ, Ginder ND, Ford CF, Honzatko RB, Reilly PJ. 2010. Tertiary structure and characterization of a glycoside hydrolase family 44 endoglucanase from Clostridium acetobutylicum. *Appl Environ Microbiol*, 76:338-346.

Woods RJ. 2005-2011. GLYCAM Web (http://glycam.ccrc.uga.edu/ccrc/pages/3dspt.html). Athens, GA: Woods Group.

Yang JK, Yoon HJ, Ahn HJ, Lee BI, Pedelacq JD, Liong EC, Berendzen J, Laivenieks M, Vieille C, Zeikus GJ*, et al.* 2004. Crystal structure of beta-D-xylosidase from Thermoanaerobacterium saccharolyticum, a family 39 glycoside hydrolase. *J Mol Biol*, 335:155-165.

Yu G, Gunay NS, Linhardt RJ, Toida T, Fareed J, Hoppensteadt DA, Shadid H, Ferro V, Li C, Fewings K*, et al.* 2002. Preparation and anticoagulant activity of the phosphosulfomannan PI-88. *Eur J Med Chem*, 37:783-791.

Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, 9:40.

Zhang Y. 2009. I-TASSER: fully automated protein structure prediction in CASP8. *Proteins*, 77 Suppl 9:100-113.

Zhou Z, Bates M, Madura JD. 2006. Structure modeling, ligand binding, and binding affinity calculation (LR-MM-PBSA) of human heparanase for inhibition and drug design. *Proteins*, 65:580-592.

Glycobiology, 2011, in press

Table S1. Orthologous sequences and their accession numbers obtained from the Swiss-Prot database.

	Species	Swiss-Prot Accession Number	Protein Name
1	*Hordeum vulgare* (Barley)	Q70YJ3	Heparanase, putative
2	*Zea mays* (Maize)	B6SRR5	Heparanase-like protein 2
3	*Oryza sativa subsp. japonica* (Rice)	Q2QN56	Heparanase-like protein 2, putative, expressed
4	*Arabidopsis lyrata subsp. lyrata*	D7M0M7	Putative uncharacterised protein
5	*Arabidopsis thaliana* (Mouse-ear cress)	B9DH99	AT5G61250 protein
6	*Arabidopsis lyrata subsp. lyrata*	D7MUJ0	Glycosyl hydrolase family 79 N-terminal domain-containing protein
7	*Populus trichocarpa* (Western balsam poplar) (*Populus balsamifera subsp. trichocarpa*)	B9IE10	Predicted protein
8	*Medicago truncatula* (Barrel medic)	B7FL81	Putative uncharacterised protein
9	*Oryza sativa subsp. japonica* (Rice)	B9FY91	Putative uncharacterised protein
10	*Ricinus communis* (Castor bean)	B9RCW6	Heparanase, putative
11	*Populus trichocarpa* (Western balsam poplar) (*Populus balsamifera subsp. trichocarpa*)	B9HXL4	Predicted protein
12	*Picea sitchensis* (Sitka spruce)	B8LM26	Putative

	(Pinus sitchensis)		uncharacterised protein
13	*Sorghum bicolor* (Sorghum) (Sorghum vulgare)	C5XV25	Putative uncharacterised protein Sb04g036200
14	*Oryza sativa subsp. indica* (Rice)	A2XAN8	Heparanase
15	*Oryza sativa subsp. japonica* (Rice)	Q5SNA6	Putative beta-glucuronidase
16	*Zea mays* (Maize)	B6U8L8	Heparanase-like protein 3
17	*Zea mays* (Maize)	B4FGA1	Putative uncharacterised protein
18	*Sorghum bicolor* (Sorghum) (Sorghum vulgare)	C5Z5E2	Putative uncharacterised protein Sb10g005260
19	*Ricinus communis* (Castor bean)	B9RP09	Heparanase, putative
20	*Glycine max* (Soybean) (Glycine hispida)	C6TEX5	Putative uncharacterised protein
21	*Vitis vinifera* (Grape)	A5C7G0	Putative uncharacterised protein
22	*Populus trichocarpa* (Western balsam poplar) (Populus balsamifera subsp. trichocarpa)	B9N700	Predicted protein
23	*Ricinus communis* (Castor bean)	B9T602	Heparanase, putative
24	*Selaginella moellendorffii* (Spikemoss)	D8SBJ1	Putative uncharacterised protein
25	*Monosiga brevicollis* (Choanoflagellate)	A9V372	Predicted protein
26	*Bombyx mori* (Silk moth)	Q8T108	Heparanase-like

			protein
27	*Pediculus humanus subsp. corporis* (Body louse)	E0VFW4	Heparanase, putative
28	*Homo sapiens* (Human)	Q8WWQ2	Heparanase-2
29	*Danio rerio* (Zebrafish) (Brachydanio rerio)	A4QNY9	HPSE protein
30	*Xenopus tropicalis* (Western clawed frog) (Silurana tropicalis)	B1H188	LOC100145320 protein
31	*Gallus gallus* (Chicken)	Q90YK5	Heparanase
32	*Oryctolagus cuniculus* (Rabbit)	A5HC54	Heparanase
33	*Rattus norvegicus* (Rat)	Q71RP1	Heparanase
34	*Mus musculus* (Mouse)	Q6YGZ1	Heparanase
35	*Bos taurus* (Bovine)	Q9MYY0	Heparanase
36	*Spalax golani* (Northern Israeli blind subterranean mole rat)	Q333X8	Heparanse
37	*Ailuropoda melanoleuca* (Giant panda)	D2GUE5	Putative uncharacterised protein
38	*Homo sapiens* (Human)	Q9Y251	Heparanase, HPSE1
39	*Homo sapiens* (Human)	D9IUY7	Heparanase splice variant T5
40	*Homo sapiens* (Human)	A9JIG7	Heparanase
41	*Branchiostoma floridae* (Florida lancelet) (Amphioxus)	C3YJ42	Putative uncharacterised protein
42	*Nematostella vectensis* (Starlet sea anemone)	A7SUG9	Predicted protein
43	*Nematostella vectensis* (Starlet sea anemone)	A7S0R9	Predicted protein
44	*Monosiga brevicollis* (Choanoflagellate)	A9UPL2	Predicted protein
45	*Marinomonas sp.* (strain MWYL1)	A6VVT7	Putative

			uncharacterised protein
46	*Saccharophagus degradans* (strain 2-40 / ATCC 43961 / DSM 17024)	Q21FZ8	Retaining b-glycosidase-like protein
47	*Selaginella moellendorffii* (Spikemoss)	D8S0W9	Putative uncharacterised protein
48	*Physcomitrella patens subsp. patens* (Moss)	A9SLA7	Predicted protein
49	*Physcomitrella patens subsp. patens* (Moss)	A9SUR5	Predicted protein
50	*Physcomitrella patens subsp. patens* (Moss)	A9STU2	Predicted protein
51	***Physcomitrella patens subsp. patens* (Moss)**	**A9T825**	**Predicted protein**

Glycobiology, 2011, in press

Table S2. Identified structural analogues with similar binding sites in the homology model of heparanase.

Representative PDB structures	GH family	RMSD	% identity	Binding site residues in the predicted model
2vt0_A, 2v3f_A	30	3.56	12	159, 224, 225, 230, 231, 298, 343
2bfg_A, 1uhv	39	4.05	12	159, 225, 230, 231, 298, 343
2eex_A	44	3.10	13	157, 158, 159, 160, 161, 162, 163, 164, 165, 224, 225, 231, 298, 343, 353

Glycobiology, 2011, in press

Table S3. pKa of all titrable groups predicted by PROPKA. PROPKA was run on the energy minimised homology model of heparanase.

RESIDUE	Predicted pK_a	pKmodel
ASP 171	3.08	3.8
ASP 183	3.2	3.8
ASP 196	3.93	3.8
ASP 209	4.07	3.8
ASP 234	3.9	3.8
ASP 245	6.24	3.8
ASP 267	10.53	3.8
ASP 291	4.96	3.8
ASP 309	6.22	3.8
ASP 314	4.28	3.8
ASP 317	4.6	3.8
ASP 357	7.6	3.8
ASP 367	7.94	3.8
ASP 395	3.96	3.8
ASP 399	4.2	3.8
ASP 403	9.51	3.8
ASP 441	3.63	3.8
ASP 449	3.76	3.8
ASP 476	3.88	3.8
ASP 504	3.69	3.8
ASP 505	4.66	3.8
GLU 221	9.3	4.5
GLU 225	9.49	4.5
GLU 244	4.68	4.5
GLU 288	3.29	4.5
GLU 308	4.61	4.5
GLU 331	4.41	4.5
GLU 343	3.93	4.5
GLU 378	3.71	4.5
GLU 396	5.63	4.5
GLU 447	6.13	4.5
GLU 513	4.76	4.5
HIS 250	2.13	6.5
HIS 296	5.21	6.5
HIS 297	2.99	6.5
HIS 392	4.91	6.5
HIS 436	3.87	6.5
HIS 458	5.52	6.5
HIS 486	6.35	6.5
TYR156	11.25	10
TYR165	10.08	10
TYR174	11.1	10
TYR210	10.39	10
TYR216	10.17	10
TYR264	12.81	10
TYR298	15.71	10

TYR299	14.6	10
TYR348	11.23	10
TYR391	17.85	10
TYR404	10.64	10
TYR434	12.55	10
TYR445	10.41	10
TYR453	13.03	10
TYR463	10.24	10
TYR468	11.5	10
TYR478	11.11	10
TYR529	10.06	10
LYS 158	9.27	10.5
LYS 159	9.1	10.5
LYS 161	10.42	10.5
LYS 214	10.98	10.5
LYS 231	10.11	10.5
LYS 232	10.16	10.5
LYS 251	10.14	10.5
LYS 255	10.34	10.5
LYS 259	9.24	10.5
LYS 262	10.53	10.5
LYS 274	10.11	10.5
LYS 277	10.28	10.5
LYS 280	9.95	10.5
LYS 284	10.4	10.5
LYS 307	9.61	10.5
LYS 325	10.12	10.5
LYS 337	9.28	10.5
LYS 338	11.47	10.5
LYS 368	6.5	10.5
LYS 411	8.96	10.5
LYS 412	10.52	10.5
LYS 417	10.18	10.5
LYS 427	9.72	10.5
LYS 430	10.07	10.5
LYS 446	9.76	10.5
LYS 462	8.55	10.5
LYS 473	10.38	10.5
LYS 477	6.54	10.5
LYS 491	10.11	10.5
LYS 501	8.33	10.5
LYS 514	10.33	10.5
LYS 538	10.24	10.5
ARG 167	12.29	12.5
ARG 193	12.35	12.5
ARG 254	12.29	12.5
ARG 272	12.32	12.5
ARG 273	12.38	12.5
ARG 303	12.59	12.5
ARG 334	13.49	12.5

ARG 374	12.61	12.5
ARG 382	10.65	12.5
ARG 428	12.41	12.5
ARG 429	11.87	12.5
ARG 432	12.27	12.5
ARG 444	12.36	12.5
ARG 465	12.3	12.5
ARG 481	12.96	12.5
ARG 517	9.59	12.5
ARG 535	12.47	12.5

Glycobiology, 2011, in press

Table S4. Heparin glycosidic linkage torsion angles (denoted as ϕ and ψ) in different residues in disaccharide substrates. The experimentally determined torsional data available is for the IdoA2S α(1→4)GlcNS6S linkage. The (1→4) linkage is defined as ϕ = O5-C1-O1-C4` and ψ = C1-O1-C4`-C5`.

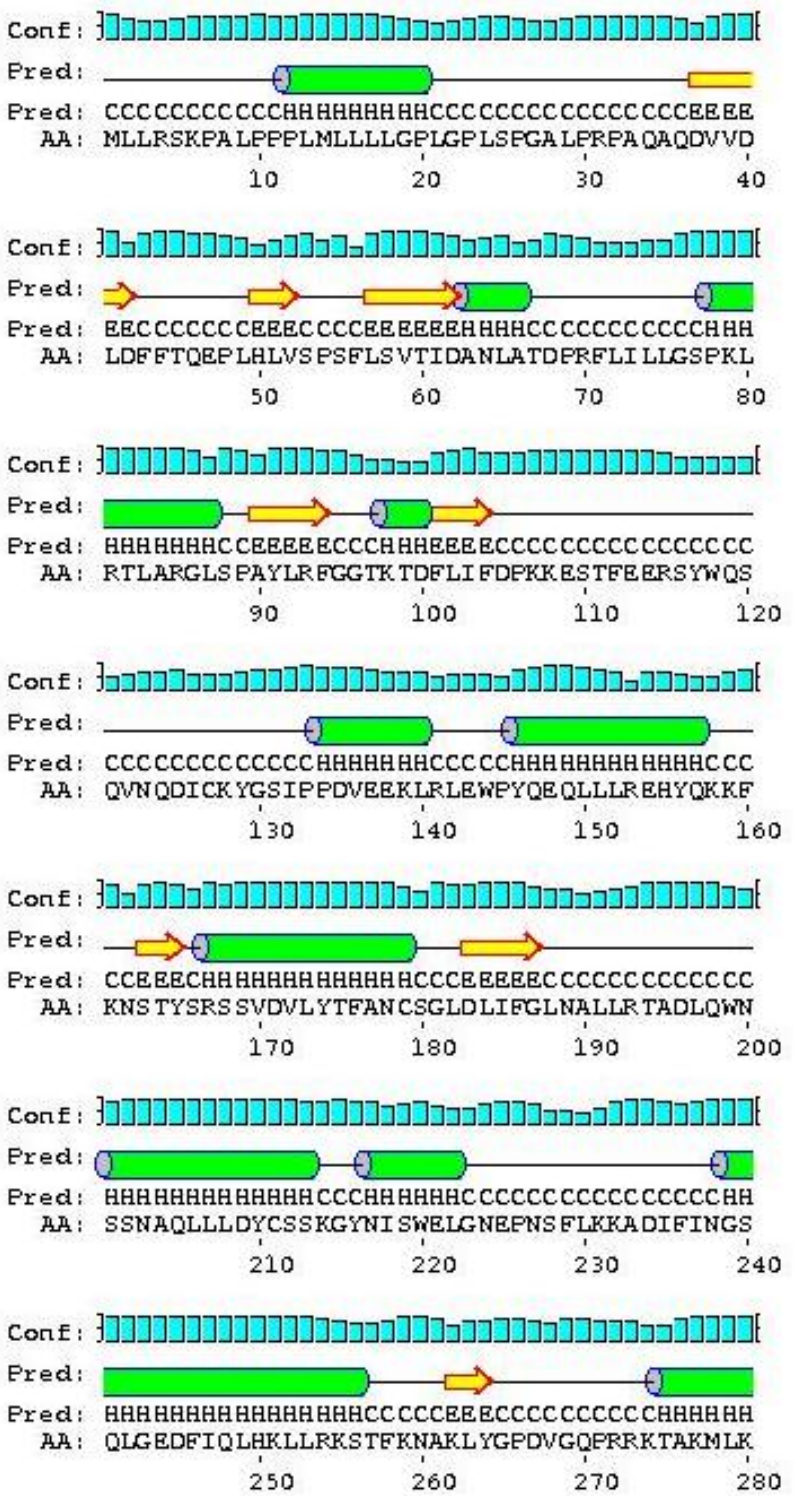
Source	Disaccharide	Torsion angle ϕ	Torsion angle ψ
[a]NMR dp12(skew boat)	IdoA2S α(1→4)GlcNS6S	-55	-107
[a]Experimental average (skew boat)	IdoA2S α(1→4)GlcNS6S	-69 ± 15	-123 ± 23
[a]Disaccharide MD (skew boat)	IdoA2S α(1→4)GlcNS6S	-84 ± 39	-126 ± 32
[a]Disaccharide MD (skew boat)	IdoA2S α(1→4)GlcNS	-95 ± 18	-127 ± 17
[a]Decasaccharide MD (skew boat)	IdoA2S α(1→4)GlcNS6S	-75 ± 11	-117 ± 11
[a]NMR dp12(chair)	IdoA2S α(1→4)GlcNS6S	-77	-110
[a]Experimental average (chair)	IdoA2S α(1→4)GlcNS6S	-69 ± 17	-118 ± 31
[a]Disaccharide MD (chair)	IdoA2S α(1→4)GlcNS6S	-113 ± 28	-142 ± 31
[a]Disaccharide MD (chair)	IdoA2S α(1→4)GlcNS	-117 ± 32	-136 ± 26
[a]Decasaccharide MD (chair)	IdoA2S α(1→4)GlcNS6S	-124 ± 18	-145 ± 11
[b]HS crystal structure	GlcA β(1→4)GlcNAc	-90.21	-110.67
[c]HS dp6	GlcA β(1→4)GlcNAc	-101 ± 38	140 ± 19*
[c]HS dp8	GlcA β(1→4)GlcNAc	-93 ± 26	121 ± 29*
[c]HS dp10	GlcA β(1→4)GlcNAc	-121 ± 26	126 ± 17*
[c]HS dp12	GlcA β(1→4)GlcNAc	-78 ± 30	114 ± 26*
[c]HS dp14	GlcA β(1→4)GlcNAc	-64 ± 32	149 ± 21*
[c]HS dp16	GlcA β(1→4)GlcNAc	-112 ± 38	136 ± 45*
[d]**1** (cluster 1) (cluster 6)	GlcA β(1→4)GlcNAc	-103.28 -69.13	-122.74 -103.72
[d]**2** (cluster 1)	GlcA β(1→4)GlcNAc6S	-143.5	-148.32

(cluster 2)		-82.96	-128.78
[d]**3** (cluster 7) (cluster 8)	GlcA β(1→4)GlcNS	-142.51 -78.39	-161.72 -136.44
[d]**4** (cluster 1)	GlcA β(1→4)GlcNS6S	-61.95	-75.89
[d]**5** (cluster 3)	GlcA β(1→4)GlcNS3S	-64.85	-95.66
[d]**6** (cluster 3)	GlcA2S β(1→4)GlcNS	-139.45	-16.19
[d]**7 (skew boat)** (cluster 1)	IdoA2S α(1→4)GlcNS	-30.11	-120.48
[d]**7 (chair)** (cluster 1)	IdoA2S α(1→4)GlcNS	-49.24	-108.53

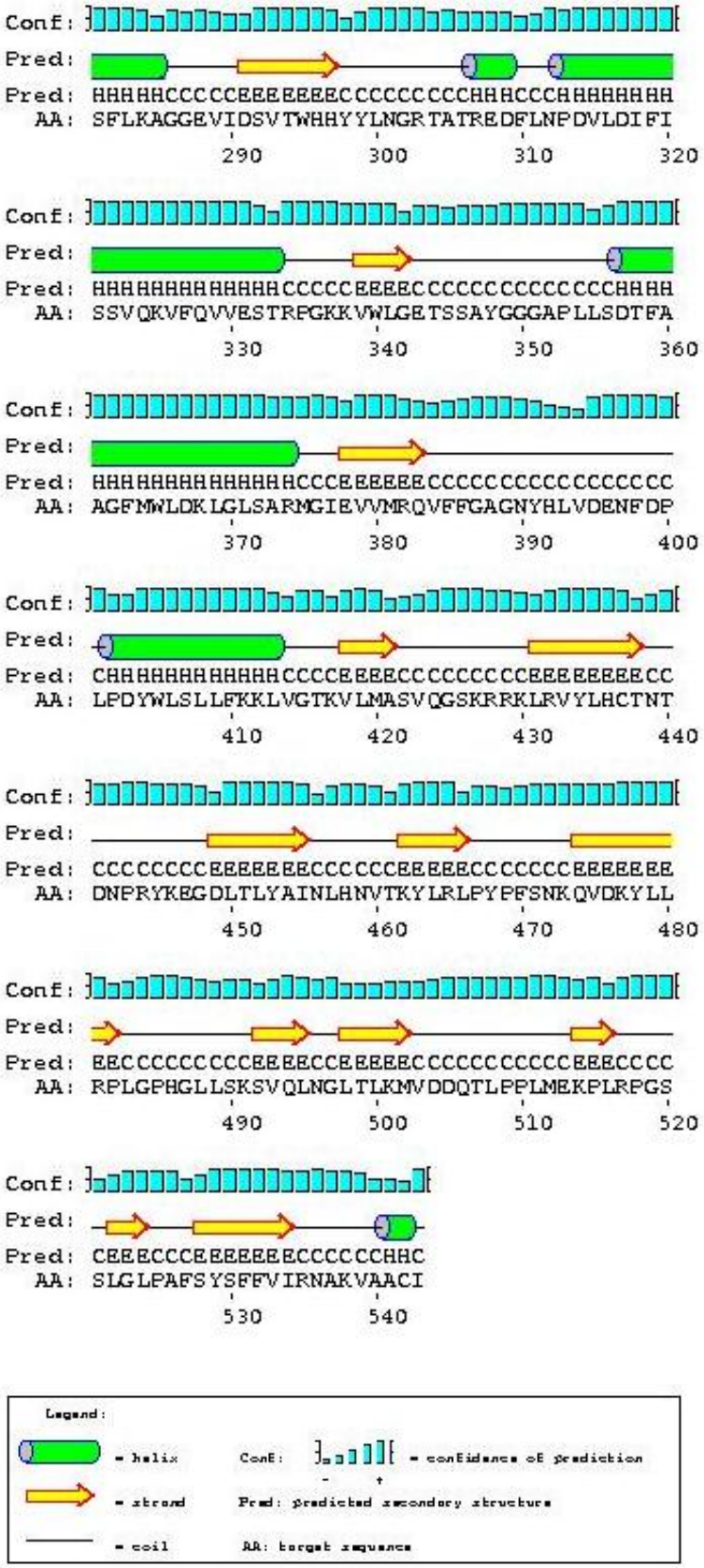
[a] MD values and experimental data adapted from "Pol-Fachin L, Verli H. 2008. Depiction of the forces participating in the 2-O-sulfo-alpha-L-iduronic acid conformational preference in heparin sequences in aqueous solutions. *Carbohydr Res*, 343:1435-1445". [b] Average values of the two tetrasaccharide molecules in the crystal structure (PDB code 3E7J). [c] Experimental data taken from "Khan, S., et al., The solution structure of heparan sulphate differs from that of heparin: implications for function. J Biol Chem, 2011 in press".* ψ = C1-O1-C4`-C3` [d] Torsion values based on the top ranking pose in the cluster as reported in Table IV.

Glycobiology, 2011, in press

Figure S 1. Secondary structure of heparanase predicted by PSIPRED.

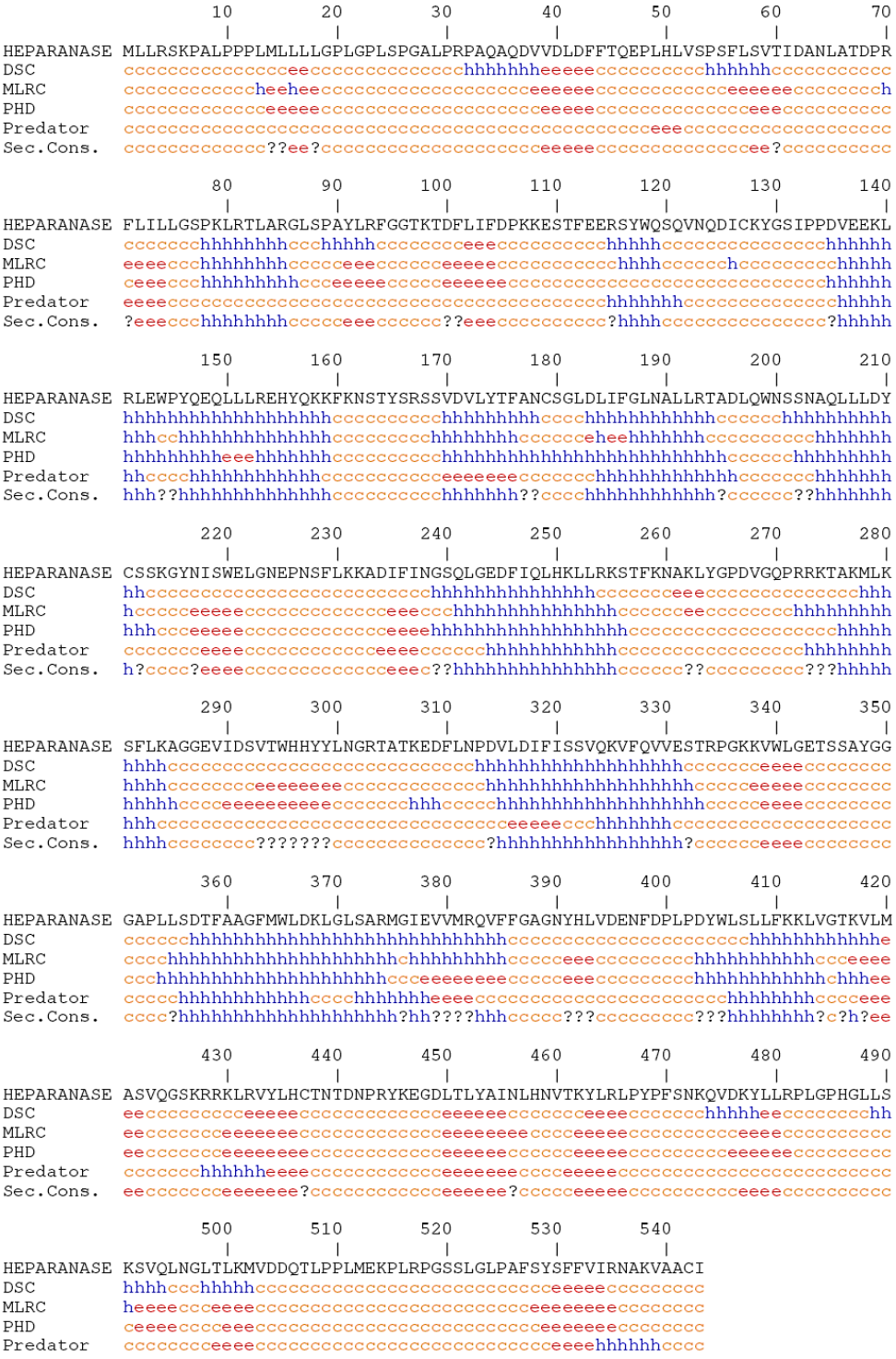


Glycobiology, 2011, in press



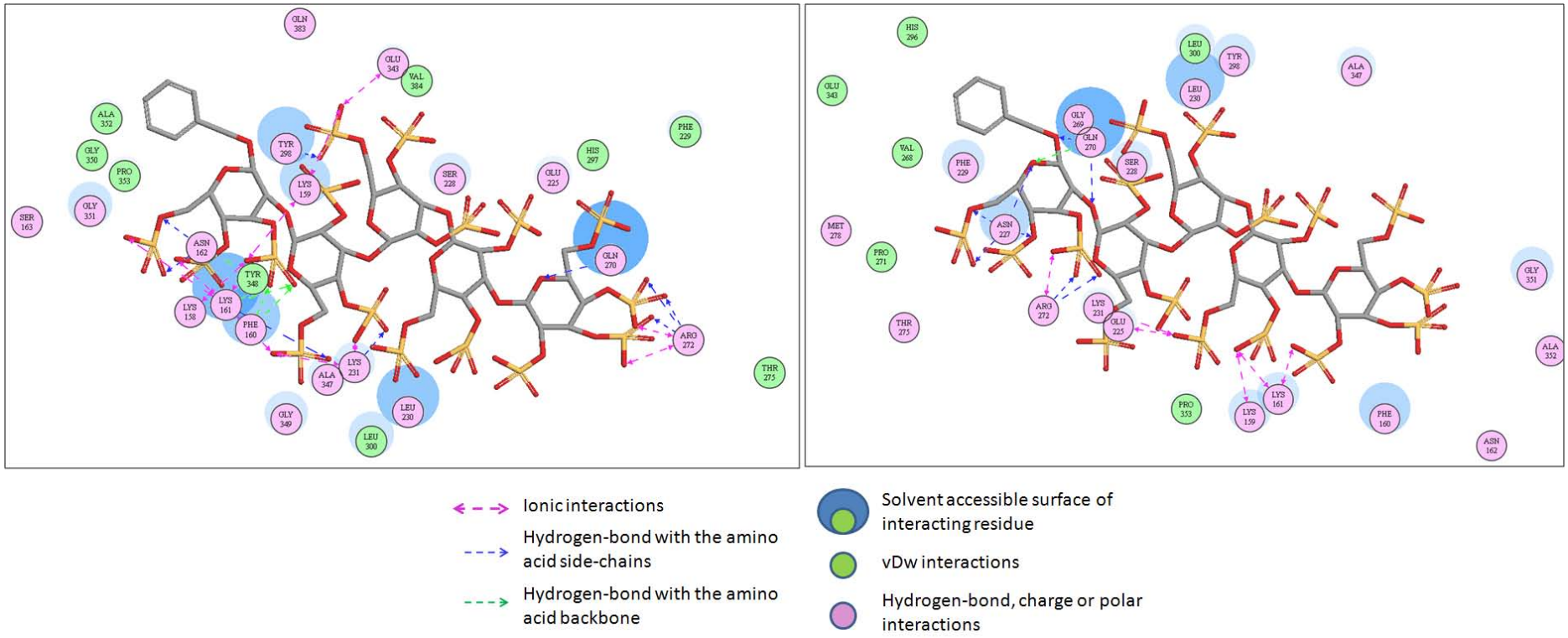
Glycobiology, 2011, in press

Figure S2. Consensus secondary structure predictions were built using the DSC, MLRC, PHD and Predator methods available at the NPSA server (http://pbil.ibcp.fr/NPSA/npsa_server.html), where 'h' represents α–helix, 'b' represents stand and 'c' stand for coiled regions.



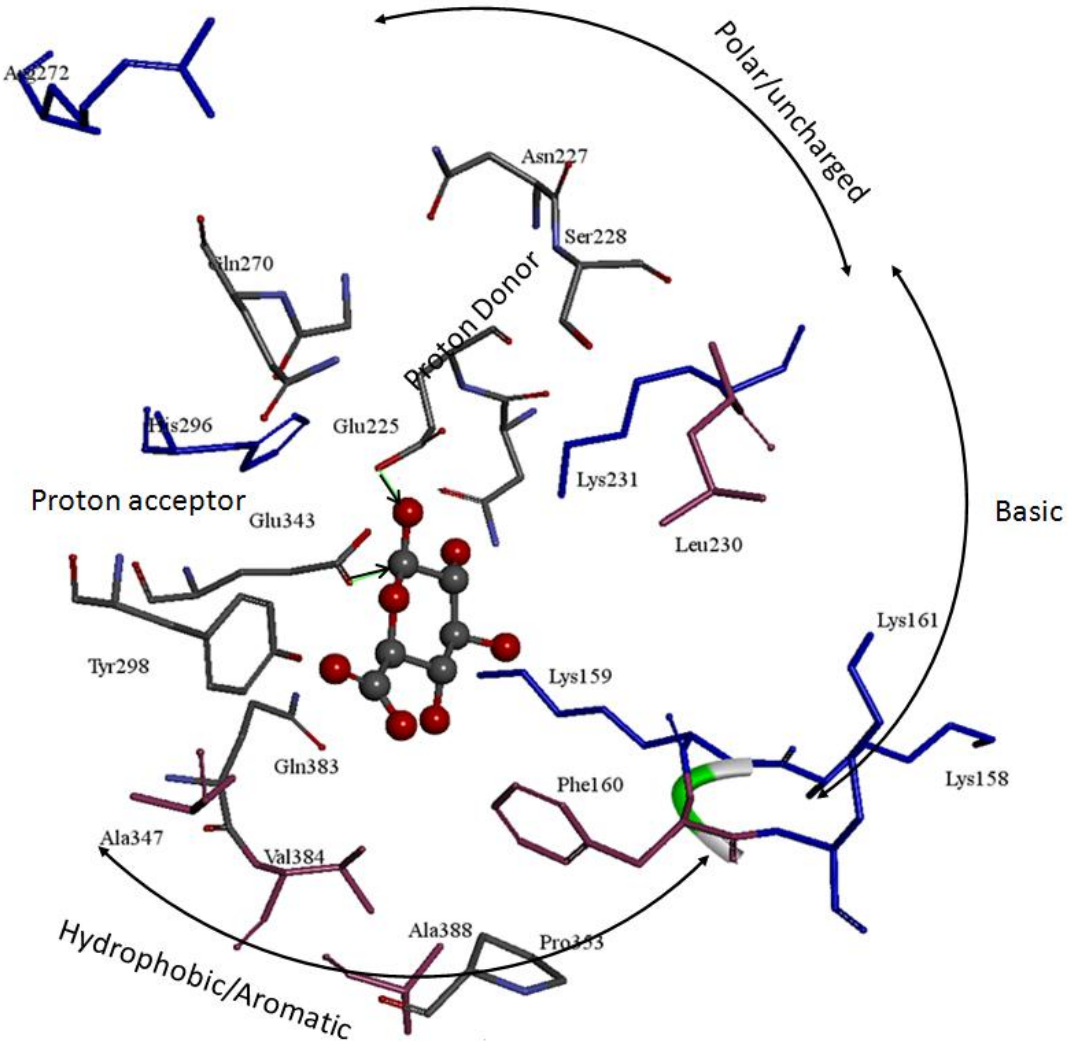
Glycobiology, 2011, in press

Figure S3. Docking of PI-88 analogue 13 into the catalytic site. Two different binding modes are observed, where the CH_2-Ph fits into the hydrophobic pocket.



Glycobiology, 2011, in press

Figure S4. Putative heparanase residues involved in enzyme-substrate interactions. Glucuronic acid is taken as a representative substrate. The colours magenta and blue indicate the hydrophobic and basic properties in the active site. Apart from the heparin binding domains consisting of basic residues (Lys158, Lys159, Lys 161, Arg272), the generalised heparanase pharmacophore consists of a proton donor/acceptor (Glu225), a hydrophobic pocket (Phe160, Leu230, Val384, Ala388) below the -1 subsite and polar neutral residues (Asn227, Ser228) above the +1 subsite. Hydrogens are not shown for clarity.



6 Free Energy Calculations of Glycosaminoglycan-Protein Interactions

Glycobiology vol. 19 no. 10 pp. 1103–1115, 2009
doi:10.1093/glycob/cwp101
Advance Access publication on July 30, 2009

Free energy calculations of glycosaminoglycan–protein interactions

Neha S Gandhi[2,3,4] and Ricardo L Mancera[1,2,3,4,5]

[2]Curtin Health Innovation Research Institute; [3]Western Australian Biomedical Research Institute; [4]School of Biomedical Sciences; and [5]School of Pharmacy, Curtin University of Technology, GPO Box U1985, Perth WA 6945, Australia

Received on May 12, 2009; revised on June 30, 2009; accepted on July 1, 2009

Glycosaminoglycans (GAGs) are complex highly charged linear polysaccharides that have a variety of roles in biological processes. We report the first use of molecular dynamics (MD) free energy calculations using the MM/PBSA method to investigate the binding of GAGs to protein molecules, namely the platelet endothelial cell adhesion molecule 1 (PECAM-1) and annexin A2. Calculations of the free energy of the binding of heparin fragments of different sizes reveal the existence of a region of low GAG-binding affinity in domains 5–6 of PECAM-1 and a region of high affinity in domains 2–3, consistent with experimental data and ligand–protein docking studies. A conformational hinge movement between domains 2 and 3 was observed, which allows the binding of heparin fragments of increasing size (pentasaccharides to octasaccharides) with an increasingly higher binding affinity. Similar simulations of the binding of a heparin fragment to annexin A2 reveal the optimization of electrostatic and hydrogen bonding interactions with the protein and protein-bound calcium ions. In general, these free energy calculations reveal that the binding of heparin to protein surfaces is dominated by strong electrostatic interactions for longer fragments, with equally important contributions from van der Waals interactions and vibrational entropy changes, against a large unfavorable desolvation penalty due to the high charge density of these molecules.

Keywords: Annexin/PECAM-1/GLYCAM/AMBER/heparin

Introduction

The molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) method (Srinivasan et al. 1998) was developed to estimate the free energy of ligand–protein (Kollman et al. 2000) and protein–protein interactions (Massova and Kollman 1999). This method combines the calculation of interaction energies from explicit solvent MD simulations with Poisson–Boltzmann calculations of the solvation energy (Gilson and Honig 1988; Honig and Nicholls 1995) and molecular surface area-based calculations of the nonpolar contribution to the solvation free energy (Sanner et al. 1996). MM-GBSA (molecular mechanics-generalized Born surface area) calculations (where electrostatic calculations are performed using the generalized Born approach) (Tsui and Case 2001) have been used to investigate how electrostatic interactions dictate the high affinity of anionic carbohydrates such as Gal-β-(1,4)-GlcNAc for galectin-1. To our knowledge, these free energy methods have not been used to simulate the interactions of GAGs with proteins.

GAGs are challenging from a molecular modeling perspective because of their high negative charge density, their conformational flexibility, and the absence of well-defined binding pockets or high surface complementarities on their target protein. The accurate computational prediction of the free energy of the interaction of sulfated GAG-protein complexes is still in its infancy, particularly because of the poorly defined contribution of water (solvation/desolvation), the large electrostatic interactions involved, and limitations in the force fields and scoring functions used to represent GAG structure, dynamics, and interactions. These limitations are slowly beginning to be overcome. The Monte Carlo multiple minima (MCMM) method (Kesserü and Kolossváry 1999) has been used to sample the many degrees of conformational freedom present in large GAG molecules as part of an investigation into the possible binding modes of cyclitols (GAG-like sulfated molecules) on the fibroblast growth factors 1 and 2 (FGF-1 and FGF-2) (Cochran et al. 2005). MD simulations have been used successfully to model the structure and dynamics of various carbohydrates (Ford et al. 2003). In the case of sulfated GAGs, simulations of heparin/HS in explicit solvent and gas phase (Mulloy et al. 1993; Mikhailov et al. 1996, 1997; Verli and Guimarães 2004; Becker et al. 2005; Jin et al. 2005; Murphy et al. 2008; Zhang et al. 2008) have provided insights into their conformational flexibility, while other simulations have investigated their interactions with proteins (Krieger et al. 2004; Canales et al. 2006). A combined docking, MD simulation, and NMR study showed that a heparin hexasaccharide induced FGF-1 dimerization either in a *cis*- or *trans*-configuration but is not required for biological activity of this growth factor (Canales et al. 2006). An MD simulation study supported by isothermal fluorescence titration experiments' study suggested that a heparin disaccharide can bind to the IL-8 dimer with high affinity (Krieger et al. 2004).

Annexin A2 and PECAM-1 are two examples of proteins that can bind heparin fragments with high affinity (Kassam et al. 1997; Coombe et al. 2008). Annexin A2 plays an important role in membrane trafficking and cytoskeletal actin bundling, while extracellular annexin A2 has been proposed to play a role in the fibrinolytic pathway (Gerke et al. 2005). Annexin A2 is known to bind heparin with high affinity and in a calcium-dependent manner, being thus also involved in the regulation of thrombotic processes (Kassam et al. 1997). The crystal structure of annexin A2 has revealed that this protein binds to up to five heparin sugar residues and that two calcium ions mediate this binding interaction (Shao et al. 2006). The binding affinity of full-length

[1]To whom correspondence should be addressed: Tel: +61-8-9266-1017; Fax: +61-8-9266-2769; e-mail: R.Mancera@curtin.edu.au

© The Author 2009. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org

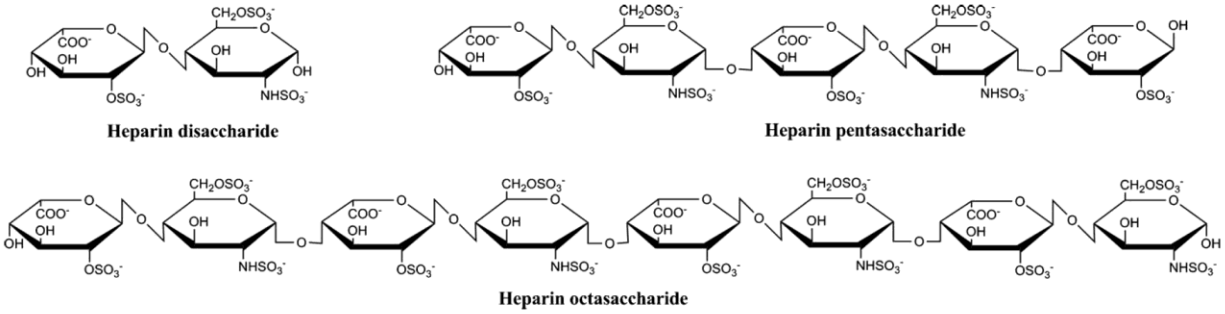


Fig. 1. Chemical structures of the heparin fragments used in the docking and MD simulations.

heparin and annexin A2 at pH 5.0 and pH 7.4 is 366 and 17 nM, respectively (Shao et al. 2006).

PECAM-1 is found on the surface of many cells of the immune system (Newman 1997) and is known to play an essential role in the transendothelial migration of leukocytes (diapedesis) (Muller 1993), angiogenesis (Cao et al. 2002), and T-cell activation (Zehnder et al. 1995). HS proteoglycans have been proposed to bind to PECAM-1 (DeLisser et al. 1993; Piali 1995; Prager 1996; Deaglio et al. 1998), and the binding of heparin fragments has recently been confirmed, with stronger binding observed at a pH of 6.0 than at a pH of 7.0 (Coombe et al. 2008; Gandhi et al. 2008). Recent molecular modeling studies of PECAM-1 and its interactions with various GAG fragments predicted the existence of high- and low-affinity GAG binding regions in PECAM-1 (Coombe et al. 2008; Gandhi et al. 2008), which was confirmed experimentally (Coombe et al. 2008).

Here, we report the first use of MD free energy calculations using the MM-PBSA method to investigate the interactions of heparin fragments (shown in Figure 1) of different sizes with annexin A2 and PECAM-1. In addition, these heparin fragments have been simulated in explicit water to characterize their structure and dynamics in the aqueous solution.

Results and discussion

Figure 2A and B shows the time evolution of the temperature and potential energy in the simulations of the heparin disaccharide complexed with Ig-domains 5 and 6 of PECAM-1. The temperature and potential energy fluctuate around converged average values. The root mean square deviation (RMSD) of the coordinates in each snapshot with respect to the coordinates in the initial snapshot was monitored for simulation of the protein and ligand complex, as shown in Figure 2C. The high RMSD values (up to $\sim$ 6 Å) observed with Ig-domains 5 and 6 indicate the occurrence of a conformational change in the backbone of these domains, as discussed in more detail further below.

Figure 3A and B shows the corresponding time evolution of the temperature and potential energy in the simulations of the heparin pentasaccharide complexed with Ig-domains 2 and 3, exhibiting fluctuations around converged values. A significant amount of backbone motion (up to $\sim$ 5 Å) in Ig-domains 2 and 3 can be observed in Figure 3C, as measured from the separate simulations of the protein and ligand. The high RMSD values indicate that there is a significant conformational change in these domains, and this is also reflected in the molecular mechanics and solvation energies of the complex. The time evolution of the energy contributions of the protein during the simulation is shown in Figure 4. The relationship between the relaxation energy (sum of solvation and molecular mechanics energies) and protein flexibility (conformational free energy) was recently investigated using the MM-PBSA method and shown to be important for the accurate estimation of free energies of binding (Swanson et al. 2004). The binding affinity of a small ligand to FKBP12 was found to be 10 kJ/mol lower than that measured experimentally when the protein relaxation energy and configurational free energy were ignored, both of which are expected to be slightly positive (Swanson et al. 2004).

Conformations of the free and protein-bound heparin pentasaccharide

Analyses of the structures of the heparin pentasaccharide bound to annexin A2 and to Ig-domains 2–3 of PECAM-1, and in aqueous solution were carried out and compared with the heparin fragment obtained from the crystal structure to investigate whether there are any differences in the conformation of the glycosidic linkages between each oligosaccharide monomer. The dihedral angles (φ, Ψ) for the GlcNS6S(1→4)IdoA2S and IdoA2S(1→4)GlcNS6S glycosidic linkages are between 80° and 110° and −70° and 120°, respectively, for heparin in the aqueous solution (Mulloy et al. 1993). The dihedral angle observed in X-ray or NMR structures has been shown to differ from these values by up to 50° (Mulloy and Forster 2000). The conformations of the pentasaccharide fragment in all three cases were analyzed from the first 4.0 ns of the MD trajectories. Table I lists the average values of the four glycosidic torsion angles (as described in Mulloy and Forster (2000)) of the pentasaccharide in each case.

All glycosidic linkages, $\alpha(1,4)_1$, $\alpha(1,4)_2$, $\alpha(1,4)_3$ and $\alpha(1,4)_4$, of the pentasaccharide exhibited greater fluctuations in the MD simulations when bound to PECAM-1 and annexin A2 than in the aqueous solution. The torsional angles Ψ of the $\alpha(1,4)_1$ linkage in the aqueous solution and in the protein-bound structures exhibited large fluctuations due to the change in the conformation of the first iduronic acid residue (present in the $^1H_2$ conformation). The conformation of the glycosidic linkage $\alpha(1,4)_2$ in the aqueous solution and in the protein-bound structures forms

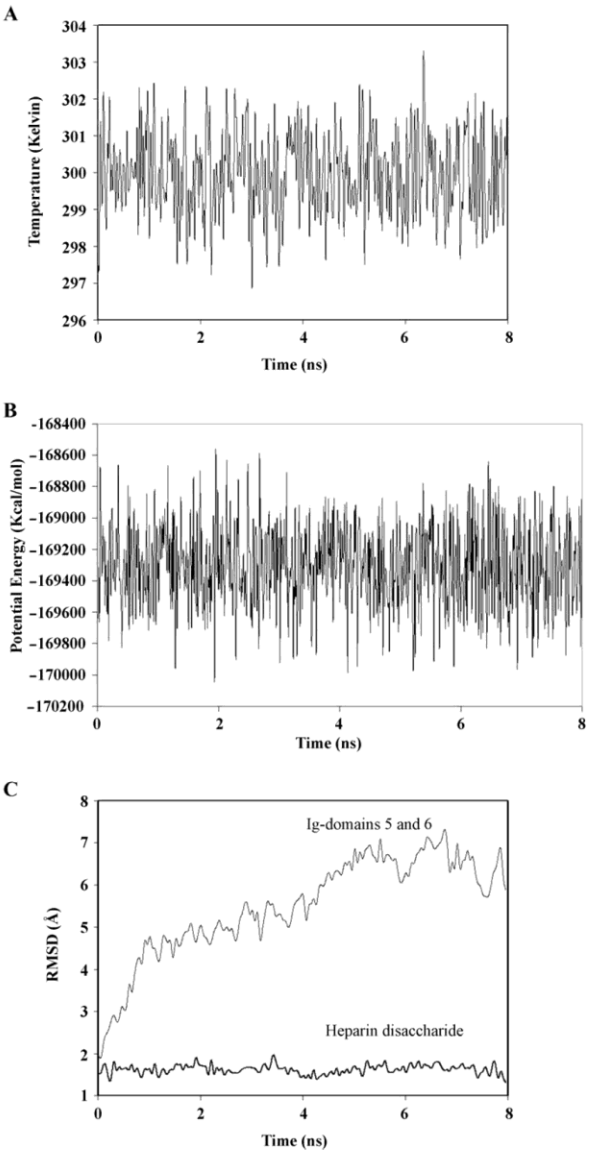


Fig. 2. Time evolution of MD simulation of a heparin disaccharide complexed with Ig-domains 5 and 6 of PECAM-1. (**A**) Temperature, (**B**) potential energy, and (**C**) root-mean-square deviation (RMSD) of the coordinates of protein main chain atoms (C_α, C, and N) and heparin disaccharide in each snapshot of their separate simulations with respect to the starting coordinates.

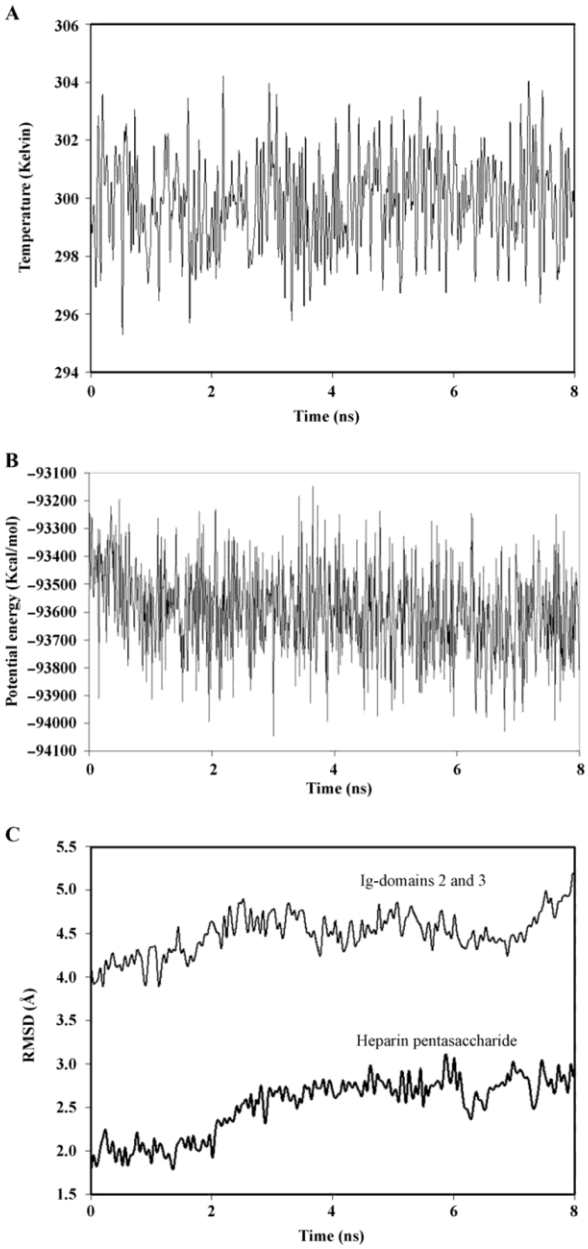


Fig. 3. Time evolution of MD simulation of a heparin pentasaccharide complexed with Ig-domains 2 and 3 of PECAM-1. (**A**) temperature, (**B**) potential energy, and (**C**) root-mean-square deviation (RMSD) of the coordinates of protein main chain atoms (C_α, C, and N) and heparin pentasaccharide in each snapshot of their separate simulations with respect to the starting coordinates.

are all similar to each other. There is no clear pattern in the conformations of linkage $\alpha(1,4)_3$ across the three cases. A similar conformation is observed for the $\alpha(1,4)_4$ glycosidic linkage in the crystal structure of annexin A2 and in the aqueous solution, but not in the MD simulations of annexin A2 and PECAM-1.

The above-mentioned data suggest that the heparin pentasaccharide undergoes a conformational change upon binding to either protein. In the case of the annexin A2 simulation, this conformational change may be due to the various interactions of the heparin fragment with the protein loops, water molecules, and the calcium ions present on the surface of the protein, as

Table I. Average values of glycosidic torsion angles of the heparin pentasaccharide extracted from the annexin A2 crystal structure (PDB code 2HYV), the pentasaccharide complexed with annexin A2 over 4.0 ns, the pentasaccharide complexed with Ig-domains 2 and 3 of PECAM-1 over 4.0 ns, and the pentasaccharide in aqueous solution over 4.0 ns

Glycosidic linkage	Torsion angles	Annexin A2 crystal structure	Annexin A2-heparin simulations	PECAM-1-heparin simulations	Heparin in aqueous solution
$\alpha(1,4)_1$	Ψ	−0.1	−6.9 (30.0)	140.1 (95.3)	156.6 (68.7)
	φ	56.2	36.0 (28.5)	11.2 (9.4)	14.1 (8.4)
$\alpha(1,4)_2$	Ψ	−23.1	−37.7 (40.1)	−27.8 (7.9)	−38.1 (8.9)
	φ	−36.6	−47.2 (38.1)	−23.39 (10.3)	−41.1 (8.4)
$\alpha(1,4)_3$	Ψ	−39.0	−12.4 (34.3)	−1.0 (17.1)	8.9 (11.9)
	φ	0.3	−49.0 (43.9)	55.6 (8.8)	48.9 (11.7)
$\alpha(1,4)_4$	Ψ	−21.8	25.1 (37.7)	10.6 (8.0)	−24.6 (13.5)
	φ	−45.3	−39.5 (47.8)	28.0 (9.4)	−34.9 (10.5)

The Ψ and φ angles in the $\alpha(1, 4)$ linkages are defined as C1-Ox-Cx-Hx and H1-C1-Ox-Cx, respectively. Standard deviations are shown in brackets.

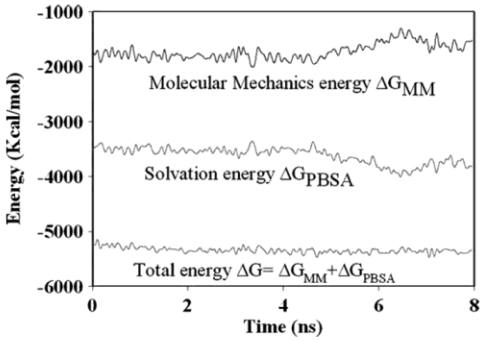


Fig. 4. (**A**) Solvation (gray line) and molecular mechanics (black line) energies of Ig-domains 2 and 3 of PECAM-1 over 8 ns of simulation. The gradual decrease in the total energy change is associated with a favorable conformational change in the hinge region between Ig-domains 2 and 3, as indicated by the total energy change (dark gray).

Table II. Residues involved in the binding of a heparin disaccharide to Ig-domains 5 and 6 of PECAM-1 in the docking and MD simulations

		Interacting amino acids and water molecules of Ig-domains 2 and 3	
Sugar Residue	Substituent	Docking	MD simulations
IdoA2S	2-*O*-sulfate	Lys 423 (s)	Water
			Asp 460 (b)
	carboxylate		Water
			Glu 470 (b)
	OH	Ser 529 (b)	Gly 528 (b)
GlcNS6S	*N*-Sulfate	Lys 423 (s)	Lys 423 (s)
	6-*O*-sulfate	Thr 533 (b)	Water
			Asn 576 (s)
			Arg 577 (b)
	sulfamide	–	Glu 527 (s)

(s) = sidechain of amino acid; (b) = backbone of amino acid.

discussed further below. The X-ray structures of bound complexes of longer heparin oligosaccharides to annexin A2, such as hexa and octasaccharides, reveal that there is no visible electron density beyond residue E. Hence, the O-1 oxygen of residue E is not really in the free hydroxyl form, as we have simulated it, but forms a 1→4 glycosidic linkage to the next residue and hence may adopt a different conformation. The dihedral angles observed in X-ray or NMR structures can vary from those observed in MD simulations of heparin or heparin-protein complexes in aqueous solutions (Mulloy and Forster 2000). Larger fluctuations in the case of the pentasaccharide fragment bound to PECAM-1 with respect to the conformation in the aqueous solution are likely to be due to the change in the receptor conformation, as described above. A comparison of the average conformations of the glycosidic linkages with average values obtained from NMR determinations, MD simulations, and crystal structures of heparin fragments bound to proteins like aFGF (Mikhailov et al. 1997) shows that the $\alpha(1,4)_2$, $\alpha(1,4)_3$, and $\alpha(1,4)_4$ linkages remain relatively stable, whereas there are relatively large changes in the $\alpha(1,4)_1$ linkage. Larger fluctuations of the $\alpha(1,4)_1$ linkage may occur because of the modification at the nonreducing end of the unsaturated UA2S required to create a 4-deoxy IdoA2S residue (4_{-D-}IdoA2S) since the residue adopts a different conformation to its original $^1H_2$ conformation (Gandhi et al. 2008).

Interactions of heparin fragments with PECAM-1

Previous docking simulations identified a number of residues in Ig-domains 5 and 6 of PECAM-1 involved in the low-affinity binding of heparin disaccharides: Lys 423, Lys 446, Lys 449, Asn 467, Thr 533, Ser 529, Arg 577, and His 580 (Gandhi et al. 2008). The amino acid numbering is in accordance with the Swiss-Prot protein sequence. The presence of this low-affinity GAG-binding region was confirmed experimentally (Coombe et al. 2008). The MD simulations of a heparin disaccharide reported here considered the third cluster reported in those docking simulations and which was predicted to have a free energy of binding and dissociation constant of −6.13 kcal/mol and 32.2 μM, respectively (Gandhi et al. 2008). Our MD simulations confirmed that the interaction between the heparin disaccharide and the protein is stable, although conformational changes in Ig- domains 5–6 resulted in some changes of the interactions predicted with docking to the rigid protein (Gandhi et al. 2008). Table II lists the residues involved in interactions of the heparin disaccharide with Ig-domains 2 and 3 of PECAM-1 in both the docking and MD simulations. The loss of interactions of the disaccharide with Ig-domains 5–6 arises from conformational changes in these domains as seen in Figure 2C. A conformational change from a predominantly β-sheet structure to a disordered random coil structure was also observed in the simulation of Ig-domains 5–6 alone.

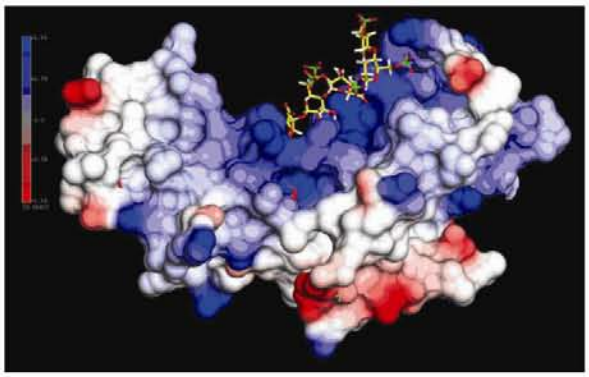


Fig. 5. Final binding mode of a heparin pentasaccharide complexed with Ig domains 2 and 3 of PECAM-1 (after 8.0 ns). The open conformation of Ig-domains 2 and 3 can interact with a longer heparin fragment through its basic residues exposed on the surface. Ig-domains 2 and 3 are represented with a solvent-accessible electrostatic potential surface (negative potential in red and positive potential in blue). The heparin pentasaccharide fragment is shown as sticks. Image created using NOC (Chen et al.).

Previous docking simulations also identified a number of residues in Ig-domains 2 and 3 of PECAM-1 involved in the high-affinity binding of heparin fragments of various sizes: Lys 176, Leu 177, Arg 179, His 239, Lys 255, Gln 259, and Ile 258 (Gandhi et al. 2008). Strong electrostatic interactions with most of these residues were found to be responsible for the computed high affinity of binding, in the presence of a high-affinity region being confirmed experimentally (Coombe et al. 2008).

Analysis of the MD trajectory of this complex revealed that a significant conformational change takes place in Ig-domains 2 and 3. Tendency for a change from an α-helical to a 3_{10} helical conformation was observed. It was also seen that the β-sheet regions are somewhat less well preserved than the helical regions, as reported previously for viscotoxin A3 (Fogolari et al. 2003). This transition is also similar to that observed in NMR studies of the globular structure of fibronectin-III (FN-III), whose β-sheeted Ig-domains adopt a random coil structure at acidic pH in solution (Penkett et al. 1997). An increase in the disordered structure in the heparin binding site and a decrease in the β-sheet content was observed during the 8 ns simulation, similar to the conformational change observed for the binding of heparin to the annexin II tetramer (AIIt) (Kassam et al. 1997). The origin of this conformational transition in Ig-domains can be attributed to the presence of large numbers of glycines (which impart conformational flexibility) and prolines (which have structure breaking properties) in their structures.

A hinge region is present in the high-affinity binding site shared by Ig-domains 2 and 3 of PECAM-1 (Gandhi et al. 2008), which can open up to expose more basic residues that may interact with a longer heparin oligosaccharide. The MD simulation indeed revealed the occurrence of a hinge movement that opened up and increased the size of the binding site (see Figure 5). This conformational change thus opened the possibility for a longer heparin fragment (such as an octasaccharide) to interact with basic residues such as Arg 179, Lys 181, Lys 255, and Gln 259, and to have a higher affinity of binding, as described further below.

Analysis of the simulation trajectory between 2–4 ns and 4–6 ns revealed that the pentasaccharide retains all the observed interactions of residues A, B, C, and D with the protein. MD simulations revealed water-mediated interactions of the sugars with the receptor in contrast to the docking studies (Gandhi et al. 2008). The interactions of 2-*O*-sulfate groups of iduronic acids with the protein are mainly water mediated. The 6-*O*-sulfate of residue B is fully solvated but does not interact with the protein only via water-mediated contacts. The pyranose ring of residue A is observed in the boat conformation. The chair ring conformations of the GlcNS6S and IdoA2S at positions B, C, D, and E were retained during the initial 2 ns period of the simulation. Table III lists the residues involved in interactions of the heparin pentasaccharide with Ig-domains 2 and 3 of PECAM-1 in both the docking and MD simulations.

Free energies of the binding of heparin fragments to PECAM-1

Earlier docking studies predicted that the free energy of the binding of a heparin disaccharide to Ig-domains 5–6 of PECAM-1 is −6.5 kcal/mol, resulting in a dissociation constant of around 15 μM, suggesting weak binding (Gandhi et al. 2008). This was confirmed experimentally (Coombe et al. 2008). Table S1 summarizes the results of the calculations of the free energy of binding using the MM-PBSA and MM-GBSA methods. The predicted free energies of binding (−1.03 kcal/mol with MM-PBSA and +6.63 kcal/mol with MM-GBSA) translate into dissociation constants in the mM range, indicating very weak binding. The calculations reveal that the favorable sum of the interaction and solvation energy terms (PBTOTAL/GBTOTAL) was not large enough to overcome the unfavorable contribution of the vibrational entropy change. It is interesting to note that the interaction energy term is not large enough because the electrostatic component of the free energy of binding (PBELE/GBELE) is large and positive, revealing that the unfavorable desolvation cost is larger than the favorable direct electrostatic interactions. Nonetheless, van der Waals (VDW) interactions provide a significant favorable contribution to the affinity of binding.

Earlier docking studies predicted that the free energy of the binding of a heparin pentasaccharide to Ig-domains 2–3 is −11.2 kcal/mol, resulting in a dissociation constant of around 5 nM, suggesting strong binding (Gandhi et al. 2008). This was confirmed experimentally (Coombe et al. 2008). Table S2 summarizes the results of the calculations of the free energy of binding using the MM-PBSA and MM-GBSA methods. The predicted free energies of binding (−16.18 kcal/mol with MM-PBSA and −13.42 kcal/mol with MM-GBSA) translate into dissociation constants of 158 and 163 pM, respectively. In this case, the interaction energy term is significantly larger because the electrostatic component of the free energy of binding (PBELE/GBELE) is large and negative, revealing that the unfavorable desolvation cost is now smaller than the favorable direct electrostatic interactions. For this longer heparin fragment, VDW interactions also provide a larger favorable contribution to the affinity of binding.

The above affinity values for the heparin pentasaccharide are likely to be influenced by those changes in the conformation of Ig-domains 2–3 described above. Hence, the MM-PBSA and MM-GBSA analyses were also carried out separately for each of four 2.0 ns portions of the simulation trajectory. The calculated free energies of binding (MM-PBSA) were −6.7 kcal/mol

Table III. Residues involved in the binding of a heparin pentasaccharide to Ig-domains 2 and 3 of PECAM-1 in the docking and MD simulations

		Interacting amino acids and water molecules of Ig-domains 2 and 3				
Sugar residue	Substituent	Docking	0–2 ns	2–4 ns	4–6 ns	6–8 ns
A-IdoA2S	2-*O*-Sulfate	Ile 258 (b)	Water	Water Lys 313 (s)	Water Lys 313 (s)	Water Lys 255
	Carboxylate	–	His 239 (s)	His 239 (s)	His 239 (s)	His 239 (s)
B-GlcNS6S	*N*-Sulfate	Gln 259 (s)	Glu 172 (b)	Glu 172 (b)	Glu 172 (b)	Lys 237 (s)
	6-*O*-Sulfate	Lys 255 (b)	Water	Water Lys 176 (s)	Water Lys 176 (s)	Water Lys 176 (s)
C-IdoA2S	2-*O*-Sulfate	His 239 (s)	Water	Water	Water	Water
	Carboxylate	–	Lys 173 (s) Met 174 (b) Val 175 (b)	Lys 173 (s) Met 174 (b) Val 175 (b)	Lys 173 (s) Met 174 (b) Val 175 (b)	Lys 173 (s) Met 174 (b) Val 175 (b)
D-GlcNS6S	*N*-Sulfate	–	–	–	–	Water
	6-*O*-Sulfate	Arg 179 (s)	Lys 173 (s)	Lys 173 (s)	Lys 173 (s)	Lys 158 (s)
E-IdoA2S	2-*O*-Sulfate	Lys 176 (s)	Water	Water	Lys 176 (s)	His 162 (s)
	Carboxylate	–	Water	His 162 (s)	His 162 (s)	Lys 176 (s)

(s) = sidechain of amino acid; (b) = backbone of amino acid.

for 0–2 ns (Table S3), −10.76 kcal/mol for 2–4 ns (Table S4), −22.81 kcal/mol for 4–6 ns (Table S5), and −19.79 kcal/mol for 6–8 ns (Table S6). There is a gradual decrease in the free energy of binding as the simulation progresses. The binding affinity increases during the last 4.0 ns as Ig-domains 2–3 adopt an "open" conformation with an increase in the electrostatic interactions of the heparin fragment with these domains. At the same time, there is a gradual decrease in the PBTOTAL/GBTOTAL (sum of molecular mechanics energy and polar and nonpolar solvation energies) of Ig domains 2–3 along the simulation trajectory (Figure 4). This is also accompanied by a gradual decrease in the free energy of solvation of the pentasaccharide and the protein (Tables S3–S5). These results suggest that the conformational change of domains 2–3 is thermodynamically favorable and, importantly, does not depend on the interactions with the heparin fragment.

It is important to point out that in all these free energy calculations, the vibrational entropy change contributions are of approximately the same magnitude as the interaction terms, revealing that both enthalpy and entropy play a key role in determining the free energy of the binding of heparin fragments.

In all the simulations of heparin fragments involving interactions with PECAM-1, entropy calculations were also carried out using quasi-harmonic analysis in order to try to account for conformational entropy contributions. However, the resulting entropy changes were too large and appeared to be unrealistic as the resulting free energies of binding were large and positive.

Docking and MD simulations of the interactions of a heparin octasaccharide with PECAM-1

The conformational changes associated with going from the "open" to the "closed" conformation of Ig-domains 2–3 of PECAM-1 described above suggested that larger heparin fragments may bind with even stronger affinity to the receptor due to additional electrostatic interactions with other residues. Hence, a heparin octasaccharide fragment was docked onto the protein, and MD simulations were carried out to fully characterize their interactions.

Docking simulations indicated that a heparin octasaccharide can interact with Ig-domains 2–3 mainly through ionic interactions (Figure 6). The 2-*O*-sulfate of residue A forms ionic interactions with the positively charged sidechain of Lys 176. The 6-*O*-sulfate of residue B forms a hydrogen bond with the backbone of Lys 176 while the *N*-sulfate forms an ionic interaction with the $N_{\varepsilon 2}$ of His 162. The 2-*O*-sulfate of residue C (IdoA2S) forms ionic and hydrogen bonding interactions with the $N_{\varepsilon 2}$ of His 239 and the sidechain of Gln 259, respectively. Also, the carboxylate of residue C was found to form an electrostatic interaction with the positively charged sidechain of Lys 237. The 6-*O*-sulfate of residue D forms ionic and hydrogen bonding interactions with the $N_{\varepsilon 1}$ of His 239 and the backbone of Ile 240, respectively. The 2-*O*-sulfate of residue E forms an ionic interaction with the positively charged sidechain of Lys 287. The positively charged sidechain of Lys 255 interacts with the oxygen in the fifth glycosidic linkage and the 6-*O*-sulfate of residue F. The 6-*O*-sulfate of residue F also forms hydrogen bonds with the hydroxyl of Tyr 290 and the backbone of Ile 254. The carboxylate of the IdoA2S and the *N*-sulfate of the

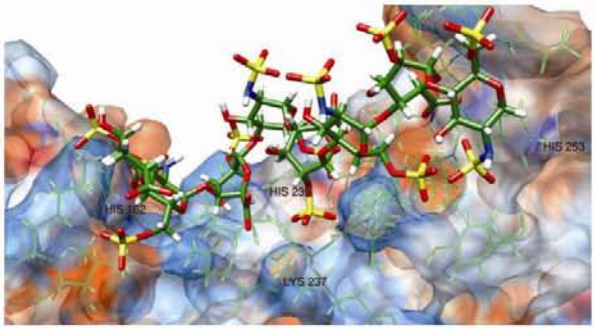


Fig. 6. Interactions of a heparin octasaccharide complexed with Ig domains 2 and 3 of PECAM-1 as predicted with AutoDock. The longer heparin fragment can interact with the open conformation of Ig-domains 2 and 3 through its basic residues exposed on the surface. Ig-domains 2 and 3 are represented with a solvent accessible electrostatic potential surface (negative potential in red and positive potential in blue). The octasaccharide fragment is shown as sticks and the basic residues are shown as lines. Image created using Chimera (Pettersen et al. 2004).

GlcNS6S in the last two residues of the octasaccharide form ionic interactions with the protonated nitrogen $N_{\varepsilon 1}$ and $N_{\varepsilon 2}$ of His 253, respectively. The free energy of binding and the dissociation constant are predicted to be −13.3 kcal/mol and 0.2 nM, respectively, revealing stronger affinity of the octasaccharide compared to the pentasaccharide, as expected.

MM-PBSA and MM-GBSA simulations of the heparin octasaccharide complexed with Ig-domains 2–3 resulted in predicted free energies of the binding of −24.17 kcal/mol and −32.23 kcal/mol, respectively (Table S7). Table S7 reveals favorable entropy term using quasi-harmonic analysis resulting in predicted energies of the binding of −6.97 kcal/mol and −15.01 kcal/mol. A structural analysis (not shown) revealed that no further opening up of Ig-domains 2–3 was observed in the receptor during the simulations. Analysis of the trajectory revealed that the 2-*O*-sulfate of residue A retains its ionic interactions with the positively charged sidechain of Lys 176, as observed in the above docking simulation, whereas the carboxylate was seen to interact strongly with a Na^+ cation through a water molecule. The interaction of the 6-*O*-sulfate of residue B with the backbone of Thr 261 is mediated by a water molecule, while the hydrogen bond observed in the docking simulation with the backbone of Lys 176 is lost. The *N*-sulfate of residue B forms ionic interactions with the $N_{\varepsilon 2}$ of His 162 and the side chain of Gln 259. The 2-*O*-sulfate of residue C forms ionic and hydrogen bonding interactions with the $N_{\varepsilon 2}$ of His 239 and the sidechain of Gln 259, respectively. Also, the carboxylate of residue C was found to interact with the backbone of Val 260 through a water molecule. The ionic and hydrogen bond interactions of the 6-*O*-sulfate of residue D with the $N_{\varepsilon 1}$ of His 239 and the backbone of Ile 240, respectively, as observed in the docking simulation are retained, while the *N*-sulfate is fully solvated and does not interact with the protein through water contacts. The 2-*O*-sulfate of residue E forms an ionic interaction with the positively charged sidechain of Lys 287 and Lys 255. The 6-*O*-sulfate of residue F interacts with the positively charged sidechain of Lys 255. The carboxylate of the IdoA2S (residue G) and the *N*-sulfate of the GlcNS6S (residue H) retain their ionic interactions with the protonated $N_{\varepsilon 1}$ and $N_{\varepsilon 2}$ of His 253, respectively, with additional interactions with water molecules.

The free energy of binding indicates that a longer oligosaccharide can easily bind to Ig-domains 2–3 of PECAM-1 with higher affinity. As in the case of the pentasaccharide fragment, the electrostatic component of the free energy of binding (PBELE/GBELE) is large and negative, once again revealing that the unfavorable energy cost of desolvation is smaller than the favorable strong direct electrostatic interactions. In addition, VDW interactions provide an even larger favorable contribution to the affinity of binding than for the pentasaccharide. It can also be seen that residues Lys 181 and Arg 184 are exposed to the surface and may be available to interact with a heparin fragment longer than an octasaccharide, consistent with experimental data (Coombe et al. 2008).

MD simulations of the interactions of a heparin pentasaccharide with annexin A2

MM-PBSA and MM-GBSA simulations were also performed on the complex made by a heparin pentasaccharide and annexin A2, as extracted from its crystal structure. This was done to provide an additional example of the use of such free energy calculations on GAG–protein interactions, particularly as the presence of calcium-mediated interactions between heparin and annexin A2 constitutes an interesting test case.

During the MD simulations, the preservation of typical secondary structural elements was monitored using Ramachandran plots. All helical structures were reasonably well preserved during the simulations. This is consistent with X-ray crystallography studies that show that the binding of heparin does not elicit a significant conformational change in annexin A2 (Shao et al. 2006).

The interactions of the heparin fragment with annexin A2 during the simulation were consistent with those observed in the crystal structure (Table IV). The heparin fragment retained its interactions via its first three saccharides (residues A–C) with Lys 280, Gly 281, Ca-8, Lys 323, and the $N_{\varepsilon 1}$ of His 93. The interaction of Ca-7 and the carboxylate of residue A mediated by a water molecule was also observed during the MD simulation. The coordination shell of Ca-8 in the crystal structure of annexin A2 is formed by the oxygen atoms from the side chain of Thr 282 and the backbone C = O of Asp 321 and two water molecules. This coordination shell was retained during the MD simulation. In the crystal structure, residue E of the heparin fragment shows no direct or water-mediated interactions with any protein atom or bound Ca^{2+} ion. However, analysis of the MD simulation trajectory revealed the formation of intermittent interactions between residue E with either water (Figure 7A) or directly with Ca-1. After 4.0 ns, residue E forms weak interactions with Ca-1 (Figure 7B). This interaction is intermittent due to the flexibility in the glycosidic linkage between residues D and E, changes in the ring conformation of residue E. Analysis of the trajectory also revealed that residue E adopts a chair conformation when it forms direct interactions with Ca-1. A small but significant conformational change in the I-AB loop is also observed during the simulation. This loop is sensitive to the presence of bound Ca ions and is known to induce a requisite conformational change in annexin V for heparin binding (Capila et al. 2001). The 2-*O*-sulfate of residue E, Gly 49, the backbone of Val 50, and the carboxylate of Glu 52 form the co-ordination shell of Ca-1. When this interaction is formed, the 6-*O*-sulfate of residue D loses its interaction with the protonated $N_{\varepsilon 1}$ of His 93, while the *N*-sulfate of residue D interacts with the backbone of Tyr 326 and the backbone NH of Asp 325, as observed in the crystal structure of the complex of a heparin tetrasaccharide with annexin A2 (Shao et al. 2006).

Both the crystal structure and the MD simulation of annexin A2 reveal that the IdoA2S residues are in the $^1C_4$ conformation. The carboxylate of IdoA2S of residue A and the *N*-sulfate of GlcNS6S of residue B form the coordination shell of Ca-7 and Ca-8, respectively. Also, residue C coordinates Ca-8 via its *O*-sulfo oxygen, whereas the sulfate at position 2 of IdoA2S of residue A is not essential for metal binding. These observations suggest that the binding preferences of heparin for Ca^{2+} on the protein surface of annexin A2 are similar to those reported for the coordination shell of Ca^{2+} for heparin in solution (Chevalier et al. 2002, 2004).

The free energy of binding analysis using MM-PBSA and MM-GBSA for the interaction of this heparin pentasaccharide to annexin A2 is reported in Table S8. The free energy of binding and dissociation constant are predicted to be −7.68 kcal/mol and 2.51 μM, respectively, by the MM-PBSA method whereas the free energy of binding is predicted to be positive by the

Table IV. Residues involved in the binding of a heparin fragment to annexin A2 as reported in the crystal structure and observed during the MD simulations

		Interacting amino acids, metal ions and water molecules of annexin A2	
Sugar Residue	Substituent	Crystal structure (Shao et al. 2006)	MD simulations
ΔUAp2S (A)	Sulfo-ester	Lys 280 (s)	Lys 280 (s)
	2-*O*-sulfate	Lys 280 (s)	Lys 280 (s)
	Carboxylate	Gly 281 (s)	Gly 281 (s)
GlcNS6S (B)	*N*-sulfate	Ca-8	Ca-8
		His 93 (s)	His 93 (s)
IdoA2S (C)	*O*-sulfate	Lys 323 (s)	Lys 323 (s)
	Pyranose O	Lys 323 (s)	Lys 323 (s)
	Carboxylate	Lys 323 (s)	Lys 323 (s)
GlcNS6S (D)	6-*O*-sulfate	His 93 (s)	–
	N-sulfate	–	Tyr 326 (b)Asp 325 (b)
IdoA2S (E)	2-*O*-sulfate	–	Ca-1

*(s) = sidechain of amino acid; (b) = backbone of amino acid.

MM-GBSA method. Normal modes analysis provided a more reasonable estimate of the entropy than the quasi-harmonic approach for the heparin–annexin system. As in the previous simulations of PECAM-1, the electrostatic component of the free energy of binding (PBELE/GBELE) is large and negative, providing the basis for the predicted high affinity of binding due to strong electrostatic interactions. The contribution of VDW interactions to the free energy is similar to that predicted for the binding of the heparin pentasaccharide to PECAM-1. Similar total energies of the interaction (PBTOTAL/GBTOTAL) have been reported for the EF-hand protein parvalbumin, where the substitution of Ser with Asp in the Ca^{2+} binding site exhibited stability and further enhanced the binding affinity for the cation (Zhao et al. 2006).

Prediction of free energies of binding for GAG–protein complexes

The free energy calculations presented in this paper are able to provide estimates of the binding affinity that are semi-quantitatively consistent with experimental and docking data. While the predicted free energies of binding have in general been overestimated, it appears that the MM-PBSA and MM-GBSA methods can provide reasonably accurate estimates of the relative binding affinities of GAG fragments to proteins. In the examples that have been presented, these methods were able to distinguish between high and low GAG-binding affinity sites and to predict increased binding affinities for larger heparin fragments, consistent with experimental data. Overall entropy changes can be measured either by normal modes analysis or by a quasi-harmonic approach. However, when applied to heparin–protein systems, the quasi-harmonic approach yielded much larger entropy values that appear to be unrealistic as they lead to large and positive free energies of binding in most cases. This quasi-harmonic approach to the calculation of conformational entropies has already been reported to suffer from sampling problems in protein–protein complexes (Holger and David 2004; Hsu et al. 2005). Entropy calculations have therefore resorted to normal modes analysis as a simplification to the problem and which appear to give better results (Case 1994; Tidor and Karplus 1994; Holger and David 2004).

Accurate prediction of free energies of the binding of GAGs will require further development and parameterization of the

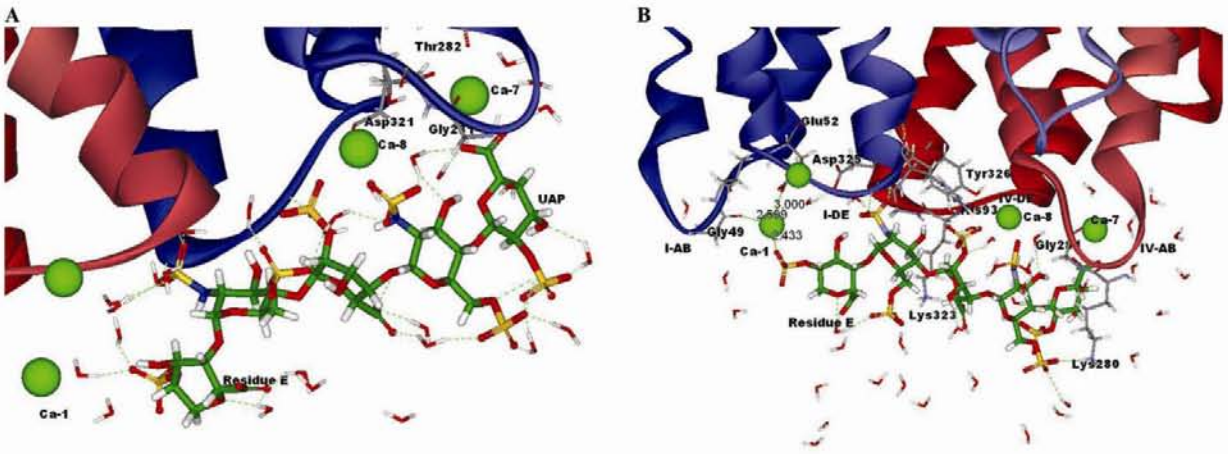


Fig. 7. Binding mode of a heparin pentasaccharide-complexed annexin A2 using trajectory analysis. **(A)** A water-mediated interaction of heparin pentasaccharide. **(B)** Analysis of the trajectories revealed the intermittent interaction of residue E of the heparin pentasaccharide with Ca-1 after 6.0 ns. The protein is shown as ribbons, metal ions as spheres, and the pentasaccharide fragment is shown as sticks. Image created using Discovery Studio 2.1.

force fields used, particularly if an appropriate description of the likely polarization effects in these systems (due to their high charge density) is to be achieved. All GLYCAM04 charges are developed from a thermally derived ensemble of conformations from long simulations performed in the presence of explicit solvent to represent the average behavior of the molecule in solution (Basma et al. 2001).

Furthermore, the RESP point charges used here for the ligand, which are taken from a single conformation, are unlikely to adequately represent the electrostatic potential around sulfate groups. Moreover, the presence of a high level of molecular flexibility in GAGs associated with hydroxyl groups makes it difficult to determine a representative ensemble of conformations in solution and this might affect the internal energies. Again, the charges derived from the single conformation might not correctly represent the *gauche–gauche*, *gauche–trans*, and *trans–gauche* rotamers of each anomer. An analogous problem arises with the simulation of nucleic acids (Cheatham and Young 2001; McDowell et al. 2007).

Conclusions

MD simulations using the MM-PBSA and MM-GBSA methods have been successfully used for the first time to investigate the interactions of heparin fragments of different size with two proteins, PECAM-1 and annexin A2. These simulations revealed that the network of ionic and hydrogen bonding interactions of heparin with proteins is optimized during the simulations, resulting in high affinities of binding for the larger fragments.

In the case of PECAM-1, these free energy calculations are in good agreement with earlier docking and experimental studies that showed the existence of high- and low-affinity GAG-binding regions in the receptor. MD simulations revealed the existence of a hinge-type conformational change affecting Ig-domains 2–3, which exposes more basic residues on the surface and thus facilitates the binding of longer-sized heparin fragments (such as an octasaccharide) with higher binding affinity.

While VDW interactions make an increasingly favorable contribution to the free energy of binding with increasing length of the heparin fragment, it appears that high affinity binding to PECAM-1 is dominated by strong electrostatic interactions. This is to be expected given the polyanionic nature of heparin and the cationic nature of the binding site of this protein. However, the solvation penalty associated with desolvating the charged sulfate groups in these heparin fragments is significantly large, resulting in very weak binding for small (disaccharide) heparin fragments. The vibrational entropy has a similar magnitude to the VDW and electrostatic interactions, and hence it also plays an important role in determining the free energy of binding. Calculations of the solvation free energies using the Poisson–Boltzmann approach and the generalized Born model give rise to similar predictions of the free energy of the binding of heparin fragments to PECAM-1.

In the case of annexin A2, MD simulations revealed the occurrence of an intermittent conformational change in a heparin pentasaccharide at the reducing end which allows the fragment to interact with an additional Ca ion on the protein surface. This is due to the intrinsic flexibility of the associated glycosidic bond and the sugar ring, as well as small but significant changes to the conformation of a neighboring loop in the protein.

Thus far, only rarely has the design of GAG-based therapeutic agents made extensive use of the computational methods such as docking and free energy simulations. In this work, free energy simulations using the MM-PBSA and MM-GBSA methods are successful in describing the interactions of heparin fragments with proteins. However, future work still needs to be done to improve the modeling of these interactions by including polarization effects and a better representation of the electrostatic properties of these molecules in the aqueous solution.

Material and methods

Starting structure of the heparin–protein complex

The coordinates of the immunoglobulin (Ig) domains 2–3 and 5–6 were extracted from our previously reported homology model of the extracellular domains of PECAM-1 (Gandhi et al. 2008). Coordinates of the complexes of a heparin pentasaccharide (Figure 1) with the closed conformation of Ig-domains 2–3 and of a disaccharide (Figure 1) with Ig-domains 5–6 were taken from our previous docking simulations (Gandhi et al. 2008). All histidine sidechains in the binding regions were protonated as the interaction of GAGs with PECAM-1 is stronger at slightly acidic pH (Coombe et al. 2008; Gandhi et al. 2008).

The heparin pentasaccharide (Figure 1) consisted of IdoA2S(1→4)GlcNS6S(1→4)IdoA2S(1→4)GlcNS6S(1→4)IdoA2S. We have modeled this pentasaccharide with its iduronic acids in the $^1C_4$ chair conformation as reported in the crystal structure of annexin A2 (PDB structure 2HYV) complexed with a heparin hexasaccharide (Shao et al. 2006). The structure of the heparin disaccharide (Figure 1) (IdoA2S(1→4)GlcNS6S) was extracted from the reported NMR structure of a heparin dodecasaccharide fragment (PDB structure 1HPN), wherein the iduronic acid is in the $^1C_4$ chair conformation and the glucosamine is in the $^4C_1$ chair conformation (Mulloy et al. 1993). This enabled the comparison of results with the similar conformation exhibited by the heparin pentasaccharide.

The heparin octasaccharide ABCDEFGF' shown in Figure 1 (IdoA2S(1→4)GlcNS6S(1→4)IdoA2S(1→4)GlcNS6S(1→4)IdoA2S(1→4)Glc*N*S6S(1→ 4)IdoA2S(1→4)Glc*N*S6S) was built from the pentasaccharide fragment extracted from the 8.0 ns MD simulations of its complex with Ig-domains 2 and 3 of PECAM-1. For this purpose, binding modes with the lowest free energy of binding were considered in conjunction with a visual search for snapshots exhibiting an "open" conformation of Ig-domains 2 and 3, so that more basic residues could be exposed to form interactions with a longer octasaccharide fragment. The octasaccharide chain conformation was generated by taking the structure of monosaccharide units B–E from the above-mentioned 2HYV crystal structure with a pentasaccharide and adding Glc*N*S6S(1→4)IdoA2S(1→4)Glc*N*S6S to the nonreducing terminus. GlcNS6S residues were added in the $^4C_1$ conformation while IdoA2S was added in the $^1C_4$ conformation. Docking simulations of this octasaccharide fragment to Ig-domains 2 and 3 of PECAM-1 in the "open" conformation were carried out as reported previously for shorter fragments (Coombe et al. 2008; Gandhi et al. 2008).

The starting structures for the simulations of the complex of a heparin pentasaccharide with human annexin A2, which includes calcium ions, were taken from PDB structure 2HYV (Shao et al. 2006) obtained at 1.42 Å resolution. Since there is

no observed electron density for the sixth saccharide residue, the fragment is effectively a pentasaccharide, as shown in Figure 1.

Parameterization of the AMBER/GLYCAM force field for heparin fragments

The GLYCAM force field has been found to represent glycosidic linkages and conformer ensembles in good agreement with those estimated by NMR determinations for heparin fragments (Angulo et al. 2003; Zhang et al. 2008). We used a similar protocol to perform unrestrained MD simulations in explicit water for heparin fragments bound to the proteins. The Parm94 (Cornell et al. 1995) force field in AMBER 9.0 (Case et al. 2005) was used with the GLYCAM04 extension for carbohydrates (Woods et al. 1995) in all MD simulations. Existing nonbonded parameters for sulfates and sulfamates were used (Huige and Altona 1995). Force constants for bond lengths and angles as well as torsional parameters that were not available for sulfates were approximated by taking those for phosphates available in the GLYCAM04 force field. Such approximation has been successfully applied to reproduce geometries of heparin oligosaccharides in gas phase simulations (Jin et al. 2005).

Partial atomic charges for the heparin di-, penta-, and octasaccharides were obtained using the restricted electrostatic potential (RESP) method (Bayly et al. 1993; Cornell et al. 1993) with the *leap* and *sander* modules in Amber 9.0. For this purpose, all molecules were initially subjected to a full geometry optimization with a 6–31G* basis set using Gaussian 98 (Frisch et al. 1998). A SCF convergence criterion of 10^{-8} kcal/mol and a "tight" optimization threshold were used. The resulting minimum energy conformation of each saccharide was then subjected to a single point energy calculation with a 6–31G* basis set and the POP = CHelpG charge option.

MD Simulations of the heparin–protein complex

Following the MM-PBSA protocol (see below), separate MD simulations were carried out for the relevant proteins (annexin A2 and PECAM-1), the heparin fragment of interest, and a complex between the two. During heating and equilibration, weak restraints (with a force constant of 25 kcal/mol/Å^2) were applied to all heavy atoms in the protein domains, except those in the GAG binding sites previously identified: residues 176–182, 207–209, 250–260, and 278–288 of Ig-domains 2 and 3 and residues Lys 423, Thr 533, Arg 577, and Lys 423 of Ig-domains 5 and 6 (Gandhi et al. 2008). Earlier protein modeling studies indicated that the PECAM-1 receptor may exist in an "open" or "closed" conformation due to the presence of loops connecting domains 2 and 3 (Gandhi et al. 2008). As a consequence, full flexibility of the receptor and heparin fragment was allowed during the production stage of the simulations. Full protein flexibility was allowed for the entire structure of annexin A2, taken from the crystal structure of its complex with a heparin hexasaccharide (PDB code 2HYV).

All energy minimizations and MD simulations were performed using the AMBER 9.0 MD package (Case et al. 2005). A cubic box of TIP3P water molecules (Jorgensen et al. 1983) was added to solvate the complex, keeping a minimum distance of 12.0 Å between each face of the box and the solute. The number of water molecules added to the annexin A2–heparin complex was 19,587. A total of 16,039 water molecules were added to the complex of the octasaccharide and the "open" conformation of Ig-domains 2 and 3 of PECAM-1, while 7181 water molecules were added to the complex of the pentasaccharide with the "closed" conformation of these domains. A total of 5651 water molecules were added to the complex of the disaccharide with Ig-domains 5 and 6 of PECAM-1. Net charges in the protein and/or heparin fragments were neutralized by adding an appropriate number of counterions (Na^+ or Cl^-). During all simulations, the particle mesh Ewald (PME) method was used to compute long-range electrostatic interactions (Tom et al. 1993), using a 1.0 Å grid spacing and a fourth-order spline for interpolation. The nonbonded cutoff was set to 9.0 Å, and the SHAKE algorithm (Ryckaert et al. 1977) was used to constrain all bonds involving hydrogen atoms. All simulations were carried out in the isobaric-isothermal (NPT) ensemble. Temperature and pressure were maintained using the weak-coupling algorithm with coupling constants τ_T and τ_P of 0.1 and 1 ps, respectively (300 K, 1 atm-Berendsen barostat) (Berendsen et al. 1984). A timestep of 1.0 fs was used in all simulations and coordinates were saved every 10 ps. Periodic boundary conditions were applied throughout.

In each simulation, initial unfavorable contacts with the solvent were removed by energy minimization after performing 10 steps of steepest descents followed by 990 steps of conjugate gradients. A 150-ps period of simulated annealing was then carried out, during which the temperature was raised from 5 to 300 K over 50 ps, with a further 50 ps at 300 K, before cooling back to 5 K over 50 ps. The system was energy minimized again as before, followed by heating from 5 to 300 K over 50 ps, upon which the systems were deemed to have equilibrated. The production phases of the simulations without any constraints were then run at 300 K for 8.0 ns for each protein and heparin–protein complex, and for 4.0 ns for each heparin fragment alone. Various properties (density, temperature, pressure, kinetic, and potential energies) were monitored during the simulations to ensure that proper equilibration had been achieved.

MM-PBSA calculations for binding free energy

In the MM-PBSA method, the free energies are calculated for representative "snapshot" structures taken from the MD trajectories of the system of interest. This method combines explicit solvation simulations with Poisson–Boltzmann analysis and nonpolar solvation free energy calculations to estimate the free energy of binding (Srinivasan et al. 1998; Kollman et al. 2000). Free energies of binding are defined as

$$\Delta G_{\text{binding}} = \Delta G_{\text{gas}} + \Delta G_{\text{sol-cmplx}} - [\Delta G_{\text{sol-prot}} + \Delta G_{\text{sol-lig}}] \quad (1)$$

where ΔG_{gas} is the interaction energy between protein and ligand in the gas phase, which is calculated using a molecular mechanics approach. $\Delta G_{\text{sol-prot}}$, $\Delta G_{\text{sol-lig}}$, and $\Delta G_{\text{sol-cmplx}}$ are the solvation free energies of the protein, ligand, and ligand–protein complex, respectively, which are estimated using a continuum Poisson–Boltzmann/surface area approach (or the generalized Born/surface approach in the case of GBSA calculations). In MM-PBSA, the free energy of binding is calculated as

$$\Delta G_{\text{binding}} = \Delta E_{\text{MM}} + \Delta G_{\text{PBSA}} - T\Delta S \quad (2)$$

where ΔE_{MM} is the difference in the average molecular mechanics energy, which is calculated as

$$\Delta E_{\text{MM}} = \Delta E_{\text{int}} + \Delta E_{\text{vdw}} + \Delta E_{\text{elec}}, \quad (3)$$

where $\Delta E_{\text{int}} = \Delta E_{\text{bond}} + \Delta E_{\text{angle}} + \Delta E_{\text{tors}}$.

ΔE_{int} corresponds to the sum of the differences in average internal bond stretching, bond bending, and torsional angle energies. ΔE_{vdw} is the difference in average VDW energy, whilst ΔE_{elec} is the difference in average electrostatic energy.

ΔG_{PBSA} is the free energy of solvation, given by

$$\Delta G_{PBSA} = \Delta G_{PB} + \Delta G_{SA}. \quad (4)$$

where ΔG_{PB} is the electrostatic component of the free energy of solvation calculated by solving the Poisson–Boltzmann equation or the generalized Born equation in the case of GB calculations (Still et al. 1990). ΔG_{SA} is the nonpolar contribution to the free energy of solvation calculated from the solvent-accessible surface area (Sitkoff et al. 1994). This term is computed with the equation $\Delta G_{SA} = \gamma SA + \beta$, where SA is the solvent-accessible surface area calculated by the MSMS program (Sanner et al. 1996) and γ and β are parameterized constants.

The final term, $T\Delta S$, is the change in vibrational entropy upon formation of the ligand–protein complex, and is approximated by performing normal modes calculations. Alternatively, a quasi-harmonic approach was also applied to determine conformational entropies (Teeter and Case 1990; Schlitter 1993).

The MM-PBSA module of AMBER 9.0 was used to compute the above-mentioned components of the free energy. For the 8.0 ns simulations, 800 snapshots of the coordinates of the system were taken at 10 ps intervals. All solvent molecules and counterions were removed prior to analysis. The snapshots were analyzed with the modified generalized Born (GB) solvation model (Tsui and Case 2001), modified for use with the PARM94 parameters to obtain energies of solvation. Poisson–Boltzmann calculations were also used to obtain solvation energies, with an ionic strength of 0.14, a dielectric constant (ε) of 1 for the solute and 80 for the solvent. A probe solvent radius of 1.4 Å and the PARSE atomic radii parameter set (Sitkoff et al. 1994) were used to determine the molecular surface. Different surface parameters were used: in the case of GB calculations, $\gamma = 0.0072$ kcal/Å^2 and $\beta = 0.0$ kcal/mol, and in the case of PB calculations, $\gamma = 0.00542$ kcal/Å^2 and $\beta = 0.92$ kcal/mol (Sitkoff et al. 1994; Sanner et al. 1996).

The vibrational, rotational, and translational entropies of the systems were computed by performing normal modes calculations (Nguyen and Case 1985) using the Nmode module of AMBER on snapshots collected every 100 ps, resulting in 80 snapshots for the 8.0 ns simulations for heparin complexed with Ig-domains 2 and 3, 5 and 6, and annexin A2. In the case of the simulations of Ig-domains 2–3 of PECAM-1, 20 snapshots were collected for normal mode analysis for each 2 ns portion of the trajectory. Prior to these normal modes calculations, the selected snapshots of the complex, protein, and heparin fragment were subjected to a full conjugate gradients energy minimization using a distance-dependent dielectric ($\varepsilon = 4r$) and a convergence criterion of 0.0001 kcal/mol. The reported vibrational, rotational, and translational entropies are the averages over all selected snapshots.

Entropy differences were also determined using a quasi-harmonic approach at 300 K using a single trajectory approach, as previously reported (Holger and David 2004). A total of 80 snapshots were used for the entropy calculations over the whole 8.0 ns trajectory, with only 20 snapshots being used for each 2 ns portion of the trajectory. Water molecules and ions were removed prior to these calculations. All atoms of the complex were used with equal weights for the fitting of the conformations in each snapshot. The coordinate frames were exported into the AMBER file format and passed on to the *ptraj* program of Amber 9.0 for the calculation of the mass-weighted covariance matrix and determination of vibrational, rotational, and translational entropies.

MM-PBSA calculations were carried out using the trajectories of the complexes of the heparin fragments with each protein. This was done as the proteins (in particular PECAM-1) exhibited significant conformational changes, as described in the results section.

Supplementary Data

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Funding

Internal Curtin Grant (RLM) and an Endeavor International Postgraduate Research Scholarship (NSG).

Acknowledgement

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

Conflict of interest statement

None declared.

Abbreviations

FN-III, fibronectin-III; GAGs, glycosaminoglycans; MD, molecular dynamics; MM-PBSA, molecular mechanics Poisson–Boltzmann surface area (MM-PBSA); MM-GBSA, molecular mechanics-generalized Born surface area; PECAM-1, platelet endothelial cell adhesion molecule 1.

References

Angulo J, Nieto PM, Martín-Lomas M. 2003. A molecular dynamics description of the conformational flexibility of the L-iduronate ring in glycosaminoglycans. *Chem Commun*. 1512–1513.

Basma M, Sundara S, Çalgan D, Vernali T, Woods RJ. 2001. Solvated ensemble averaging in the calculation of partial atomic charges. *J Comput Chem*. 22:1125–1137.

Bayly CI, Cieplak P, Cornell W, Kollman PA. 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J Phys Chem*. 97:10269–10280.

Becker CF, Guimarães JA, Verli H. 2005. Molecular dynamics and atomic charge calculations in the study of heparin conformation in aqueous solution. *Carbohydr Res*. 340:1499–1507.

Berendsen HJC, Postma JPM, Gunsteren WFv, DiNola A, Haak JR. 1984. Molecular dynamics with coupling to an external bath. *J Chem Phys*. 81:3684–3690.

Canales A, Lozano R, Lopez-Mendez B, Angulo J, Ojeda R, Nieto PM, Martin-Lomas M, Gimenez-Gallego G, Jimenez-Barbero J. 2006. Solution NMR structure of a human FGF-1 monomer, activated by a hexasaccharide heparin-analogue. *FEBS J*. 273:4716.

Cao G, O'Brien CD, Zhou Z, Sanders SM, Greenbaum JN, Makrigiannakis A, DeLisser HM. 2002. Involvement of human PECAM-1 in angiogenesis and in vitro endothelial cell migration. *Am J Physiol Cell Physiol.* 282:1181–1190.

Capila I, Hernaiz MJ, Mo YD, Mealy TR, Campos B, Dedman JR, Linhardt RJ, Seaton BA. 2001. Annexin V-heparin oligosaccharide complex suggests heparan sulfate-mediated assembly on cell surfaces. *Structure.* 9: 57–64.

Case DA. 1994. Normal mode analysis of protein dynamics. *Curr Opin Struct Biol.* 4:285–290.

Case DA, Cheatham III TE, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ. 2005. The Amber biomolecular simulation programs. *J Comput Chem.* 26:1668–1688.

Cheatham III TE, Young MA. 2001. Molecular dynamics simulation of nucleic acids: Successes, limitations, and promise. *Biopolymers.* 56:232–256.

Chen ME, Cang HX, Nymeyer H. http://noch.sourceforge.net/.

Chevalier F, Angulo J, Lucas R, Nieto PM, Martin M. 2002. The Heparin Ca^{2+} interaction: Structure of the Ca^{2+} binding site. *Eur J Org Chem.* 2367: 2376.

Chevalier F, Lucas R, Angulo J, Martin-Lomas M, Nieto PM. 2004. The heparin-Ca (2+) interaction: The influence of the *O*-sulfation pattern on binding. *Carbohydr Res.* 339:975–983.

Cochran S, Li CP, Bytheway I. 2005. An experimental and molecular-modeling study of the binding of linked sulfated tetracyclitols to FGF-1 and FGF-2. *ChemBioChem.* 6:1882–1890.

Coombe DR, Stevenson SM, Kinnear BF, Gandhi NS, Mancera RL, Osmond RIW, Kett WC. 2008. Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 2. Biochemical analyses. *Biochemistry.* 47:4863–4875.

Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc.* 117:5179–5197.

Cornell WD, Cieplak P, Bayly CI, Kollmann PA. 1993. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J Am Chem Soc.* 115:9620–9631.

Deaglio S, Morra M, Mallone R, Ausiello CM, Prager E, Garbarino G, Dianzani U, Stockinger H, Malavasi F. 1998. Human CD38 (ADP-Ribosyl Cyclase) is a counter-receptor of CD31, an Ig superfamily member 1. *J Immunol.* 160:395–402.

DeLisser HM, Yan HC, Newman PJ, Muller WA, Buck CA, Albelda SM. 1993. Platelet/endothelial cell adhesion molecule-1 (CD31)-mediated cellular aggregation involves cell surface glycosaminoglycans. *J Biol Chem.* 268:16037–16046.

Fogolari F, Brigo A, Molinari H. 2003. Protocol for MM/PBSA molecular dynamics simulations of proteins. *Biophys J.* 85:159–166.

Ford MG, Weimar T, Köhli T, Woods RJ. 2003. Molecular dynamics simulations of galectin-1-oligosaccharide complexes reveal the molecular basis for ligand diversity. *Proteins.* 53:229–240.

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA Jr, Stratmann RE, Burant JC. 1998. *Gaussian 98. Revision A. 7.* Pittsburgh (PA): Gaussian Inc.

Gandhi NS, Coombe DR, Mancera RL. 2008. Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its Interactions with glycosaminoglycans: 1. Molecular modeling studies. *Biochemistry.* 47:4851–4862.

Gerke V, Creutz CE, Moss SE. 2005. Annexins: Linking Ca^{2+} signalling to membrane dynamics. *Nat Rev Mol Cell Biol.* 6:449–461.

Gilson MK, Honig B. 1988. Calculation of the total electrostatic energy of a macromolecular system: Solvation energies, binding energies, and conformational analysis. *Proteins.* 4:7–18.

Holger G, David AC. 2004. Converging free energy estimates: MM-PB(GB)SA studies on the protein–protein complex Ras-Raf. *J Comput Chem.* 25:238–250.

Honig B, Nicholls A. 1995. Classical electrostatics in biology and chemistry. *Science.* 268:1144.

Hsu S-TD, Peter C, van Gunsteren WF, Bonvin AMJJ. 2005. Entropy calculation of HIV-1 Env gp120, its receptor CD4, and their complex: An analysis of configurational entropy changes upon complexation. *Biophys J.* 88:15–24.

Huige CJM, Altona C. 1995. Force field parameters for sulfates and sulfamates based on Ab initio calculations: Extensions of AMBER and CHARMm fields. *J Comput Chem.* 16:56–79.

Jin L, Barran PE, Deakin JA, Lyon M, an Uhrín D, Matter S. 2005. Conformation of glycosaminoglycans by ion mobility mass spectrometry and molecular modelling. *Phys Chem Chem Phys.* 7:3464–3471.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 79:926.

Kassam G, Manro A, Braat CE, Louie P, Fitzpatrick SL, Waisman DM. 1997. Characterization of the heparin binding properties of annexin II tetramer. *J Biol Chem.* 272:15093–15100.

Keserü GM, Kolossváry I. 1999. Molecular Mechanics and Conformational Analysis in Drug Design. Oxford: Blackwell Publishing.

Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W. 2000. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc Chem Res.* 33:889–897.

Krieger F, Geretti F, Brandner B, Goger B, Wells TN, Kungl AJ. 2004. A structural and dynamic model for the interaction of interleukin-8 and glycosaminoglycans: Support from isothermal fluorescence titrations. *Proteins.* 54:768–775.

Massova I, Kollman PA. 1999. Computational alanine scanning to probe protein–protein interactions: A novel approach to evaluate binding free energies. *J Am Chem Soc.* 121:8133–8143.

McDowell SE, Spackova N, Sponer J, Walter NG. 2007. Molecular dynamics simulations of RNA: An in silico single molecule approach. *Biopolymers.* 85:169–184.

Mikhailov D, Linhardt RJ, Mayo KH. 1997. NMR solution conformation of heparin-derived hexasaccharide. *Biochem J.* 328:51–61.

Mikhailov D, Mayo KH, Vlahov IR, Toida T, Pervin A, Linhardt RJ. 1996. NMR solution conformation of heparin-derived tetrasaccharide. *Biochem J.* 318:93–102.

Muller WA. 1993. PECAM-1 is required for transendothelial migration of leukocytes. *J Exp Med.* 178:449–460.

Mulloy B, Forster MJ. 2000. Conformation and dynamics of heparin and heparan sulfate. *Glycobiology.* 10:1147–1156.

Mulloy B, Forster MJ, Jones C, Davies DB. 1993. Nmr and molecular-modelling studies of the solution conformation of heparin. *Biochem J.* 293:849–858.

Murphy KJ, McLay N, Pye DA. 2008. Structural studies of heparan sulfate hexasaccharides: New insights into iduronate conformational behavior. *J Am Chem Soc.* 130:12435–12444.

Newman PJ. 1997. The biology of PECAM-1. *J Clin Invest.* 99:3–8.

Nguyen DT, Case DA. 1985. On finding stationary states on large-molecule potential energy surfaces. *J Phys Chem.* 89:4020–4026.

Penkett CJ, Redfield C, Dodd I, Hubbard J, McBay DL, Mossakowska DE, Smith RAG, Dobson CM, Smith LJ. 1997. NMR analysis of main-chain conformational preferences in an unfolded fibronectin-binding protein. *J Mol Biol.* 274:152–159.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem.* 25:1605–1612.

Piali L. 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.

Prager E. 1996. Interaction of CD31 with a heterophilic counterreceptor involved in downregulation of human T cell responses. *J Exp Med.* 184:41–50.

Ryckaert J-P, Ciccotti G, Berendsen HJC. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J Comput Phys.* 23:327–341.

Sanner MF, Olson AJ, Spehner J-C. 1996. Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers.* 38:305–320.

Schlitter J. 1993. Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chem Phys Lett.* 215:617–621.

Shao C, Zhang F, Kemp MM, Linhardt RJ, Waisman DM, Head JF, Seaton BA. 2006. Crystallographic analysis of calcium-dependent heparin binding to annexin A2. *J Biol Chem.* 281:31689.

Sitkoff D, Sharp KA, Honig B. 1994. Accurate calculation of hydration free energies using macroscopic solvent models. *J Phys Chem.* 98:1978–1988.

Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA. 1998. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J Am Chem Soc.* 120:9401–9409.

Still WC, Tempczyk A, Hawley RC, Hendrickson T. 1990. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J Am Chem Soc.* 112:6127–6129.

Swanson JM, Henchman RH, McCammon JA. 2004. Revisiting free energy calculations: A theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophys J.* 86:67–74.

Teeter MM, Case DA. 1990. Harmonic and quasiharmonic descriptions of crambin. *J Phys Chem.* 94:8091–8097.

Tidor B, Karplus M. 1994. The contribution of vibrational entropy to molecular association: The dimerization of insulin. *J Mol Biol*. 238:405–414.
Tom D, Darrin Y, Lee P. 1993. Particle mesh Ewald: An *N* [center-dot] log(N) method for Ewald sums in large systems. *J Chem Phys*. 98:10089–10092.
Tsui V, Case DA. 2001. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers*. 56:257–291.
Verli H, Guimarães JA. 2004. Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution. *Carbohydr Res*. 339:281–290.
Woods RJ, Dwek RA, Edge CJ, Fraser-Reid B. 1995. Molecular mechanical and molecular dynamic simulations of glycoproteins and oligosaccharides: 1. GLYCAM_93 parameter development. *J Phys Chem*. 99:3832–3846.
Zehnder JL, Shatsky M, Leung LL, Butcher EC, McGregor JL, Levitt LJ. 1995. Involvement of CD31 in lymphocyte-mediated immune responses: Importance of the membrane-proximal immunoglobulin domain and identification of an inhibiting CD31 peptide. *Blood*. 85:1282.
Zhang Z, McCallum SA, Xie J, Nieto L, Corzana F, Jiménez-Barbero J, Chen M, Liu J, Linhardt RJ. 2008. Solution structures of chemoenzymatically synthesized heparin and its precursors. *J Am Chem Soc*. 130:12998–13007.
Zhao J, Nelson DJ, Huo S. 2006. Potential influence of Asp in the Ca^{2+} coordination position 5 of parvalbumin on the calcium-binding affinity: A computational study. *J Inorg Biochem*. 100:1879–1887.

Glycobiology, 2009, 19, 1103-1115

Table S1. MM/PBSA energy component analysis of the interactions of a heparin disaccharide with Ig-domains 5-6 of PECAM-1 averaged over 8.0 ns[a].

	Complex		*Receptor*		*Ligand*		*Δ*	
	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
ELE	-4827.81	111.08	-4711.05	112.35	64.67	8.76	-181.43	31.25
VDW	-600.46	20.82	-587.52	20.34	8.08	3.54	-21.03	3.9
INT	2323.65	44.32	2299.91	43.92	23.74	6.07	0	0
GAS	-3104.62	118.29	-2998.65	121.27	96.49	9.11	-202.46	31.56
PBSUR	79.15	1.07	78.5	1.09	4.58	0.04	-3.93	0.26
PBCAL	-3726.98	93.86	-3373.83	98.8	-548.34	6.7	195.19	30.28
PBSOL	-3647.83	93.41	-3295.33	98.25	-543.76	6.72	191.25	30.14
PBELE	-8554.79	35.68	-8084.88	30.73	-483.67	4.79	13.76	7.67
PBTOT	-6752.45	53.64	-6293.98	51.99	-447.27	5.36	-11.2	5.96
GBSUR	103.92	1.43	103.06	1.45	4.86	0.06	-4	0.35
GBCAL	-3741.22	99.36	-3417.39	101.68	-524.7	6.56	200.86	29.6
GBSOL	-3637.3	98.82	-3314.32	100.97	-519.83	6.59	196.86	29.45
GBELE	-8569.03	26.42	-8128.44	24.99	-460.03	4.49	19.43	4.45
GBTOT	-6741.91	48.33	-6312.98	49.01	-423.34	5.36	-5.6	4.47
$T\Delta S^{b}$	2302.42	11.52	2250.36	10.2	67.46	0.53	-12.23	
$T\Delta S^{c}$							-20.66	
$\Delta G_{binding\text{-}PBSA}{}^{b}$							-1.03	
$\Delta G_{binding\text{-}GBSA}{}^{b}$							+6.63	
$\Delta G_{binding\text{-}PBSA}{}^{c}$							+9.46	
$\Delta G_{binding\text{-}GBSA}{}^{c}$							+15.06	

a Average over 800 snapshots of the MD simulation trajectory.

b Entropy calculations were based on normal modes analysis using 80 snapshots.

c Entropy calculations were based on quasiharmonic analysis using 80 snapshots.

ELE, non-bonded electrostatic energy; VDW, non-bonded van der Waals energy; INT, bond, angle, dihedral energies; GAS, ELE+VDW+INT; PBSUR, hydrophobic contribution to solvation free energy for PB calculations; PBCAL, reaction field energy calculated by PB; PBSOL=PBSUR+PBCAL; PBELE=PBCAL+ELE; PBTOTAL=PBSOL+GAS; GBSUR, hydrophobic contributions to solvation free energy for GB calculations; GBCAL, reaction field energy calculated by GB; GBSOL=GBSUR+GBCAL; GBELE=GBCAL+ELE; GBTOTAL=GBSOL+GAS; T,ΔS, T(temperature), ΔS (sum of rotational, translational and vibrational entropies); $\Delta G_{binding}$, total binding energy of the system. All energies are in kcal/mol.

Glycobiology, 2009, 19, 1103-1115

Table S2. MM/PBSA energy component analysis of the interactions of a heparin pentasaccharide with Ig-domains 2-3 of PECAM-1 averaged over 8.0 ns[a].

	Complex		*Receptor*		*Ligand*		Δ	
	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
ELE	-4696.88	118.36	-3374.69	142.57	1277.96	32.82	-2600.15	113.69
VDW	-476.13	27.96	-445.64	24.08	3.95	5.28	-34.44	8.48
INT	2154.66	45.29	2092.51	43.86	62.14	9.88	0	0
GAS	-3018.36	118.67	-1727.81	143.34	1344.05	31.43	-2634.6	111.89
PBSUR	73.9	2.67	71.8	2.5	8.11	0.14	-6.01	0.51
PBCAL	-3338.24	127.35	-3635.36	150.61	-2287.91	28.23	2585.03	108.51
PBSOL	-3264.34	125.31	-3563.56	148.67	-2279.8	28.34	2579.02	108.49
PBELE	-8035.12	43.56	-7010.05	36.25	-1009.95	8.92	-15.12	14.83
PBTOT	-6282.69	51.84	-5291.37	49.14	-935.75	8.3	-55.57	8.68
GBSUR	96.95	3.55	94.16	3.32	9.55	0.18	-6.76	0.68
GBCAL	-3321.74	127.89	-3687.6	154.28	-2222.68	28.94	2588.55	108.7
GBSOL	-3224.79	125.25	-3593.44	151.73	-2213.13	29.08	2581.79	108.71
GBELE	-8018.62	38.91	-7062.29	34.05	-944.73	8.7	-11.6	11.9
GBTOT	-6243.15	49.77	-5321.26	49.2	-869.08	8.96	-52.81	7.89
$T\Delta S^{b}$	2090.38	10.03	1987.91	10.47	139.80	0.37	-39.39	
$T\Delta S^{c}$							-103.141	
$\Delta G_{binding\text{-}PBSA}$ [b]							-16.18	
$\Delta G_{binding\text{-}GBSA}$ [b]							-13.42	
$\Delta G_{binding\text{-}PBSA}$ [c]							+47.57	
$\Delta G_{binding\text{-}GBSA}$ [c]							+50.33	

a Average over 800 snapshots of the MD simulation trajectory.

b Entropy calculations were based on normal modes analysis using 80 snapshots.

c Entropy calculations were based on quasiharmonic analysis using 80 snapshots.

ELE, non-bonded electrostatic energy; VDW, non-bonded van der Waals energy; INT, bond, angle, dihedral energies; GAS, ELE+VDW+INT; PBSUR, hydrophobic contribution to solvation free energy for PB calculations; PBCAL, reaction field energy calculated by PB; PBSOL=PBSUR+PBCAL; PBELE=PBCAL+ELE; PBTOTAL=PBSOL+GAS; GBSUR, hydrophobic contributions to solvation free energy for GB calculations; GBCAL, reaction field energy calculated by GB; GBSOL=GBSUR+GBCAL; GBELE=GBCAL+ELE; GBTOTAL=GBSOL+GAS; T,S; Ttemperature); ΔS/sum of rotational, translational and vibrational entropies); $\Delta G_{binding}$, total binding energy of the system. All energies are in Kcal/mol.

Glycobiology, 2009, 19, 1103-1115

Table S3. MM/PBSA energy component analysis of the interactions of a heparin pentasaccharide with Ig-domains 2-3 of PECAM-1 averaged during 0-2 ns[a].

	Complex		*Receptor*		*Ligand*		*Δ*	
	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
ELE	-4751.91	103.21	-3457.97	69.47	1267.03	22.39	-2560.96	107.24
VDW	-499.8	22.7	-462.07	24.48	2.43	4.63	-40.16	7.69
INT	2904.75	46.41	2842.88	46.13	61.87	9.76	0	0
GAS	-2346.95	101.21	-1077.16	75.5	1331.33	20.92	-2601.12	103.57
PBSUR	71.05	1.74	69.14	1.92	8.14	0.11	-6.23	0.36
PBCAL	-3239.32	110.57	-3518.58	63.8	-2275.54	17.98	2554.8	100.39
PBSOL	-3168.26	109.3	-3449.44	63.23	-2267.4	18.05	2548.57	100.47
PBELE	-7991.22	33.79	-6976.55	35.84	-1008.51	8.59	-6.16	12.54
PBTOT	-5515.22	47.64	-4526.6	47.99	-936.07	8.14	-52.54	7.98
GBSUR	93.17	2.31	90.62	2.55	9.59	0.14	-7.05	0.48
GBCAL	-3225.66	109.32	-3568.82	65.23	-2211.85	18.6	2555.01	101.2
GBSOL	-3132.49	107.72	-3478.19	64.66	-2202.26	18.69	2547.96	101.31
GBELE	-7977.56	27.14	-7026.79	28.98	-944.83	8.44	-5.95	10.58
GBTOT	-5479.44	46.07	-4555.35	45.87	-870.93	8.57	-53.16	6.23
$T\Delta S^{b}$	2078.93	9.55	1982.92	5.75	141.84	0.15	-45.84	
$T\Delta S^{c}$							-66.13	
$\Delta G_{binding\text{-}PBSA}{}^{b}$							-6.7	
$\Delta G_{binding\text{-}GBSA}{}^{b}$							-7.32	
$\Delta G_{binding\text{-}PBSA}{}^{c}$							+13.59	
$\Delta G_{binding\text{-}GBSA}{}^{c}$							+12.97	

a Average over 200 snapshots of the trajectory.

b Entropy calculations were based on normal modes analysis using 20 snapshots.

c Entropy calculations were based on quasiharmonic analysis using 20 snapshots.

ELE, non-bonded electrostatic energy; VDW, non-bonded van der Waals energy; INT, bond, angle, dihedral energies; GAS, ELE+VDW+INT; PBSUR, hydrophobic contribution to solvation free energy for PB calculations; PBCAL, reaction field energy calculated by PB; PBSOL=PBSUR+PBCAL; PBELE=PBCAL+ELE; PBTOTAL=PBSOL+GAS; GBSUR, hydrophobic contributions to solvation free energy for GB calculations; GBCAL, reaction field energy calculated by GB; GBSOL=GBSUR+GBCAL; GBELE=GBCAL+ELE; GBTOTAL=GBSOL+GAS; $T\Delta S$, T(temperature), ΔS(sum of rotational, translational and vibrational entropies); $\Delta G_{binding}$, total binding energy of the system. All energies are in kcal/mol.

Glycobiology, 2009, 19, 1103-1115

Table S4. MM/PBSA energy component analysis of the interactions of a heparin pentasaccharide with Ig-domains 2-3 of PECAM-1 averaged during 2-4 ns[a].

	Complex		*Receptor*		*Ligand*		**Δ**	
	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
ELE	-4731.55	83.16	-3456.73	81.11	1260.58	20.12	-2535.4	74.86
VDW	-487.6	20.98	-452.69	18.87	4.67	4.85	-39.58	6.9
INT	2112.77	38.13	2097.54	36.06	15.23	10.01	0	0
GAS	-3106.39	88.36	-1811.88	86.81	1280.48	21.13	-2574.99	75.55
PBSUR	72.2	1.01	70.49	0.91	8.17	0.07	-6.46	0.3
PBCAL	-3292.78	73.21	-3549.73	67.5	-2273.59	17.82	2530.54	74.52
PBSOL	-3220.58	72.99	-3479.24	67.48	-2265.41	17.87	2524.08	74.41
PBELE	-8024.33	33.58	-7006.46	27.99	-1013.01	7.83	-4.86	12.04
PBTOT	-6326.97	40.86	-5291.13	38.09	-984.93	8.25	-50.91	7.27
GBSUR	94.69	1.34	92.42	1.21	9.64	0.1	-7.36	0.4
GBCAL	-3273.23	72.16	-3595.96	69.86	-2208.2	17.8	2530.93	73.58
GBSOL	-3178.54	71.98	-3503.54	69.89	-2198.57	17.86	2523.57	73.46
GBELE	-8004.79	26.8	-7052.69	22.93	-947.62	8.1	-4.48	8.52
GBTOT	-6284.93	38.49	-5315.42	36.35	-918.08	8.82	-51.42	5.81
$T\Delta S^{b}$	2093.52	4.49	1992.42	8.06	141.25	0.04	-40.15	
$T\Delta S^{c}$							-61.22	
$\Delta G_{binding\text{-}PBSA}{}^{b}$							-10.76	
$\Delta G_{binding\text{-}GBSA}{}^{b}$							-11.27	
$\Delta G_{binding\text{-}PBSA}{}^{c}$							+10.31	
$\Delta G_{binding\text{-}GBSA}{}^{c}$							+9.81	

a Average over 200 snapshots of the MD simulation trajectory.

b Entropy calculations were based on normal modes analysis using 20 snapshots.

c Entropy calculations were based on quasiharmonic analysis using 20 snapshots.

ELE, non-bonded electrostatic energy; VDW, non-bonded van der Waals energy; INT, bond, angle, dihedral energies; GAS, ELE+VDW+INT; PBSUR, hydrophobic contribution to solvation free energy for PB calculations; PBCAL, reaction field energy calculated by PB; PBSOL=PBSUR+PBCAL; PBELE=PBCAL+ELE; PBTOTAL=PBSOL+GAS; GBSUR, hydrophobic contributions to solvation free energy for GB calculations; GBCAL, reaction field energy calculated by GB; GBSOL=GBSUR+GBCAL; GBELE=GBCAL+ELE; GBTOTAL=GBSOL+GAS; $T\Delta S$, T(temperature); ΔS(sum of rotational, translational and vibrational entropies); $\Delta G_{binding}$, total binding energy of the system. All energies are in kcal/mol.

Glycobiology, 2009, 19, 1103-1115

Table S5. MM/PBSA energy component analysis of the interactions of a heparin pentasaccharide with Ig-domains 2-3 of PECAM-1 averaged during 4-6 ns[a].

	Complex		*Receptor*		*Ligand*		Δ	
	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
ELE	-4713.93	128.46	-3397.42	110.24	1267.91	23.43	-2584.42	91.61
VDW	-456.35	20.26	-437.48	18.97	7.11	4.83	-25.98	4.91
INT	2141.04	39.59	2078.84	37.75	62.19	9.73	0	0
GAS	-3029.24	129.92	-1756.06	110.9	1337.21	24.21	-2610.39	92.92
PBSUR	75.56	1.12	72.88	0.98	8.17	0.11	-5.49	0.44
PBCAL	-3356.68	119.87	-3629.63	104.31	-2281.69	20.58	2554.64	88.15
PBSOL	-3281.12	119.12	-3556.75	103.71	-2273.52	20.65	2549.15	87.85
PBELE	-8070.6	27.34	-7027.05	25.18	-1013.78	7.84	-29.78	9.32
PBTOT	-6310.36	38.99	-5312.8	36.29	-936.31	8.27	-61.25	7.92
GBSUR	99.16	1.49	95.6	1.3	9.63	0.15	-6.07	0.59
GBCAL	-3338.42	122.08	-3686.56	108.9	-2213.76	21	2561.89	86.11
GBSOL	-3239.27	121.09	-3590.96	108.12	-2204.13	21.09	2555.82	85.72
GBELE	-8052.35	23.6	-7083.98	20.63	-945.85	8.19	-22.53	9.48
GBTOT	-6268.51	40.48	-5347.02	36.96	-866.92	8.69	-54.58	9.51
$T\Delta S^{b}$	2090.56	5.36	1989.04	12.41	139.96	0.21	-38.44	
$T\Delta S^{c}$							-65.29	
$\Delta G_{binding\text{-}PBSA}$ [b]							-22.81	
$\Delta G_{binding\text{-}GBSA}$ [b]							-16.14	
$\Delta G_{binding\text{-}PBSA}$ [c]							+4.04	
$\Delta G_{binding\text{-}GBSA}$ [c]							+10.71	

a Average over 200 snapshots of the MD simulation trajectory.

b Entropy calculations were based on normal modes analysis using 20 snapshots.

c Entropy calculations were based on quasiharmonic analysis using 20 snapshots.

ELE, non-bonded electrostatic energy; VDW, non-bonded van der Waals energy; INT, bond, angle, dihedral energies; GAS, ELE+VDW+INT; PBSUR, hydrophobic contribution to solvation free energy for PB calculations; PBCAL, reaction field energy calculated by PB; PBSOL=PBSUR+PBCAL; PBELE=PBCAL+ELE; PBTOTAL=PBSOL+GAS; GBSUR, hydrophobic contributions to solvation free energy for GB calculations; GBCAL, reaction field energy calculated by GB; GBSOL=GBSUR+GBCAL; GBELE=GBCAL+ELE; GBTOTAL=GBSOL+GAS; $T\Delta S$, T(temperature). ΔS(sum of rotational, translational and vibrational entropies); $\Delta G_{binding}$, total binding energy of the system. All energies are in kcal/mol.

Glycobiology, 2009, 19, 1103-1115

Table S6. MM/PBSA energy component analysis of the interactions of a heparin pentasaccharide with Ig-domains 2-3 of PECAM-1 averaged during 6-8 ns[a].

	Complex		*Receptor*		*Ligand*		Δ	
	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
ELE	-4590.15	78.65	-3186.62	90.85	1316.31	29.44	-2719.84	77.44
VDW	-460.77	21.12	-430.33	19.73	1.6	5.01	-32.05	4.54
INT	2128.75	39.04	2073.07	36.53	55.68	10.27	0	0
GAS	-2922.17	85.1	-1543.88	98.71	1373.59	27.8	-2751.88	78
PBSUR	76.78	1.08	74.69	1.09	7.96	0.11	-5.86	0.32
PBCAL	-3464.16	70.5	-3843.49	83.17	-2320.83	25.43	2700.16	75.62
PBSOL	-3387.38	69.91	-3768.81	82.55	-2312.88	25.52	2694.3	75.43
PBELE	-8054.31	30.23	-7030.11	27.16	-1004.52	8.17	-19.68	8.43
PBTOT	-6309.56	42.58	-5312.69	40.48	-939.28	8.22	-57.59	7.47
GBSUR	100.77	1.43	97.99	1.44	9.35	0.15	-6.57	0.42
GBCAL	-3449.63	72.46	-3899.08	83.66	-2256.93	25.89	2706.37	72.39
GBSOL	-3348.86	71.71	-3801.08	82.87	-2247.58	26.01	2699.81	72.14
GBELE	-8039.78	24.25	-7085.7	21.93	-940.62	8.56	-13.46	9.37
GBTOT	-6271.03	39.89	-5344.96	39.25	-873.99	9.15	-52.08	9.02
$T\Delta S^{b}$	2096.65	9.33	1993.95	7.64	140.5	0.61	-37.8	
$T\Delta S^{c}$							-65.69	
$\Delta G_{binding\text{-}PBSA}$ [b]							-19.79	
$\Delta G_{binding\text{-}GBSA}$ [b]							-14.28	
$\Delta G_{binding\text{-}PBSA}$ [c]							+8.1	
$\Delta G_{binding\text{-}GBSA}$ [c]							+13.6	

a Average over 200 snapshots of the MD simulation trajectory.

b Entropy calculations were based on normal modes analysis using 20 snapshots.

c Entropy calculations were based on quasiharmonic analysis using 20 snapshots.

ELE, non-bonded electrostatic energy; VDW, non-bonded van der Waals energy; INT, bond, angle, dihedral energies; GAS, ELE+VDW+INT; PBSUR, hydrophobic contribution to solvation free energy for PB calculations; PBCAL, reaction field energy calculated by PB; PBSOL=PBSUR+PBCAL; PBELE=PBCAL+ELE; PBTOTAL=PBSOL+GAS; GBSUR, hydrophobic contributions to solvation free energy for GB calculations; GBCAL, reaction field energy calculated by GB; GBSOL=GBSUR+GBCAL; GBELE=GBCAL+ELE; GBTOTAL=GBSOL+GAS; TΔS, T (temperature), ΔS (sum of rotational, translational and vibrational entropies); $\Delta G_{binding}$, total binding energy of the system. All energies are in kcal/mol.

Glycobiology, 2009, 19, 1103-1115

Table S7. MM/PBSA energy component analysis of the interactions of a heparin octasaccharide with the extended conformation of Ig-domains 2-3 of PECAM-1 averaged over 8.0 ns[a].

	Complex		*Receptor*		*Ligand*		Δ	
	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
ELE	-4583.62	121.08	-3422.19	119.31	2710.59	52.4	-3872.02	115.16
VDW	-478.04	24.31	-436.94	22.92	10.26	6.59	-51.37	6.53
INT	2681.08	44.38	2710.4	41.69	-29.32	13.89	0	0
GAS	-2380.58	124.92	-1148.73	129.93	2691.53	51.44	-3923.39	115.84
PBSUR	97.97	2.47	92.47	2.44	14.91	0.2	-9.41	0.51
PBCAL	-4347.5	102.69	-3608.89	109.2	-4590.23	48.9	3851.62	112.38
PBSOL	-4249.53	102.88	-3516.42	109.01	-4575.31	49	3842.21	112.13
PBELE	-8931.12	41.68	-7031.08	30.94	-1879.64	12.06	-20.4	17.02
PBTOT	-6630.11	51.42	-4665.15	45.09	-1883.78	11.77	-81.18	13.79
GBSUR	97.97	2.47	92.47	2.44	14.91	0.2	-9.41	0.51
GBCAL	-4276	104.45	-3656.83	111.77	-4462.74	49.01	3843.57	110.21
GBSOL	-4178.03	104.69	-3564.36	111.59	-4447.82	49.11	3834.15	109.97
GBELE	-8859.62	32.41	-7079.02	23.57	-1752.15	12.09	-28.45	13.59
GBTOT	-6558.62	46.55	-4713.09	41.44	-1756.29	12.73	-89.24	11.26
$T\Delta S^{b}$	2157.86	11.1575	1989.453	8.0575	223.4275	0.71	-57.01	
$T\Delta S^{c}$							-74.23	
$\Delta G_{binding\text{-}PBSA}$ [b]							-24.17	
$\Delta G_{binding\text{-}GBSA}$ [b]							-32.23	
$\Delta G_{binding\text{-}PBSA}$ [c]							-6.95	
$\Delta G_{binding\text{-}GBSA}$ [c]							-15.01	

a Average over 800 snapshots of the MD simulation trajectory.

b Entropy calculations were based on normal modes analysis using 80 snapshots.

c Entropy calculations were based on quasiharmonic analysis using 80 snapshots.

ELE, non-bonded electrostatic energy; VDW, non-bonded van der Waals energy; INT, bond, angle, dihedral energies; GAS, ELE+VDW+INT; PBSUR, hydrophobic contribution to solvation free energy for PB calculations; PBCAL, reaction field energy calculated by PB; PBSOL=PBSUR+PBCAL; PBELE=PBCAL+ELE; PBTOTAL=PBSOL+GAS. GBSUR, hydrophobic contributions to solvation free energy for GB calculations; GBCAL, reaction field energy calculated by GB; GBSOL=GBSUR+GBCAL; GBELE=GBCAL+ELE; GBTOTAL=GBSOL+GAS; $T\Delta S$, T(temperature); ΔS(sum of rotational, translational and vibrational entropies); $\Delta G_{binding}$, total binding energy of the system. All energies are in kcal/mol.

Glycobiology, 2009, 19, 1103-1115

Table S8. MM/PBSA energy component analysis of the interactions of a heparin pentasaccharide with the crystal structure of annexin A2 averaged over 8.0 ns[a].

	Complex		*Receptor*		*Ligand*		Δ	
	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
ELE	-12513	106.86	-11091.3	112.97	1479.71	-12513	-2901.43	92.3
VDW	-1346.43	28.18	-1310.09	28.17	-8.51	-1346.43	-27.83	4.78
INT	4653.6	47.41	4610.68	47.28	42.92	4653.6	0	0
GAS	-9205.85	112.71	-7790.71	117.45	1514.13	-9205.85	-2929.27	92.02
PBSUR	92.49	0.92	90.01	0.91	8.12	92.49	-5.64	0.28
PBCAL	-5010.01	99.51	-5648.54	104.6	-2252.1	-5010.01	2890.62	88.28
PBSOL	-4917.52	99.04	-5558.53	104.18	-2243.97	-4917.52	2884.99	88.17
PBELE	-17523	35.16	-16739.8	33.91	-772.39	-17523	-10.81	10.84
PBTOT	-14123.4	52.57	-13349.2	51.4	-729.85	-14123.4	-44.28	9.74
GBSUR	92.49	0.92	90.01	0.91	8.12	92.49	-5.64	0.28
GBCAL	-4949.5	101.39	-5655.85	107.5	-2196.24	-4949.5	2902.58	88.45
GBSOL	-4857.01	100.95	-5565.84	107.09	-2188.12	-4857.01	2896.95	88.34
GBELE	-17462.5	27.87	-16747.2	27.36	-716.53	-17462.5	1.15	8.02
GBTOT	-14062.9	50.1	-13356.6	49.51	-673.99	-14062.9	-32.32	7.15
$T\Delta S^{b}$	3652.46	5.96	3547.76	6.01	142.66	1.98	-36.60	
$T\Delta S^{c}$							-70.56	
$\Delta G_{binding\text{-}PBSA}{}^{b}$							-7.68	
$\Delta G_{binding\text{-}GBSA}{}^{b}$							+4.28	
$\Delta G_{binding\text{-}PBSA}{}^{c}$							+26.28	
$\Delta G_{binding\text{-}GBSA}{}^{c}$							+38.24	

a Average over 800 snapshots of the MD simulation trajectory.

b Entropy calculations were based on normal modes analysis using 80 snapshots.

c Entropy calculations were based on quasiharmonic analysis using 80 snapshots.

ELE, non-bonded electrostatic energy; VDW, non-bonded van der Waals energy; INT, bond, angle, dihedral energies; GAS, ELE+VDW+INT; PBSUR, hydrophobic contribution to solvation free energy for PB calculations; PBCAL, reaction field energy calculated by PB; PBSOL=PBSUR+PBCAL; PBELE=PBCAL+ELE; PBTOTAL=PBSOL+GAS; GBSUR, hydrophobic contributions to solvation free energy for GB calculations; GBCAL, reaction field energy calculated by GB; GBSOL=GBSUR+GBCAL; GBELE=GBCAL+ELE; GBTOTAL=GBSOL+GAS; $T\Delta S$, T(temperature); ΔS(sum of rotational, translational and vibrational entropies); $\Delta G_{binding}$, total binding energy of the system. All energies are in kcal/mol.

7

7 Can Current Force Fields Reproduce Ring Puckering in 2-*O*-Sulfo-α-L-Iduronic Acid? A Molecular Dynamics Simulation Study

Carbohydrate Research 345 (2010) 689–695

Contents lists available at ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

Can current force fields reproduce ring puckering in 2-*O*-sulfo-α-L-iduronic acid? A molecular dynamics simulation study

Neha S. Gandhi [a], Ricardo L. Mancera [a,b,*]

[a] Curtin Health Innovation Research Institute, Western Australian Biomedical Research Institute, School of Biomedical Sciences, Curtin University of Technology, GPO Box U1987 Perth, Western Australia 6845, Australia
[b] School of Pharmacy, Curtin University of Technology, GPO Box U1987 Perth, Western Australia 6845, Australia

ARTICLE INFO

Article history:
Received 6 October 2009
Received in revised form 30 November 2009
Accepted 20 December 2009
Available online 29 December 2009

Keywords:
Iduronic acid
GLYCAM
GROMOS
Molecular dynamics
Puckering

ABSTRACT

The monosaccharide 2-*O*-sulfo-α-L-iduronic acid (IdoA2S) is one of the major components of glycosaminoglycans. The ability of molecular mechanics force fields to reproduce ring-puckering conformational equilibrium is important for the successful prediction of the free energies of interaction of these carbohydrates with proteins. Here we report unconstrained molecular dynamics simulations of IdoA2S monosaccharide that were carried out to investigate the ability of commonly used force fields to reproduce its ring conformational flexibility in aqueous solution. In particular, the distribution of ring conformer populations of IdoA2S was determined. The GROMOS96 force field with the SPC/E water potential can predict successfully the dominant skew-boat to chair conformational transition of the IdoA2S monosaccharide in aqueous solution. On the other hand, the GLYCAM06 force field with the TIP3P water potential sampled transitional conformations between the boat and chair forms. Simulations using the GROMOS96 force field showed no pseudorotational equilibrium fluctuations and hence no inter-conversion between the boat and twist boat ring conformers. Calculations of theoretical proton NMR coupling constants showed that the GROMOS96 force field can predict the skew-boat to chair conformational ratio in good agreement with the experiment, whereas GLYCAM06 shows worse agreement. The omega rotamer distribution about the C5–C6 bond was predicted by both force fields to have torsions around 10°, 190°, and 360°.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

α-L-Iduronic acid (IdoA) and D-glucosamine are the major components of glycosaminoglycans (GAGs) such as heparin.[1] Heparin consists of differently sulfated alternating disaccharide units such as 2-*O*-sulfo-α-L-iduronic acid (IdoA2S) (Fig. 1) which contains an *O*-sulfate group at carbon position 2 and 2-deoxy-2-sulfamido-α-D-glucopyranosyl-6-*O*-sulfate (GlcNS6S) which contains an *N*-sulfate and an *O*-sulfate at carbon positions 2 and 6, respectively. Analysis of the conformations of these sugar residues within heparin has revealed that unsubstituted IdoA residues exist predominantly in the $^4C_1$ or $^1C_4$ chair form in aqueous solution.[2] On the other hand, IdoA2S exists in equilibrium between a number of different conformations, the most important being the chair ($^1C_4$) and skew-boat ($^2S_0$) forms, whereas GlcNS6S prefers the $^4C_1$ chair conformation.[3] In particular, IdoA2S exhibits a distribution between the $^2S_0$ and $^1C_4$ conformational states of 10% and 90%, respectively.[2] This distribution shifts, respectively, to 35% and 65% when IdoA2S is located between two GlcNS6S residues in a trisaccharide.[2] The conformational equilibrium of IdoA and IdoA2S residues depends on the substitution pattern of sulfates and the presence of an adjacent glucosamine residue in heparin chains.[4]

A graphical procedure based on a torsion angle formalism, NMR couplings, and intra-ring hydrogen distances of IdoA ring conformations has been used to investigate the inter-conversion between different conformations of the iduronate residue beyond the canonical chair and skew-boat conformations.[5] It was found that the $^4C_1$ form at the north pole of the conformational space can easily convert to a region between the $^2S_0$ and $^1S_3$ conformations at the equator, whereas inter-conversion from $^1C_4$ at the south pole goes to a region between the $^0S_2$ and $^3S_1$ conformations present at the equatorial region.[5] Ragazzi et al. plotted a projection of the isoenergy contours on a spherical surface for methyl 4-*O*-methyl-α-L-idopyranosyluronic acid 2-sulfate (DMIS) using molecular mechanics calculations in order to correlate the peculiar vicinal proton coupling constants observed in the iduronate ring to the conformations of the ring. However, they constrained one of the Cremer–Pople[6] ring-puckering parameters. It was found that DMIS exists as a mixture of boat and chair ring geometries, and that these interconvert on the submicrosecond time scale.[7]

* Corresponding author. Tel.: +61 8 9266 1017; fax: +61 8 9266 2342.
E-mail address: R.Mancera@curtin.edu.au (R.L. Mancera).

0008-6215/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2009.12.020

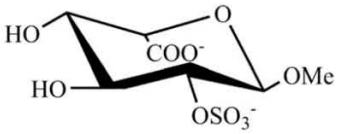


Figure 1. Structure of IdoA2S.

Molecular dynamics (MD) simulations have been used to investigate the conformational dynamics of individual pyranose rings present in heparin in aqueous solution, complemented by experimental NMR data.[8–17] These simulations, which have used force fields such as GLYCAM[18] and GROMOS,[19] have successfully reproduced various conformational and structural properties of GAG polysaccharides such as intra-ring torsions, glycosidic linkage torsion angles, conformational energies, and the conformational equilibrium between the chair and skew-boat forms.[8–12,15–17] Molecular dynamics simulations of IdoA2S-containing oligosaccharides have indicated an important role for intra-molecular hydrogen bonds observed to occur in glycosidic linkages as the driving force for the conformational equilibrium of IdoA2S.[17] This in turn appears to be influenced by the surrounding solvent structure. Combined quantum mechanical studies and MD simulations have also suggested that heparin glycosidic linkages are most susceptible to the influence of ions as well as the solvent.[17] Most of these MD simulations have successfully reproduced the glycosidic torsion angles φ and ψ with the use of ring conformational constraints; however, unacceptable distortions have been observed for simulations of pyranose ring conformations[14] depending on the type of adjacent monosaccharide. Time average restraints have been used to reproduce ring geometry, puckering, and torsion φ, ψ values for undersulfated heparin and heparin[9] as well as five-membered carbohydrate ring systems.[20] The successful representation of the conformational inter-conversion of hexopyranoses using MD simulations without the need for any positional constraints, time average, or harmonic restraints is likely to be dependent on the choice of partial atomic charges, the use of explicit solvent, and/or the presence of counter-ions.

An evaluation of the ability of existing force fields to successfully predict ring-puckering conformational equilibrium has not been reported for the IdoA2S monosaccharide. Here we report our studies of the ring-puckering profile of the IdoA2S monosaccharide using unconstrained MD simulations in aqueous solution using two of the most popular force fields currently in use to simulate GAGs: GROMOS96[19] and GLYCAM06.[21] We determined that simulations with the GROMOS96 force field in combination with the SPC/E[22] water potential were able to successfully predict the chair and skew-boat conformations of IdoA2S with a ratio of 83:17. On the other hand, simulations with the GLYCAM06 force field in combination with the TIP3P[23] water potential tended to sample transitional conformations between the chair and boat forms, with unrealistic wide and rapid fluctuations in the Cremer–Pople ring-puckering parameters.

Another challenging task in the study of conformational properties of carbohydrates is the determination of rotamer populations about the C5–C6 bond (the O6–C6–C5–O5 ω-torsion angle). Three rotamers are typically present: *gauche–trans* (*gt*), *trans–gauche* (*tg*), and *gauche–gauche* (*gg*). The distribution of rotamers is influenced by *gg* effects for IdoA2S and GlcNS6S present in heparin sequences.[24] Our simulations reveal that both the GROMOS96 and the GLYCAM06 force fields were able to reproduce the ω-torsion angle distribution for the IdoA2S monosaccharide in a similar fashion.

2. Results and discussion

2.1. Cremer–Pople ring-puckering parameter θ using GROMOS96

Values of the Cremer–Pople ring-puckering parameter θ for IdoA2S obtained during the initial 500 ps of the simulation using the GROMOS96 force field and starting from the $^2S_0$ skew-boat ring conformation are plotted in Figure 2. After 81 ps, a transition of the iduronate ring to the $^1C_4$ chair form was observed. No further transitions from the chair to any other distinct conformation were observed at all during the subsequent 199.5 ns of the simulations (data not shown). No transitions from this $^1C_4$ chair conformation to another distinct ring conformation were observed during the MD simulations, which is consistent with the studies of the iduronate ring (as a methyl glycoside and as the central residue of a trisaccharide).[15]

Torsional parameters in the GROMOS force field have been shown to lead to the stabilization of the hexopyranose ring in the $^4C_1$ chair conformation in the absence of dihedral-angle restraints;[25] however, in this case we have observed a stabilization in the alternative $^1C_4$ chair conformation. This is not surprising as in D-pyranoses the $^4C_1$ form is usually favored, whereas for L-pyranoses the $^1C_4$ form is usually the most energetically favored conformation. For pyranoses like maltose, conformational inversions of the pyranose ring have been observed within 1 ns simulations but at a higher temperature of 600 K.[26] We were also able to observe conformational transitions from the $^1C_4$ to the $^2S_0$ form in IdoA2S in simulations carried out at 600 K (results shown in Supplementary data), although simulations at such a high temperature would not be practical for any biologically relevant situation.

The Cremer–Pople φ_2 values measured for the $^1C_4$ chair conformation exhibited large variations (results not shown) for IdoA2S, consistent with the previous reports of the iduronate ring (as a methyl glycoside and as the central residue of a trisaccharide).[15] The $^2S_0$ skew-boat ring conformation, observed during the first 81 ps of our simulation, exhibited an average φ_2 value of 145 ± 19° (Fig. 3). This reveals that pseudorotational equilibrium transitions did not occur during the time scale of the simulation and that the conformational space of the $^2S_0$ skew-boat form was the only

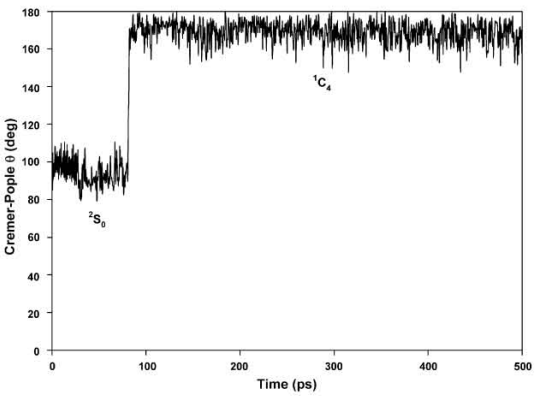


Figure 2. Values of the Cremer–Pople ring-puckering parameter θ for IdoA2S starting from the $^2S_0$ conformation using the GROMOS96 force field. The MD simulations were performed for 200 ns but the puckering parameter is shown only for the initial 500 ps for clarity, as no further changes were observed. Ring $^2S_0$ and $^1C_4$ conformations are characterized by θ values of around 90° and 180°, respectively.

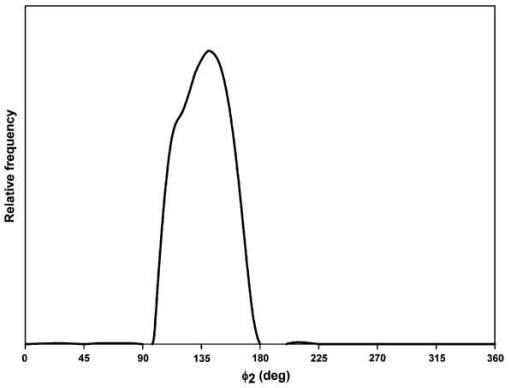


Figure 3. Distribution of the φ_2 angle for the iduronate ring in IdoA2S using the GROMOS96 force field. This represents the conformational space of the $^2S_0$ skew-boat form. These conformations are characterized by average Cremer–Pople φ_2 values of around 150°.

one sampled. This is in agreement with MD simulations of heparin fragments in aqueous solution, where no other contributions from skew-boat pseudorotational transitions have been observed.[13]

2.2. Cremer–Pople ring-puckering parameter θ using GLYCAM06

The conformer populations of IdoA2S were also analyzed for the simulation using the GLYCAM06 force field on the basis of Cremer–Pople ring-puckering angles. Most of the conformers had puckering angles with values of $\theta = 130°$ and $\theta = 60°$ (Fig. 4). The use of the GLYCAM06 force field thus seems to lead to worse agreement with the experiment, as the conformational ring equilibrium is shifted toward transitional conformations rather than the stable $^1C_4$ chair conformation (i.e., $\theta = 180°$). Similar fluctuations were observed if the simulations were started using the $^1C_4$ form. These results show clear differences with the above-described simulations using the GROMOS96 force field, wherein IdoA2S is stabilized in the $^1C_4$ chair conformation after its transition from the starting $^2S_0$ skew-boat conformation. The Cremer–Pople φ_2 values measured for the whole trajectories exhibited large variations for IdoA2S using the GLYCAM06 approach.

As shown in Figure 4, a striking difference between this simulation with the GLYCAM06 force field and the previously described simulation using the GROMOS96 force field is that the Cremer–Pople ring-puckering angles fluctuate largely and rapidly revealing a very rapid inter-conversion from one conformer to the other and suggesting that no conformer is stable enough and that the energy barrier between the conformers is unrealistically low. Our results are not consistent with those of an earlier 2.0 ns MD simulation study of a heparin hexasaccharide in aqueous solution using the GLYCAM93 force field.[13] In that study, the average coupling constants for simulations in the $^1C_4$ form for the three IdoA2S units present in a hexasaccharide showed relative conformational populations of 63%, 58%, and 62%, whereas the average coupling constants for simulations in the $^2S_0$ form showed relative conformer populations of 37%, 42%, and 38%, respectively, without any contribution from skew-boat pseudorotational transitions. The theoretical ^1H–^1H NMR coupling constants were also reported to be successfully reproduced, in excellent agreement with the experimental data.[13] Our results are also inconsistent with those of another 1.0 ns MD simulation study of IdoA2S and IdoA carried out with the GAFF parameter set and AM1-BCC charges. These simulations exhibited conformational transitions from the $^1C_4$ chair to the $^2S_0$ skew-boat form after 599 ps and 99 s for IdoA2S and IdoA, respectively, and those conformations persisted during the remainder of the simulations.[16] Such an inter-conversion from the stable $^1C_4$ chair to the $^2S_0$ skew-boat conformation would appear to be too fast as it is known to be a slow (millisecond) process.[5,27] In addition, these simulations did not exhibit any transitions from the $^2S_0$ and $^4C_1$ starting conformations. Furthermore, no rapid fluctuations of the Cremer–Pople ring-puckering angle were observed, in contrast to our own observations.

It is possible that the accurate reproduction of ring-puckering parameters and the underlying conformer population for IdoA2S using the GLYCAM06 force field may require the use of partial charges derived from various ring conformations, dihedral angle constraints within the ring during geometry optimizations and MD simulations, a different water potential, and/or much longer simulation times. Ring-averaged charge derivations and alternative water potentials such as TIP4P and TIP5P have already been used successfully to improve the distribution of conformer populations of the α-D-arabinofuranoside[28] and methyl β-D-arabinofuranoside[29] rings over the standard GLYCAM approach, which involves parameterization using the TIP3P water potential and a single conformation charge derivation approach.[30]

The successful simulation of the distribution of conformer populations might also be dependent on the choice of the parameters for the sulfate and sulfamate groups. Different parameters for sulfates and sulfamates have been reported.[31] These parameters where derived from gas-phase simulations and showed the presence of large differences in the conformation of the glycosidic link-

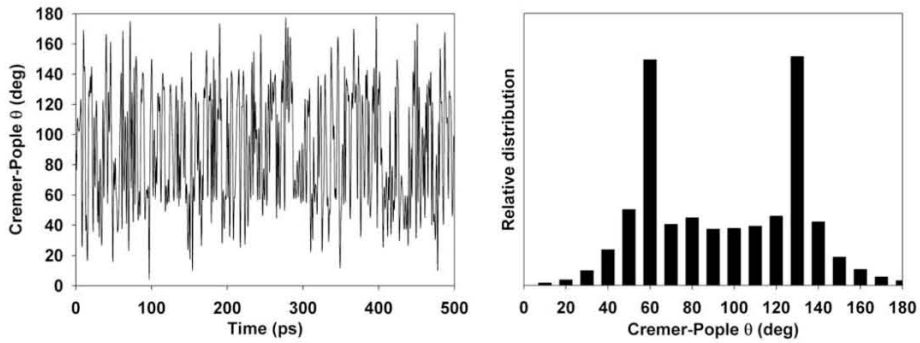


Figure 4. Values of the Cremer–Pople ring-puckering parameter θ for IdoA2S present in IdoA2S using the GLYCAM06 force field. The MD simulations were performed for 200 ns but the puckering parameter is shown only for the initial 500 ps for clarity as the same rapid inter-conversion between conformations was observed throughout the entire simulation. The plot shows a bias toward conformers in the regions $\theta = 130°$ and $\theta = 60°$.

age compared to that observed by NMR spectroscopy.[8] This parameter set has been used to simulate the conformational behavior of sugars in *N*-acetylheparosan, *N*-sulfoheparosan, and undersulfated heparin, but this was investigated using time-averaged restrained molecular dynamics (TARMD).[16] Recently, the same parameter set has been used to perform restrained and unrestrained MD simulations of a heparin tetrasaccharide in an aqueous environment,[32] with the overall shape of the heparin molecule not changing significantly, irrespective of the conformation of the internal IdoA. However, the authors reported only the time evolution of the RMSD of the ring atoms in the monosaccharides rather than ring puckering, so it was not possible to establish with certainty which ring conformations were present.

2.3. Prediction of proton coupling constants using GROMOS96 and GLYCAM06

Theoretical proton coupling constants were derived for the $^{2}S_{0}$ skew-boat and $^{4}C_{1}$ chair conformations. In the case of the simulation using the GROMACS96 force field, since a $^{2}S_{0}$–$^{1}C_{4}$ transition was observed after 81 ps, the $^{2}S_{0}$ coupling constants were averaged over the initial 81 snapshots while the $^{1}C_{4}$ coupling constants were averaged over 419 random snapshots from the remaining 419 ps. The average values of the theoretical coupling constants ($J_{average}$) for the entire ensemble were also calculated to obtain the best fit with the experimental values. For the calculation of the proton coupling constants in the simulation using the GLYCAM06 force field, 500 snapshots were collected from the 200 ns trajectory taking into account the value of the Cremer–Pople angle θ. This was done to determine the type of ring conformation present so that the values of the computed coupling constants could be allocated to the correct conformer due to the frequent ring transitions observed (Fig. 4). The theoretical values of the coupling constants and their MD ensemble averages are reported in Table 1 for the $^{2}S_{0}$ skew-boat and $^{1}C_{4}$ chair conformations. We did not observe any transition to the $^{4}C_{1}$ chair form in either of the simulations. Consequently, and since it has been reported that the inclusion of the corresponding term for this conformation in the Karplus equation does not improve the fitting,[4] it was set to zero in our calculations.

Conformer population ratios were determined by a least square fitting procedure. Significant differences were found depending on the choice of force field. Better agreement with the experiment was obtained with the simulation using the GROMOS96 force field for the conformer population ratios. With this force field IdoA2S was found predominantly in the $^{1}C_{4}$ chair conformation (83%) compared to the $^{2}S_{0}$ skew-boat conformation (17%) in relatively good agreement with the data reported previously obtained by fitting to experimental data using molecular mechanics parameters (chair to skew-boat ratio of 90:10).[4]

Simulations using the GLYCAM06 force field predicted coupling constants (Table 1) and conformer populations in poorer agreement with the experiment. The predicted coupling constants reveal a bias of the GLYCAM06 force field in favor of the $^{1}C_{4}$ boat chair conformation of IdoA2S. It was not possible to reproduce the correct ratio of the $^{2}S_{0}$ skew-boat and $^{1}C_{4}$ chair conformational states because the sampled transitional conformations exhibited wide and rapid fluctuations in the Cremer–Pople ring-puckering parameters (Fig. 4). In the GLYCAM approach, the value of $^{3}J_{H,H}$ reported for the $^{2}S_{0}$ form corresponds to the average of all the boat and skew-boat conformations and, consequently, no fit can be obtained between the experimental and theoretical coupling constants. It is interesting to note here that the theoretical coupling constants calculated for IdoA2S in a MD study of a heparin hexasaccharide[13] using the GLYCAM93 force field[18] were in good agreement with the experimental values. These contrasting results between GLYCAM93 and GLYCAM06 may be attributed to the differences in the force field parameters used for the representation of anomeric carbon atoms (the reducing carbon atoms). These carbon atoms are designated α or β depending on their configuration and in GLYCAM93 they have different atom types and force constants. By contrast, the explicit definition of the anomeric carbon atom was removed in GLYCAM06.

Table 1
Calculated proton NMR coupling constants ($^{3}J_{HH}$, Hz) for IdoA2S and calculated MD ensemble averages

	Experiment[4]	GROMOS96			GLYCAM06		
		$J_{average}$	$^{1}C_{4}$	$^{2}S_{0}$	$J_{average}$	$^{1}C_{4}$	$^{2}S_{0}$
$J_{1,2}$	1.76	3.29	1.75	6.75	1.835	1.16	2.13
$J_{2,3}$	3.34	4.98	2.23	7.87	1.785	1.03	2.22
$J_{3,4}$	3.44	3.69	2.16	5.53	3.231	3.19	2.24
$J_{4,5}$	2.22	1.64	1.23	2.66	1.573	0.66	1.62

$J_{average}$ is the average of the theoretical coupling constants on the entire ensemble.

2.4. Prediction of ω-torsion angle rotamer distributions using GROMOS96 and GLYCAM06

It is generally recognized that the *gauche* effect[33] in carbohydrates is a solvent-dependent phenomenon and that the correct reproduction of the experimental ω-torsion angle rotamer distribution is dependent on the simulation lengths and the explicit solvent models used in MD simulations.[29,34] A number of studies using the GLYCAM approach have shown that the ω-torsion angle displays a preference for *gauche* orientations in glucopyranosides[34] and arabinofuranosides.[28,29]

The resulting ω-torsion angle distributions of IdoA2S obtained from our 200 ns trajectories using the GROMOS96 and GLYCAM06 force fields are shown in Figure 5a and b, respectively. The ω-torsion angle was defined with respect to only one of the oxygen atoms of the COO^{-} group in IdoA2S. The histogram for the simulation with the GROMOS96 force field (Fig. 5a) reveals that the distribution of rotamers is predicted with ω-torsion angle peaks centered at approximately 10° (*gt* rotamer), 190°, and 360°. On the other hand, the corresponding histogram for the simulation with the GLYCAM06 force field (Fig. 5b) reveals that the distribution of rotamers is also predicted with ω-torsion angle peaks centered at approximately 10°, 190°, and 360°. The *gg* rotamer, which dominates the rotamer distribution for GlcNS6S and other pyranosides, is predicted to be less populated in both simulations compared to the *tg* and *gt* rotamers. There are no experimental values available for the coupling constants around C5–C6 for uronic acids; however, the ω-torsion angles around C5–C6 bond of IdoA2S in the heparin NMR structure[3] are ~164° and ~156° for the $^{2}S_{0}$ and $^{1}C_{4}$ ring conformations, respectively. These results suggest that neither force field can predict the rotamer distribution about the C5–C6 bond in IdoA2S in agreement with NMR data for IdoA2S monosaccharide.

The results of our analysis of the ω-torsion angle distribution in IdoA2S do not bear resemblance to those reported using the GLYCAM force field for non-sulfated glucopyranosides such as methyl α-D-glucopyranoside.[34] Frequent short-lived transitions between the *gt* and *tg* rotamers were observed for methyl α-D-glucopyranoside, whereas *gg* was the most dynamically stable rotamer. However, the trajectory for methyl α-D-galactopyranoside displayed infrequent and short-lived transitions to the *gg* rotamer, with the *gt* and *tg* being the most populated rotamers, in contrast to glucopyranose.[34] The distribution of ω-torsional angles reported for gluco- and galactopyranosides carrying $CH_{2}OH$ groups cannot be compared with the distribution of ω-torsional angles for IdoA2S

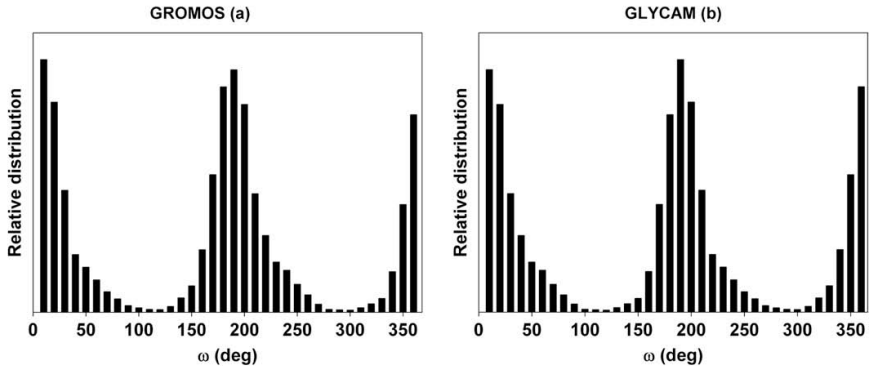


Figure 5. (a) Histogram of the distribution of the ω-angle in IdoA2S using the GROMOS96 force field. (b) Histogram of the distribution of the ω-angle in IdoA2S using the GLYCAM06 force field.

bearing a COO^- group because the ^3J parameters are not reported for the C5–C6 bond in uronic acids.

No apparent correlation between the ω-torsion angle distribution and the Cremer–Pople ring-puckering parameters could be determined for either of the simulations with the GROMOS96 or GLYCAM06 force fields for IdoA2S. This suggests that there is no direct link between the ω-torsion angle rotamer and the ring conformation in IdoA2S.

3. Conclusions

MD simulations of the IdoA2S monosaccharide have been performed in aqueous solution using both the GROMOS96 and GLYCAM06 force fields. The distribution of ring conformations was determined by measuring the Cremer–Pople ring-puckering parameters and computing theoretical proton NMR coupling parameters. Simulations with the GROMOS96 force field in combination with the SPC/E water potential can predict successfully the dominant skew-boat to chair conformational transition in aqueous solution, although no chair to skew-boat transition could be observed during the 200 ns simulation at 300 K. In addition, these simulations showed no pseudorotational equilibrium transitions and hence no inter-conversion between the boat and twist-boat ring conformers. On the other hand, simulations with the GLYCAM06 force field in combination with the TIP3P water potential sampled transitional conformations, with wide and rapid fluctuations in the Cremer–Pople ring-puckering parameters.

To determine the relative proportion of skew-boat and chair conformations, calculations of theoretical proton NMR coupling constants were performed and the results were compared with the reported experimental values. The theoretical coupling constant values determined from simulations using the GROMOS96 force field resulted in a prediction of a 17:83 skew-boat to chair ratio, in relative good agreement with the experimental ratio of 10:90. On the other hand, predictions using the GLYCAM06 force field resulted in significantly worse agreement.

The ω-torsional angle rotamer distribution about the C5–C6 bond in IdoA2S was predicted by both force fields to have values around 10°, 190°, and 360°.

The values of the ring-puckering parameters for IdoA2S predicted with simulations using the GLYCAM06 force field suggest that a ring-average charge derivation procedure, the use of alternative water potentials, and/or the derivation of new parameters for sulfates and sulfamates may be required for the correct representation of the conformational dynamics of this carbohydrate molecule.

Long simulation times of at least 200 ns combined with efficient conformational search methods such as replica exchange MD (REMD) have been reported to be required to account for the inherent flexibility of carbohydrates.[35] Ab initio metadynamics has been used to compute the conformational free energy landscape of β-D-glucopyranose as a function of the Cartesian puckering coordinates.[36] A metadynamics calculation of the puckering free energy surface of a glucuronic acid ring using the GROMOS G45A4 force field suggests that only polar coordinates should be considered as a proper set of collective variables when studying the puckered conformations of ring structures and the free energy landscape in contrast to the Cartesian coordinates.[37] The metadynamics technique has also been applied to study the conformational free energy surfaces and to reproduce ω-torsions of methyl α-D-glucopyranoside and methyl α-D-galactopyranoside.[38] A new puckering coordinate system can be used to describe the pucker of an *N*-member monocyclic ring using N-3 parameters (triangular decomposition)[39] as an alternative over Cremer–Pople[6] and Bercés et al. systems.[40] These methods might be useful over conventional MD simulation methods to accurately model the ring-puckering dynamics of sulfated pyranoses such as IdoA2S.

It appears that the use of a Karplus equation specific to IdoA2S may be necessary. The Karplus equation has been parameterized for sulfated GAGs based on proton–proton and proton–carbon spin–spin coupling constants ($^3J_{H-C-C-H}$).[41,42] Theoretical calculations using DFT with the B3LYP functional and a 6-311++G** basis set were reported using the parameterized Karplus equation.[42,43] However, the energies from these DFT computations indicated the existence of a major contribution from the $^1C_4$ form as well as a small contribution from the $^4C_1$ form to the conformational equilibrium of IdoA2S, while the $^2S_0$ form did not contribute significantly, in contrast to experiment.

Nonetheless the successful simulation of the ring conformational equilibrium of IdoA2S in aqueous solution using the GROMOS96 force field is important because both the $^1C_4$ chair and $^2S_0$ skew-boat conformations have been observed in GAGs in the crystal structures of their complexes with proteins. Consequently, this force field may be better suited to model GAG–protein interactions. However, our results provide strong evidence for the need of new parameterization for GAGs that can reproduce accurately their conformational equilibrium. Current force fields have been parameterized to reproduce the stable $^4C_1$ conformation as it is the only one populated with unfunctionalized D-hexopyranoses and, consequently, are unable to describe ring puckering.

4. Experimental

The structure of the IdoA2S monosaccharide in the skew-boat conformation ($^2S_0$) was extracted from the NMR structure of a heparin dodecasaccharide (PDB code 1HPN).[3] This structure was then submitted to the Dundee PRODRG2 server[44] to generate initial geometries and topologies for the simulations.

A first set of MD simulations was performed using the GROMOS96 force field in the GROMACS 3.3 simulation package.[45] Löwdin atomic charges for IdoA2S were derived using a HF/6-31G** basis set, as reported earlier.[10,11] Additionally, improper dihedral angles imposed to prevent $^1C_4$ to $^2S_0$ transitions as earlier reported[10] were not used so that the unconstrained dynamics of the pyranose ring system could be investigated. The SPC/E water potential has been reported to represent better the conformational profile of heparin compared to the SPC water potential.[10] Consequently the structure of IdoA2S reported in this work was solvated with 1436 SPC/E[22] water molecules in a cubic box using periodic boundary conditions. Placement of sodium ions to neutralize the charge in the system has been reported to be necessary to successfully reproduce torsion angles in agreement with the experimental data.[10] Consequently the module *addions* in AMBER 9.0[46] was used to place sodium cations close to each negatively charged group in the saccharide. Electrostatic interactions were calculated using the Particle Mesh Ewald (PME) method[47] using a fourth-order spline for interpolation. A cutoff of 9.0 Å was set for non-bonded interactions. The temperature was set to 300 K and the pressure to 1 atm. The temperature and pressure were maintained constant using the Berendsen thermostat[48] and the Berendsen barostat,[48] with coupling constants of $\tau = 0.1$ and 0.5 ps, respectively. Translational and rotational motions of the center of mass of the simulation box were removed periodically throughout the simulations. Initially the system was subjected to energy minimization of the water environment followed by minimization of the whole system. Subsequently simulated annealing was used to gently heat up the system from 50 K to the desired temperature of 300 K in steps of 5 ps by increasing the reference temperature by 50 K every step. A time step of 1.0 fs was used in all simulations without any bond or angle constraints. MD simulations were then conducted in the NPT ensemble for 200 ns. The data were collected every 1.0 ps.

A second set of simulations of the IdoA2S monosaccharide was performed using the Sander module in AMBER 9.0 (with the PARM 94 parameter set).[46] The SHAKE algorithm was applied to all bonds with hydrogen atoms. A cubic box consisting of 1318 TIP3P[23] water molecules was added to solvate the molecule, keeping a minimum distance of 12.0 Å between each face of the box and the solute. Net charges in the monosaccharide were neutralized by adding an appropriate number of sodium counter-ions. The GLYCAM06 force field[21] augmented with non-bonded parameters for sulfates and sulfamates was used for these simulations.[49,50] The SCNB and SCEE scaling parameters (unit scale factor for all 1–4 non-bonded electrostatic and van der Waals interactions) were both set to unity in accordance with the GLYCAM force field parameterization approach[34] to avoid potential imbalances in the internal energies of five- and six-member intra-molecular hydrogen bonds. A 9.0 Å atom-based cutoff was applied to all van der Waals interactions. A constant temperature of 300 K was maintained using the Berendsen thermostat[48] with a time constant of 0.1 ps. All MD simulations were carried out in the NPT ensemble at a pressure of 1 atm using the Berendsen barostat.[48] Translational and rotational motions of the center of mass of the simulation box were removed periodically through out the simulations. A 1.0 fs time step was used throughout. MD simulations were preceded by 1000 cycles of steepest descent energy minimization. The system was then slowly heated from 5 to the desired temperature of 300 K and allowed to equilibrate at this temperature for 100 ps. All subsequent simulations were performed using the PME method[47] for long-range electrostatic interactions,[47] using a 1.0 Å grid spacing and a fourth-order spline for interpolation. The simulations were run for 200 ns under NPT conditions. The data were collected every 0.5 ps.

To carry out these MD simulations with the GLYCAM06 force field, partial atomic charges for IdoA2S were obtained using the restricted electrostatic potential (RESP) method[51] using R.E.D.-III tools[52] with a restraint weight of 0.01.[53] These charges for IdoA2S are reported in Supplementary data S2. All GLYCAM charges are developed from a thermally derived ensemble of conformations from long simulations performed in the presence of explicit solvent to represent the average behavior of the molecule in solution.[30] Consequently an initial set of partial charges was derived using the RESP approach. After 50 ns simulations, conformations of IdoA2S obtained from the simulation were subjected to a full quantum mechanics geometry optimization with a HF/6-31G* basis set using GAUSSIAN 03.[54] The multi-conformation RESP charge fitting approach was used to derive charges with 100 conformations from random snapshots and the final charges were obtained as an average.[21] Aliphatic protons were set to have a zero net charge to ensure compatibility between this work and the GLYCAM06 force field.

The Cremer–Pople[6] azimuthal angles θ and φ_2 for the pyranose ring of IdoA2S was measured by selecting atoms in the orders O5, C1, C2, C3, C4, and C5.[55] With this convention, the $^2S_0$ skew-boat has a value of $\theta = 90°$ with the parameter pseudorotation phase angle $\varphi_2 = 150°$ indicating the shape of the boat or skew whereas the more stable $^1C_4$ form has a value of $\theta = 180°$. In particular, a value of $\varphi_2 = 270°$ has been reported[7] for the $^2S_0$ form arising from the different schemes for the selection of atoms, that is, C1, C2, ..., O5. Initial $^2S_0$ skew-boat conformations were used for all MD simulations to try to determine ring-puckering parameters for this conformation. This was done because the $^1C_4$ to $^2S_0$ inter-conversion is known to occur in the millisecond range[5,13,27] and thus could not be expected to be sampled by these simulations. The program MDXVU[56] was used to analyze the Cremer–Pople ring-puckering parameters θ and φ_2.

The MD simulation trajectories were also analyzed by performing a calculation of NMR proton coupling constants using the Altona and Haasnoot formalism[57] and comparing them with the experimental values reported for IdoA2S.[4] The sum of the squares of the differences between each of the experimentally observed coupling constants and a weighted average of the theoretical coupling constants was minimized. For IdoA2S, the weighted average of the theoretical constants was represented by the Karplus formula

$$aJ_{x,y}\,{}^1C_4 + bJ_{x,y}\,{}^2S_0 + cJ_{x,y}\,{}^4C_1 \quad (1)$$

where $J_{x,y}$ is the theoretical value for each of the coupling constants $J_{1,2}$, $J_{2,3}$, $J_{3,4}$, and $J_{4,5}$.

The sum of the squares of the differences was minimized by varying parameters a, b, and c subject to the constraints that their sum should equal 1 and that each constant can only take a value $\geqslant 0$ but $\leqslant 1$. The percentages of each conformation reported are the parameters a, b, and c expressed as a percentage of their sum, which totals 1.

Acknowledgments

N.S.G. is grateful for the award of an Endeavour International Postgraduate Research Studentship. We also gratefully acknowledge the Interactive Virtual Environments Centre (iVEC) of Western Australia and the National Computational Infrastructure (NCI) Facility for access to high-performance computing. The

authors thank Laercio Pol-Fachin (Federal University of Rio Grande do Sul, Brazil) for useful discussions and for providing some of the parameters for our calculations. We also acknowledge Professor Francois Dupradeau (Université de Picardie, Jules Verne, France) for his help with R.E.D.-III tools.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.carres.2009.12.020.

References

1. Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Hart, G. W.; Etzler, M. E. *Essentials of Glycobiology*, 2nd ed.; CSHL Press: New York, 2009. p. 784.
2. Ferro, D. R.; Provasoli, A.; Ragazzi, M.; Casu, B.; Torri, G.; Bossennec, V.; Perly, B.; Sinay, P.; Petitou, M.; Choay, J. *Carbohydr. Res.* **1990**, *195*, 157–167.
3. Mulloy, B.; Forster, M. J.; Jones, C.; Davies, D. B. *Biochem. J.* **1993**, *293*, 849–858.
4. Ferro, D. R.; Provasoli, A.; Ragazzi, M.; Torri, G.; Casu, B.; Gatti, G.; Jacquinet, J. C.; Sinay, P.; Petitou, M.; Choay, J. *J. Am. Chem. Soc.* **1986**, *108*, 6773–6778.
5. Ernst, S.; Venkataraman, G.; Sasisekharan, V.; Langer, R.; Cooney, C. L.; Sasisekharan, R. *J. Am. Chem. Soc.* **1998**, *120*, 2099–2107.
6. Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358.
7. Massimo, R.; Dino, R. F.; Augusto, P. *J. Comput. Chem.* **1986**, *7*, 105–112.
8. Jin, L.; Barran, P. E.; Deakin, J. A.; Lyon, M.; Uhrin, D. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3464–3471.
9. Zhang, Z.; McCallum, S. A.; Xie, J.; Nieto, L.; Corzana, F.; Jiménez-Barbero, J. S.; Chen, M.; Liu, J.; Linhardt, R. J. *J. Am. Chem. Soc.* **2008**, *130*, 12998–13007.
10. Verli, H.; Guimarães, J. A. *Carbohydr. Res.* **2004**, *339*, 281–290.
11. Becker, C. F.; Guimarães, J. A.; Verli, H. *Carbohydr. Res.* **2005**, *340*, 1499.
12. Mikhailov, D.; Mayo, K. H.; Vlahov, I. R.; Toida, T.; Pervin, A.; Linhardt, R. J. *Biochem. J.* **1996**, *318*, 93–102.
13. Angulo, J.; Nieto, P. M.; Martín-Lomas, M. *Chem. Commun.* **2003**, 1512–1513.
14. Mikhailov, D.; Linhardt, R. J.; Mayo, K. H. *Biochem. J.* **1997**, *328*, 51–61.
15. Forster, M. J.; Mulloy, B. *Biopolymers* **1993**, *33*, 575–588.
16. Murphy, K. J.; McLay, N.; Pye, D. A. *J. Am. Chem. Soc.* **2008**, *130*, 12435–12444.
17. Pol-Fachin, L.; Verli, H. *Carbohydr. Res.* **2008**, *343*, 1435–1445.
18. Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832–3846.
19. van Gunsteren, W.; Billeter, S.; Eising, A.; Hünenberger, P.; Krüger, P.; Mark, A.; Scott, W.; Tironi, I. *Biomolecular Simulations: the GROMOS96 Manual and User Guide*, 1996.
20. Pieter, M. S. H.; Francisco, C.; Stefaan, D.; Dirk, A. T.; Koen, A.; Josê, C. M. *J. Comput. Chem.*, in press.
21. Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622–655.
22. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
23. William, L. J.; Jayaraman, C.; Jeffry, D. M.; Roger, W. I.; Michael, L. K. *J. Chem. Phys.* **1983**, *79*, 926–935.
24. Gatti, G.; Casu, B.; Hamer, G. K.; Perlin, A. S. *Macromolecules* **2002**, *12*, 1001–1007.
25. Lins, R. D.; Hünenberger, P. H. *J. Comput. Chem.* **2005**, *26*, 1400–1412.
26. Ott, K.-H.; Meyer, B. *J. Comput. Chem.* **1996**, *17*, 1068–1084.
27. Mulloy, B.; Forster, M. J. *Glycobiology* **2000**, *10*, 1147–1156.
28. Seo, M.; Castillo, N.; Ganzynkowicz, R.; Daniels, C. R.; Woods, R. J.; Lowary, T. L.; Roy, P.-N. *J. Chem. Theory Comput.* **2008**, *4*, 184–191.
29. Taha, H. A.; Castillo, N.; Roy, P.-N.; Lowary, T. L. *J. Chem. Theory Comput.* **2009**, *5*, 430–438.
30. Basma, M.; Sundara, S.; Çalgan, D.; Vernali, T.; Woods, R. J. *J. Comput. Chem.* **2001**, *22*, 1125–1137.
31. Jin, L. Ph.D. Thesis, University of Edinburgh, 2007.
32. Jin, L.; Hricovini, M.; Deakin, J. A.; Lyon, M.; Uhrin, D. *Glycobiology* **2009**, cwp105.
33. Wolfe, S. *Acc. Chem. Res.* **2002**, *5*, 102–111.
34. Kirschner, K. N.; Woods, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10541–10545.
35. Yongye, A. B.; Gonzalez-Outeiriño, J.; Glushka, J.; Schultheis, V.; Woods, R. J. *Biochemistry* **2008**, *47*, 12493–12514.
36. Biarnes, X.; Ardevol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2007**, *129*, 10686–10693.
37. Sega, M.; Autieri, E.; Pederiva, F. *J. Chem. Phys.* **2009**, *130*, 225102.
38. Spiwok, V.; Tvaroska, I. *Carbohydr. Res.* **2009**, *344*, 1575–1581.
39. Hill, A. D.; Reilly, P. J. *J. Chem. Inf. Model.* **2007**, *47*, 1031–1035.
40. Bêrces, A.; Whitfield, D. M.; Nukada, T. *Tetrahedron* **2001**, *57*, 477–491.
41. Tvaroska, I.; Hricovíni, M.; Petráková, E. *Carbohydr. Res.* **1989**, *189*, 359–362.
42. Hricovíni, M. *Carbohydr. Res.* **2006**, *341*, 2575–2580.
43. Hricovíni, M.; Bízik, F. *Carbohydr. Res.* **2007**, *342*, 779–783.
44. Schüttelkopf, A. W.; Van Aalten, D. M. F. *Acta Crystallogr., Sect. D* **2004**, *60*, 1355–1363.
45. van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A.; Feenstra, K. A.; van Drunen, R. GROMACS *User Manual*, 3.3; Department of Biophysical Chemistry, University of Groningen: Groningen, The Netherlands, 2006.
46. Case, D. A.; Darden, T. A.; Cheatham I, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. AMBER *9*, University of California, San Francisco, 2006.
47. Tom, D.; Darrin, Y.; Lee, P. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
48. Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F. v.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
49. Huige, C. J. M.; Altona, C. *J. Comput. Chem.* **1995**, *16*, 56–79.
50. Yongye, A.; Tessier, M.; Kawatkar, S.; Kirschner, K. N.; Tschampel, S.; Woods, R. J. Glycam 0f.dat: Woods Group (2005–2009). GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia, Athens, GA, 2008.
51. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
52. Pigache, A.; Cieplak, P.; Dupradeau, F. Y. In *Automatic and Highly Reproducible RESP and ESP Charge Derivation: Application to the Development of Programs* RED *and* XRED, 227th ACS National Meeting, Anaheim, CA, USA, Mar 28–Apr 1; Anaheim, CA, USA, 2004.
53. Woods, R. J.; Chappelle, R. *J. Mol. Struct.: THEOCHEM* **2000**, *527*, 149–156.
54. Frisch, M. J. T. G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A. Jr.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. GAUSSIAN 03, Revision C. 02, Gaussian, Gaussian, Inc.: Pittsburgh, PA, 2004.
55. Jeffrey, G. A.; Yates, J. H. *Carbohydr. Res.* **1979**, *74*, 319–322.
56. Forster, M. J. MDXVU-Molecular Dynamics X11 Viewer. 0.95, 2005.
57. Haasnoot, C. A. G.; de Leeuw, F. A. A. M.; Altona, C. *Tetrahedron* **1980**, *36*, 2783–2792.

Carbohydr Res, 2010, 345, 689-695

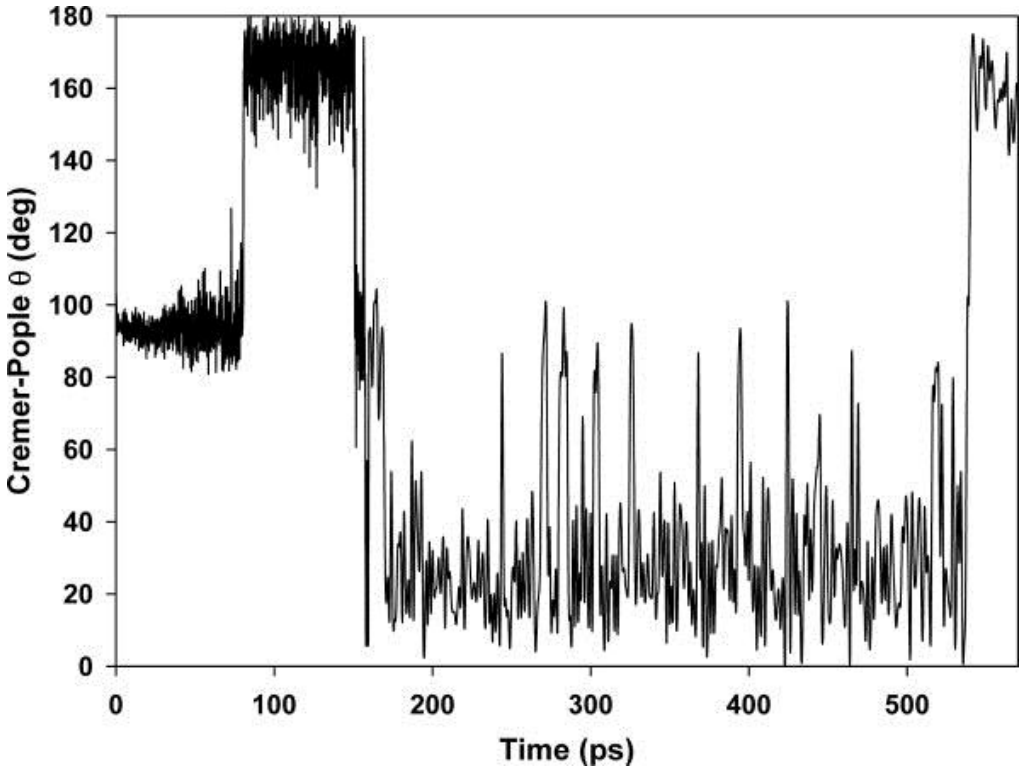
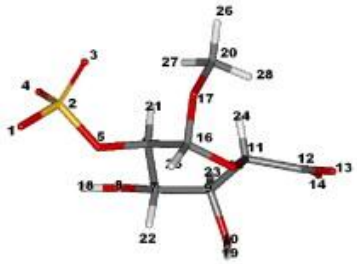


Figure S1. Values of the Cremer-Pople ring puckering parameter Θ for IdoA2S starting from the $^{2}S_{0}$ conformation using the GROMOS96 force field at 600 K. Ring $^{2}S_{0}$ and $^{1}C_{4}$ conformations are characterised by Θ values of around 90° and 180°, respectively. Only the first 570 ps are plotted, showing fluctuations between the $^{2}S_{0}$, $^{1}C_{4}$ and $^{4}C_{1}$ conformations.

Carbohydr Res, 2010, 345, 689-695



Atom No.	*Atom Name*	*Atom type*	*Charge*
1	OAN	O2	-0.6827
2	SAM	S	1.4252
3	OAO	O2	-0.6827
4	OAP	O2	-0.6827
5	O2	OS	-0.5066
6	C2	CG	0.1647
7	C3	CG	0.1761
8	O3	OH	-0.3875
9	C4	CG	0.1826
10	O4	OH	-0.3870
11	C5	CG	0.2276
12	C6	CG	0.2174
13	OAL	O2	-0.6366
14	O6	O2	-0.6366
15	O5	OS	-0.3354
16	C1	CG	0.2780
17	O1	OS	-0.3532
18	H1	HO	0.2107
19	H2	HO	0.2107
20	C3	CG	0.1980
21	H3	H1	0.0000
22	H4	H1	0.0000
23	H5	H1	0.0000
24	H6	H1	0.0000
25	H7	H1	0.0000
26	H8	H1	0.0000
27	H9	H1	0.0000
28	H10	H1	0.0000

Figure S2. RESP charges for IdoA2S labelled according to atom number.

8 Epilogue

8.1 Epilogue: Current challenges and future prospects

Several experimental methods can be used to study GAG-protein interactions, such as affinity chromatography, analytical ultracentrifugation, electrophoretic mobility shift assays, competition experiments, mass spectrometry-based approaches, isothermal titration calorimetry and surface plasmon resonance [1]. Molecular modelling and computational chemistry approaches to study sulphated GAG-protein interactions can provide in principle complementary detailed structural and thermodynamic information. These methods face, however, unique challenges because of the high conformational flexibility and negative charge of GAGs, the indispensability of water-mediated interactions, the lack of parameters for their molecular modelling and simulation, and the scarce availability of structural data on GAG-protein complexes. Although there are some general computational approaches applicable for the identification of shallow, solvent-accessible heparin binding sites on proteins [2-4], these methods experience difficulties as GAG-binding sites are not conserved among cytokines or cell adhesion molecules [4, 5]. Robust molecular modelling methods have yet to be developed for the full characterisation of GAG-protein recognition.

Only a handful of computational molecular docking approaches, such as that implemented in AutoDock, have been applied with some level of success to the prediction and analysis of heparin/HS-protein interactions [4, 6-8]. The extensive conformational flexibility and the presence of large negative charges in GAGs pose serious challenges for molecular docking simulations. Most current docking algorithms cannot efficiently sample more than 32 rotatable bonds [9] and, therefore, only relatively short GAGs (up to four saccharides long) have been so far docked reliably without applying undesirable conformational constraints. The orientations of reducing and non-reducing termini and their high symmetry are also challenges for currently available docking methods. Furthermore, due to the conformational flexibility of IdoA2S (i.e. $^{1}C_{4}$ or $^{2}S_{0}$), two different docking calculations are needed, one for

each conformation. AutoDock 4.2 allows the exploration of ring flexibility by converting ring systems into their corresponding acyclic forms by breaking a bond [10]. In this method, the ligand is treated by means of special (G, glue) atom types. Ring opening can increase dramatically the total number of rotatable bonds, but it requires significantly longer calculation times and the large number of G atoms required makes this method a less desirable choice for docking GAGs with inherent flexibility. Similarly to the genetic algorithm (GA) and Lamarckian genetic algorithm (LGA) in Autodock, the programs SHAPE [11] and GLYGAL (GLYcosidic bonds Genetic Algorithm) [12, 13] implement a GA approach for the conformational search of oligosaccharides using MM3 and MM4 force fields, respectively, but none of them have been validated for sGAGs. The precursor MM2 force field has parameters for sulphates and carbohydrates [14] and, therefore, these programs might perform well in conformational searches of heparin/HS fragments.

Before performing molecular docking simulations it is essential to take into consideration the oligomeric states of the GAG binding proteins. The protein/receptor may exist in oligomeric states, resulting in more than one heparin-binding site and requiring two sulphated regions with a non-sulphated 'spacer' between them for maximum affinity, as in the case of CXCL-8. To be able to simulate larger GAGs, it appears that a first step involving fragment-based docking (i.e. docking of mono and disaccharides) may be needed to obtain clues about possible binding poses, to be followed by the refinement of these predicted binding poses with more sophisticated methods, such as molecular dynamics simulations of the predicted GAG-protein complexes in aqueous solution.

The reliability of predictions of binding affinity using scoring functions or free energy methods has been limited by the complexity and heterogeneity of GAGs. For example, the majority of crystal structures contain fully sulphated heparin oligosaccharides and, hence, the results of docking simulations of heparin fragments cannot be validated with experimental

data wherein physiological GAGs (typically HS, which is less sulphated than heparin) have been used. The difference in predicted and experimental affinity is thus attributed to the size and sulphation patterns of HS. Furthermore, HS displays an extended bent conformation that is significantly distinct from that of heparin, indicating that HS may interact differently with proteins compared to heparin [15]. In relatively recent years two scoring functions have been trained on a dataset of non-charged saccharides [16, 17]. However, these scoring functions do not consider parameters for sulphate moieties and therefore cannot yet be used for docking simulations of GAGs.

The docking packages can are capable of reproducing 70–80% of the crystallographic binding modes to within 2 Å root mean square deviation (RMSD) across a diverse set of protein targets. Despite this, scoring functions perform poorly in predicting the absolute free energies of binding [18]. A docking program may identify the lowest energy conformation but it will fail to differentiate reliably between nanomolar and micromolar ligands leaving significant room for improvements in scoring functions. Furthermore, the inherent inaccuracy of the experimental data remains the culprit to the limited accuracy in predictions of binding affinity through docking. The training of scoring function would benefit from use of a larger and more structurally and thermodynamically diverse set of complexes.

AutoDock also experiences difficulties from a conformational point of view with the docking of long oligosaccharides given the large number of flexible glycosidic linkages. In these cases it is thus recommended to fix the glycosidic linkages first and allow flexibility in the sulphates, hydroxyls and methyl groups [19]. The best predicted binding pose can be subsequently investigated using molecular dynamics simulations in explicit water with no restraints applied. An alternative protocol is to use simulated annealing, as has been reported for the docking of substrates to amylase [20, 21], wherein the maximal rotation per step for

the exocyclic torsion angles can be reduced by a 0.9875 factor per cycle from an initial value of 15°, and a very small translation (i.e. step of 0.2 Å) is allowed.

In GAGs, usually only the recognition unit binds to the protein and the rest of the structure makes favourable interactions with water. Fundamentally, molecular docking (in the absence of water) tries to find the conformation which can make maximum favourable contacts with the protein. Therefore, docking in the presence of water or explicit solvent MD simulation approaches can help to further probe the significance of water-mediated interactions in GAG-protein interfaces. In a recent study, AutoDock 3 outperformed MOE, eHiTS and FlexX in the prediction of eleven GAG-protein complexes in the presence of crystallographically-determined water molecules, although it experienced difficulties in quantitatively reproducing experimental data on the specificity of heparin/HS disaccharides binding to CXCL-8 [22]. A correlation between experimental data and the calculated docking energies could not be established, similarly to the findings reported in this thesis (Chapter 4). This is likely to be attributable to poor clustering of binding modes and the inability to search effectively the binding modes of GAG oligosaccharides in a relatively large binding site comprising many positively charged residues (resulting in very high negative values for the free energies of binding). Although methods such as GRID have been applied to investigate the role of water in carbohydrate recognition by a bacterial enterotoxin [23], the placement of water molecules in the absence of experimental X-ray structures remains a challenging task for molecular docking.

A variety of implicit solvation models which take into account the role of water in carbohydrate binding, such as the Generalised Born (MM/GBSA) and the Poisson–Boltzmann (MM/PBSA) approaches have been applied relatively successfully to study carbohydrate recognition by lectins [24-26] and GAG-protein interactions (Chapter 6); however, these methods do not explicitly derive the locations of structurally conserved water

molecules [27]. Furthermore, the reorganisation of water molecules upon ligand binding greatly affects the entropic component of the free energy of binding, which is not explicitly considered by continuum solvent approaches such as MM/PB(GB)SA (reviewed in [28]). Large deviations in the calculated MM/PB(GB)SA energies are attributed to problems associated with both the electrostatics and the entropy calculations using normal modes or quasiharmonic methods [29]. In particular, the MM/PBSA approach does not work well when calculating the absolute free energy of binding of charged ligands (see Chapter 6) compared to neutral ones [30]. This is due to the fact that the polarisation of charged molecules at the binding interface is quite different to that in neutral molecules. In the MM/PBSA method, the electrostatic energy is split into three components: the average MM electrostatic interaction with the charged ligand, the PB (Poisson-Boltzmann) contribution due to the polarisation of the surrounding solvent, and the electrostatic energy due to the presence of the ligand in solution. This approach provides good estimates of the electrostatics and solvation in a neutral medium, in the absence of any net charges. However, in the presence of a heavily charged ligand, this approximation tends to fail as the solute and solvent electrostatic energy components tend to dominate the energy components. While the MM/PBSA method uses a value for the dielectric constant (ε=4) which should account for highly charged binding interfaces (giving rise to strong electrostatic interactions formed between the protein and a carbohydrate molecule), it is insufficient for this particular system. For example, the calculation of solvation free energies using the MM/PBSA method gives estimates almost twice as large as the actual value. Furthermore, as in the case of heparanase discussed in Chapter 6, the AMBER force field in conjunction with GLYCAM parameters were used with the vdW and electrostatics scaling factors set according to the protein force field, which in the case of sGAGs would overestimate electrostatic interactions and forces by a factor ~2. The above problems associated with the electrostatics (charged binding sites) can be overcome using a suitably calibrated solute dielectric constant to calculate the polar

solvation energies. Efforts are being directed at improving MD sampling techniques, solvation free energy calculations using the PBSA approach, and entropy calculations will help to address the limitations of continuum solvent methods [31-35].

Several studies have found strong inter-connections between specific structural properties of GAGs, such as ring puckering, conformational equilibria and inter-residue linkages (glycosidic bonds) and the dynamical behaviour of the surrounding solvent [36-38]. These studies have generally used molecular mechanics force fields, such as GLYCAM, CHARMM, GROMOS or OPLS-AA with Scaling of Electrostatic Interactions (SEI). The advantages of these methods are their low computational costs, transferability or generality for extension to carbohydrates outside of the original parameterisation scheme (or to glycomimetics), and compatibility with other biomolecules and solvent models [39]. However, the consistency of the properties measured by these simulations depends strongly on the force fields (Chapter 7) and requires simulation times of hundreds of nanoseconds to achieve convergence. Currently, standard MD and advanced sampling techniques can successfully simulate the conformational space and NMR vicinal couplings of iduronic acid and various hexopyranoside monosaccharides [40-42], as well as the α(1→4) β(1→3) and β(1→4) glycosidic linkage conformations in hyaluronan and chondroitin sulphate [43, 44]. While the puckering landscapes of β-D-glucuronic acid (GlcA) in water have been studied [40, 45, 46], future work should be directed at performing MD simulations of β(1→4) D-glucuronic acid-based disaccharides (HS), such as those described in Chapter **Error! Reference source not found.**, to perform a complete scan of their conformational space including their glycosidic torsions.

A number of carbohydrate-protein systems have been studied using MD simulations [36, 47, 48]. GLYCAM in conjunction with the AMBER force field are used with the Lennard–Jones (*scnb*) and electrostatics (*scee*) scaling factors set according to the protein force field. (i.e.

2.0 and 1.2 for *scnb* and *scee*, respectively). However, use of these default scaling parameters degrades the accuracy of the predicted rotamer/conformer populations for oligosaccharides. Accurate modelling of the conformational equilibrium of hexopyranoses has been achieved with the GLYCAM force field when 1–4 interactions are not rescaled (i.e. *scnb* and *scee* scaling factors are set to unity) [37, 41, 42]. The discrepancy of scaling parameters has now been resolved in AMBER 11 [49], which allows mixed scaling of the *scnb* and *scee* parameters for simulations of systems containing both carbohydrates and proteins.

Previous versions of GLYCAM and AMBER (before AMBER10) lacked parameters for sulphates and sulphamates. Therefore, parameters for sulphates and sulphamates were approximated according to Huige and Altona [50], and some parameters for sulphate groups that were not available from the work of Huige and Altona were used to be approximated by those for phosphates from Parm99. GLYCAM06 now provides parameters and charges for *O*-sulphation [51] but lacks parameters for *N*-sulphation, making it of limited use for simulations of heparin oligosaccharides. The *N*-sulphated monosaccharide would require the derivation of ensemble-averaged partial charges using the restricted electrostatic potential (RESP) method [52, 53]. The derivation of charges would involved optimising all of the geometries of the monosaccharide at the B3LYP/cc-pVTZ or HF/6-31G* quantum level, computing the ESP charges at the B3LYP/cc-pVTZ or HF/6-31G* level using RESP with a RESP weighting of 0.01 and constraining all aliphatic H's to zero charge. This method is employed to maintain consistency with AMBER and GLYCAM force fields. The charges from an ensemble are averaged if they are equal. Finally the hydroxyl H from the monosaccharide is removed and the partial charge on the linking oxygen atom is adjusted so that the net charge in the residue is the same as in the original monosaccharide residue. The above charge derivation procedure compatible with GLYCAM can be simplified using R.E.D. server [54].

The development of dedicated bioinformatics resources to handle data from a wide variety of glycobiology studies (in particular lectins) has been attempted by a number of research groups and international consortia [55, 56]; however, bioinformatics and molecular modelling methods are still poorly developed for sGAGs. There are also errors in the GAG residue and linkage nomenclature employed in the Protein Data Bank, arising from a lack of adoption of carbohydrate-specific tools for structure curation or refinement validation [57].

The goal of the research described in this thesis has not been to develop better algorithms or methods but but to address some of the bottlenecks in the molecular modelling of GAG-protein interactions that arise with existing approaches. For example, the use of two separate docking simulations with boat and chair conformations for IdoA2S is no longer needed. Fred and Omega can be used to fit optimal ring conformations during the docking simulations. Surveys of sulphate binding structural motifs can help to predict heparin binding sites on the surface of proteins. When working with a homology model (e.g. heparanase), FINDSITE is helpful to locate anchor ligands (ligands found in similar binding sites) that can serve as starting points for the docking of oligosaccharides. A similar approach was recently used in the modelling of heparanase, wherein the presence of natural substrates in the template structure (endoxylanase from *P.simplicissium* co-crystallised with a xylan oligosaccharide) allowed the prediction of a binding model for a hydrolysed HS pentasaccharide [58].

The work reported in Chapter 6 and the one by Imberty *et al.*, 2011 [58] emphasises the importance of template selection for the homology modelling of human heparanase. Both works focus on use of β(1→4) retaining, endo-acting templates and also agree that arabinofuranosidase is not an optimal template for modelling the binding site of heparanase because the α-L-arabinofuranosidase hydrolyses the α(1→3) linkage between a xylopyranose and an arabinofuranose substituent which is different from substrates of heparanase. In the work reported in this thesis, residues Lys157 to Ile543 were modelled, whereas the recent

work by Imberty *et al.* [58] reports the modelling of a monomeric construct, wherein the 8 kDa subunit (residues Gln36-Glu109) is connected to the 50 kDa subunit by the peptide GSGSGSQ. Despite the differences in the molecular models, the results from both studies complement and reinforce each other. Both studies indicate that His296 and Tyr298 are highly conserved and that His296 plays a role in the enzymatic reaction rather than in substrate binding. Both studies shed light on the direct interactions of sugar units with the HBD-2 (i.e. Gln270-Lys280). Apart from the basic residues in HBD-1 and HBD-2 that interact with anionic saccharides, these studies also highlight the role of hydrophobic residues in the catalytic pocket. Aromatic hydrophobic residues Phe385, Phe386, Phe101, Phe229, Tyr298 and Tyr299, and aliphatic residues Leu230, Ala195, Val268, Ala276 and Leu300 were predicted to form the substrate binding site in the study by Imberty *et al.* [58], where residues Phe229, Phe385 and Phe386 are further away to interact with the substrate, as predicted in our study. Furthermore, interactions with the 8 kDa domain were not modeled in this work, while interactions with the disaccharide substrate were predicted to form interactions with Leu102, Asn64 and Phe101 in the study by Imberty *et al.* [58]. Residues Tyr298, Tyr299, Leu230 and Phe229 were predicted to be part of the binding site in both studies.

There is much further work that can be done to build on the approaches and predictions reported in this thesis. Heparin, HS and GAG mimetics such as cyclitols and PI-88 are known to interact with a wide range of functionally and structurally distinct proteins, such as CXCL-8 (Chapter 4) and heparanase (Chapter 5), exhibiting greatly differing affinities for their receptors. Standard molecular docking techniques typically focus on the association of GAGs with a single target receptor, but in order to identify differences in binding specificity between different HS-binding proteins it is necessary to make comparisons across multiple receptors. Virtual screening techniques now generally applied on protein and nucleic acid systems have not kept pace with the field of carbohydrate modelling. A virtual screening

protocol for carbohydrate ligands should take into account comparative docking to multiple receptors, the conformations of oligosaccharides and the oligomeric state of the protein(s) of interest. For example, docking of cyclitols and PI-88 to known crystal structures like FGF-1, FGF-2, VEGF, CXCL-8, heparanase and CXCL-12 could be used as a benchmark. In the case of heparanase, methods such as covalent docking [59] and "substrate-imprinted docking" [60] would aid in the modelling of enzymatic catalysis, i.e. modelling the role of the enzyme in stabilising the transition state.

Understanding the structure and dynamics of sGAG-protein complexes requires the development of new computational methods for (1) conformational searching, (2) parameterisation of scoring functions and solvation terms and (3) implementation of docking algorithms on massively parallel computing infrastructure. The later issue should be now easily achievable due to the development of MPI (message passing interface)-enabled AutoDock 4.2 for high-performance computing (HPC) architectures [61]. In addition to scoring and clustering analyses, it is important to obtain validation that the predictions of docking simulations are consistent with various experimental data, such as contact sites identified by saturation transfer difference (STD), nuclear Overhauser enhancement (NOE), J-coupling NMR experiments, protein mutagenesis studies or epitope mapping studies for the same or closely related systems [62].

Apart from the various roles of sGAGs discussed in this thesis, heparin/HS also play an important role in protein aggregation. A number of neurodegenerative diseases such as Alzheimer's disease (AD), spongiform encephalopathies and tauopathies, have been directly associated with altered conformations of proteins such as Tau, Prion, α-synuclein, and amyloid-β (Aβ), that assemble to form highly organised aggregates [63]. The minimal core sequences responsible for the polymerisation of these disordered proteins contain cationic motifs, which are implicated in the interaction with sGAGs [64]. Atomistic simulations in

the presence of explicit water in the microsecond timescale are widely used to test the stability of possible β-strand arrangements in amyloid and prion proteins, and to explore the conformational changes of native or intermediate states [65]. Coarse-grained models where protein chains are represented as joint beads to reduce torsional flexibility have been used to simulate protein aggregation [65, 66]. These approaches can be expanded to characterise the formation of fibrillar and amyloid aggregates in the presence of sGAGs to understand the mechanism of heparin-induced aggregation. Considerable work has been done in the parameterisation of force fields for GAGs for use in atomistic simulations, with a few reports already describing coarse-grained potentials for aggrecan molecules [67] and GAGs such as chondroitin, CS and HA [68]. Future work should be directed towards development and validation of coarse-grained parameters for heparin/HS to enable the prediction of equilibrium solution properties of high molecular weight GAGs using molecular dynamics simulations.

8.2 References

1. Imberty, A., H. Lortat-Jacob, and S. Perez, *Structural view of glycosaminoglycan-protein interactions.* Carbohydr. Res., 2007. **342**(3-4): p. 430-9.
2. Cardin, A.D. and H.J. Weintraub, *Molecular modeling of protein-glycosaminoglycan interactions.* Arteriosclerosis, 1989. **9**(1): p. 21-32.
3. Hileman, R.E., et al., *Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins.* Bioessays, 1998. **20**(2): p. 156-67.
4. Gandhi, N.S., D.R. Coombe, and R.L. Mancera, *Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 1. Molecular modeling studies.* Biochemistry, 2008. **47**(17): p. 4851-62.
5. Coombe, D.R., *Biological implications of glycosaminoglycan interactions with haemopoietic cytokines.* Immunol. Cell Biol., 2008. **86**(7): p. 598-607.
6. Coombe, D.R., et al., *Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 2. Biochemical analyses.* Biochemistry, 2008. **47**(17): p. 4863-75.
7. Forster, M. and B. Mulloy, *Computational approaches to the identification of heparin-binding sites on the surfaces of proteins.* Biochem. Soc. Trans., 2006. **34**(Pt 3): p. 431-4.
8. Bitomsky, W. and R.C. Wade, *Docking of glycosaminoglycans to heparin-binding proteins: Validation for aFGF, bFGF, and antithrombin and application to IL-8.* J. Am. Chem. Soc., 1999. **121**(13): p. 3004-3013.

9. Waszkowycz, B., D.E. Clark, and E. Gancia, *Outstanding challenges in protein–ligand docking and structure-based virtual screening.* Wiley Interdisciplinary Reviews: Computational Molecular Science, 2011. **1**(2): p. 229-259.
10. Forli, S. and M. Botta, *Lennard-Jones potential and dummy atom settings to overcome the AUTODOCK limitation in treating flexible ring systems.* J. Chem. Inf. Model., 2007. **47**(4): p. 1481-92.
11. Rosen, J., L. Miguet, and S. Perez, *Shape: automatic conformation prediction of carbohydrates using a genetic algorithm.* J Cheminform, 2009. **1**(1): p. 16.
12. Nahmany, A., et al., *The use of a genetic algorithm search for molecular mechanics (MM3)-based conformational analysis of oligosaccharides.* Carbohydr. Res., 2005. **340**(5): p. 1059-64.
13. Strino, F., et al., *Selenoglycosides in silico: ab initio-derived reparameterization of MM4, conformational analysis using histo-blood group ABH antigens and lectin docking as indication for potential of bioactivity.* J. Comput. Aided Mol. Des., 2010. **24**(12): p. 1009-21.
14. Ferro, D.R., et al., *Treatment of ionic species in force-field calculations: sulfate and carboxylate groups in carbohydrates.* Int. J. Biol. Macromol., 1995. **17**(3-4): p. 131-6.
15. Khan, S., et al., *The solution structure of heparan sulphate differs from that of heparin: implications for function.* J Biol Chem, 2011.
16. Kerzmann, A., D. Neumann, and O. Kohlbacher, *SLICK--scoring and energy functions for protein-carbohydrate interactions.* J. Chem. Inf. Model., 2006. **46**(4): p. 1635-42.
17. Hill, A.D. and P.J. Reilly, *A Gibbs free energy correlation for automated docking of carbohydrates.* J Comput Chem, 2008. **29**(7): p. 1131-41.
18. Tang, Y.T. and G.R. Marshall, *PHOENIX: a scoring function for affinity prediction derived using high-resolution crystal structures and calorimetry measurements.* J Chem Inf Model, 2011. **51**(2): p. 214-28.
19. Sapay, N., et al., *Molecular modeling of the interaction between heparan sulfate and cellular growth factors: bringing pieces together.* Glycobiology, 2011.
20. Laederach, A., et al., *Automated docking of maltose, 2-deoxymaltose, and maltotetraose into the soybean β-amylase active site.* Proteins, 1999. **37**(2): p. 166-175.
21. Coutinho, P.M., M.K. Dowd, and P.J. Reilly, *Automated docking of glucosyl disaccharides in the glucoamylase active site.* Proteins, 1997. **28**(2): p. 162-173.
22. Samsonov, S.A., J. Teyra, and M.T. Pisabarro, *Docking glycosaminoglycans to proteins: analysis of solvent inclusion.* J. Comput. Aided Mol. Des., 2011. **25**(5): p. 477-89.
23. Minke, W.E., et al., *The role of waters in docking strategies with incremental flexibility for carbohydrate derivatives: heat-labile enterotoxin, a multivalent test case.* J. Med. Chem., 1999. **42**(10): p. 1778-88.
24. Liu, Z. and Y. Zhang, *Molecular dynamics simulations and MM-PBSA calculations of the lectin from snowdrop (Galanthus nivalis).* J Mol Model, 2009. **15**(12): p. 1501-7.
25. Mishra, N.K., et al., *Recognition of selected monosaccharides by Pseudomonas aeruginosa Lectin II analyzed by molecular dynamics and free energy calculations.* Carbohydr. Res., 2010. **345**(10): p. 1432-41.
26. Ford, M.G., et al., *Molecular dynamics simulations of galectin-1-oligosaccharide complexes reveal the molecular basis for ligand diversity.* Proteins, 2003. **53**(2): p. 229-40.

27. Agostino, M., E. Yuriev, and P.A. Ramsland, *A computational approach for exploring carbohydrate recognition by lectins in innate immunity.* Front Immunol, 2011. **2**.
28. Yuriev, E., M. Agostino, and P.A. Ramsland, *Challenges and advances in computational docking: 2009 in review.* J. Mol. Recognit., 2011. **24**(2): p. 149-64.
29. Singh, N. and A. Warshel, *Absolute binding free energy calculations: on the accuracy of computational scoring of protein-ligand interactions.* Proteins, 2010. **78**(7): p. 1705-23.
30. ÅQvist, J., *Calculation of absolute binding free energies for charged ligands and effects of long-range electrostatic interactions.* J Comput Chem, 1996. **17**(14): p. 1587-1597.
31. Green, D.F., *Optimized parameters for continuum solvation calculations with carbohydrates.* J Phys Chem B, 2008. **112**(16): p. 5238-49.
32. Hou, T., et al., *Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations.* J. Chem. Inf. Model., 2011. **51**(1): p. 69-82.
33. Rastelli, G., et al., *Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA.* J Comput Chem, 2010. **31**(4): p. 797-810.
34. Kongsted, J. and U. Ryde, *An improved method to predict the entropy term with the MM/PBSA approach.* J. Comput. Aided Mol. Des., 2009. **23**(2): p. 63-71.
35. Yang, C.Y., et al., *Importance of ligand reorganization free energy in protein-ligand binding-affinity prediction.* J. Am. Chem. Soc., 2009. **131**(38): p. 13709-21.
36. Spiwok, V. and A. D. French, *Modelling the effect of solvents on carbohydrates.* Mini Rev Org Chem, 2011. **8**(3): p. 249-255.
37. Kirschner, K.N. and R.J. Woods, *Solvent interactions determine carbohydrate conformation.* Proc. Natl. Acad. Sci. U.S.A., 2001. **98**(19): p. 10541-5.
38. Almond, A., *Towards understanding the interaction between oligosaccharides and water molecules.* Carbohydr. Res., 2005. **340**(5): p. 907-20.
39. Fadda, E. and R.J. Woods, *Molecular simulations of carbohydrates and protein-carbohydrate interactions: motivation, issues and prospects.* Drug Discov. Today, 2010. **15**(15-16): p. 596-609.
40. Sattelle, B.M., et al., *Free energy landscapes of iduronic acid and related monosaccharides.* J. Am. Chem. Soc., 2010. **132**(38): p. 13132-4.
41. Spiwok, V., B. Kralova, and I. Tvaroska, *Modelling of beta-D-glucopyranose ring distortion in different force fields: a metadynamics study.* Carbohydr. Res., 2010. **345**(4): p. 530-7.
42. Spiwok, V. and I. Tvaroska, *Metadynamics modelling of the solvent effect on primary hydroxyl rotamer equilibria in hexopyranosides.* Carbohydr. Res., 2009. **344**(12): p. 1575-81.
43. Sattelle, B.M. and A. Almond, *Less is more when simulating unsulfated glycosaminoglycan 3D-structure: comparison of GLYCAM06/TIP3P, PM3-CARB1/TIP3P, and SCC-DFTB-D/TIP3P predictions with experiment.* J Comput Chem, 2010. **31**(16): p. 2932-47.
44. Almond, A. and J.K. Sheehan, *Glycosaminoglycan conformation: do aqueous molecular dynamics simulations agree with x-ray fiber diffraction?* Glycobiology, 2000. **10**(3): p. 329-338.
45. Sega, M., E. Autieri, and F. Pederiva, *On the calculation of puckering free energy surfaces.* J Chem Phys, 2009. **130**(22): p. 225102.
46. Babin, V. and C. Sagui, *Conformational free energies of methyl-alpha-L-iduronic and methyl-beta-D-glucuronic acids in water.* J Chem Phys, 2010. **132**(10): p. 104108.

47. M. Kuttel, M., *The conformational free energy of carbohydrates.* Mini Rev Org Chem, 2011. **8**(3): p. 256-262.
48. Pol-Fachin, L. and H. Verli, *Assessment of glycoproteins dynamics from computer simulations.* Mini Rev Org Chem, 2011. **8**(3): p. 229-238.
49. Case, D.A., et al., *AmberTools 1.5 User's Manual*. 2011: The Amber Development Team (Creative Commons Attribution 2.5). 403.
50. Huige, C.J.M. and C. Altona, *Force field parameters for sulfates and sulfamates based on ab initio calculations: Extensions of AMBER and CHARMm fields.* J Comput Chem, 1995. **16**(1): p. 56-79.
51. Woods, R.J., et al. *Sulfating GLYCAM residues*. 2011 [cited 2011 27th July]; Available from: http://glycam.ccrc.uga.edu/ccrc/pages/sulfation.html.
52. Woods, R.J. and R. Chappelle, *Restrained electrostatic potential atomic partial charges for condensed-phase simulations of carbohydrates.* J Mol Struct: Theochem, 2000. **527**(1-3): p. 149-156.
53. Bayly, C.I., et al., *A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model.* J Phys Chem, 1993. **97**: p. 10269-10280.
54. Vanquelef, E., et al., *R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments.* Nucleic Acids Res., 2011. **39**(Web Server issue): p. W511-7.
55. Frank, M. and S. Schloissnig, *Bioinformatics and molecular modeling in glycobiology.* Cell Mol. Life Sci., 2010. **67**(16): p. 2749-72.
56. Lieth, C.-W.v.d., M. Frank, and T. Luetteke, *Bioinformatics for glycobiology and glycomics : An introduction*. 2009: John Wiley.
57. DeMarco, M.L. and R.J. Woods, *Structural glycobiology: A game of snakes and ladders.* Glycobiology, 2008. **18**(6): p. 426-440.
58. Sapay, N., et al., *Molecular model of human heparanase with proposed binding mode of a heparan sulfate oligosaccharide and catalytic amino acids.* Biopolymers, 2011: p. n/a-n/a.
59. Moura-Tamames, S.A., M.J. Ramos, and P.A. Fernandes, *Modelling beta-1,3-exoglucanase-saccharide interactions: structure of the enzyme-substrate complex and enzyme binding to the cell wall.* J. Mol. Graph. Model., 2009. **27**(8): p. 908-20.
60. Juhl, P.B., et al., *Modelling substrate specificity and enantioselectivity for lipases and esterases by substrate-imprinted docking.* BMC Struct. Biol., 2009. **9**: p. 39.
61. Norgan, A.P., et al., *Multilevel parallelization of AutoDock 4.2.* J Cheminform, 2011. **3**(1): p. 12.
62. Woods, R.J. and M.B. Tessier, *Computational glycoscience: characterizing the spatial and temporal properties of glycans and glycan-protein complexes.* Curr. Opin. Struct. Biol., 2010. **20**(5): p. 575-583.
63. Papy-Garcia, D., et al., *Glycosaminoglycans, protein aggregation and neurodegeneration.* Curr. Protein Pept. Sci., 2011. **12**(3): p. 258-68.
64. Diaz-Nido, J., F. Wandosell, and J. Avila, *Glycosaminoglycans and beta-amyloid, prion and tau peptides in neurodegenerative diseases.* Peptides, 2002. **23**(7): p. 1323-32.
65. Ma, B. and R. Nussinov, *Simulations as analytical tools to understand protein aggregation and predict amyloid conformation.* Curr. Opin. Chem. Biol., 2006. **10**(5): p. 445-52.
66. Urbanc, B., et al., *Molecular dynamics simulation of amyloid beta dimer formation.* Biophys. J., 2004. **87**(4): p. 2310-21.
67. Nap, R.J. and I. Szleifer, *Structure and interactions of aggrecans: statistical thermodynamic approach.* Biophys. J., 2008. **95**(10): p. 4570-83.

68. Bathe, M., et al., *A coarse-grained molecular model for glycosaminoglycans: application to chondroitin, chondroitin sulfate, and hyaluronic acid.* Biophys. J., 2005. **88**(6): p. 3870-3887.

BIBLIOGRAPHY

(2003). Chemokine/chemokine receptor nomenclature. *Cytokine*, IUIS/WHO Subcommittee on Chemokine Nomenclature. **21:** 48-49.

(2007). Chemdraw Ultra, Cambridgesoft.

(2008). FRED (Fast Rigid Exhaustive Docking). Santa Fe, NM 87508, Openeye Scientific Software.

(2008). OMEGA C++ Toolkit. Sante Fe, NM 87508, OpenEye Scientific Software, Inc.

Aamiri, A., A. Mobarek, et al. (1995). "Effect of a substituted dextran on reinnervation during regeneration of adult rat skeletal muscle." *C. R. Acad. Sci. III, Sci. Vie* **318**(10): 1037-1044.

Adjou, K. T., S. Simoneau, et al. (2003). "A novel generation of heparan sulfate mimetics for the treatment of prion diseases." *J. Gen. Virol.* **84**(9): 2595-2603.

Agostino, M., E. Yuriev, et al. (2011). "A computational approach for exploring carbohydrate recognition by lectins in innate immunity." *Front Immunol* **2**.

Ahuja, S. K. and P. M. Murphy (1996). "The CXC chemokines growth-regulated oncogene (GRO) alpha, GRObeta, GROgamma, neutrophil-activating peptide-2, and epithelial cell-derived neutrophil-activating peptide-78 are potent agonists for the type B, but not the type A, human interleukin-8 receptor." *J. Biol. Chem.* **271**(34): 20545-20550.

Aisen, P. S., S. Gauthier, et al. (2007). "Alzhemed: A potential treatment for Alzheimers disease." *Curr Alzheimer Res* **4**(4): 473-478.

Aisen, P. S., D. Saumier, et al. (2006). "A Phase II study targeting amyloid-β with 3APS in mild-to-moderate Alzheimer disease." *Neurology* **67**(10): 1757-1763.

Alban, S. (2008). "Natural and synthetic glycosaminoglycans. Molecular characteristics as the basis of distinct drug profiles." *Hamostaseologie* **28**(1-2): 51-61.

Alban, S. (2008). "Pharmacological strategies for inhibition of thrombin activity." *Curr. Pharm. Des.* **14**(12): 1152-1175.

Albo, D., C. Long, et al. (1996). "Modulation of cranial bone healing with a heparin-like dextran derivative." *J Craniofac Surg* **7**(1): 19-22.

Allinger, N. L. (1977). "Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms." *J. Am. Chem. Soc.* **99**(25): 8127-8134.

Allinger, N. L., Y. H. Yuh, et al. (1989). "Molecular mechanics. The MM3 force field for hydrocarbons. 1." *J. Am. Chem. Soc.* **111**(23): 8551-8566.

Almond, A. (2005). "Towards understanding the interaction between oligosaccharides and water molecules." *Carbohydr. Res.* **340**(5): 907-920.

Almond, A. and J. K. Sheehan (2000). "Glycosaminoglycan conformation: do aqueous molecular dynamics simulations agree with x-ray fiber diffraction?" *Glycobiology* **10**(3): 329-338.

Almond, A. and J. K. Sheehan (2003). "Predicting the molecular shape of polysaccharides from dynamic interactions with water." *Glycobiology* **13**(4): 255-264.

Altschul, S. F., T. L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." *Nucleic Acids Res.* **25**(17): 3389-3402.

Ancsin, J. B. and R. Kisilevsky (1999). "The heparin/heparan sulfate-binding site on apo-serum Amyloid A." *J. Biol. Chem.* **274**(11): 7172-7181.

Angulo, J., P. M. Nieto, et al. (2003). "A molecular dynamics description of the conformational flexibility of the L-iduronate ring in glycosaminoglycans." *Chem. Commun. (Camb.)*(13): 1512-1513.

Åqvist, J. (1996). "Calculation of absolute binding free energies for charged ligands and effects of long-range electrostatic interactions." *J Comput Chem* **17**(14): 1587-1597.

Aqvist, J. and J. Marelius (2001). "The linear interaction energy method for predicting ligand binding free energies." *Comb. Chem. High Throughput Screen.* **4**(8): 613-626.

Arocas, V., S. C. Bock, et al. (1999). "The role of Arg46 and Arg47 of antithrombin in heparin binding." *Biochemistry* **38**(31): 10196-10204.

Arocas, V., S. C. Bock, et al. (2001). "Lysine 114 of antithrombin is of crucial importance for the affinity and kinetics of heparin pentasaccharide binding." *J. Biol. Chem.* **276**(47): 43809-43817.

Arteel, G. E., S. Franken, et al. (2000). "Binding of selenoprotein P to heparin: characterization with surface plasmon resonance." *Biol. Chem.* **381**(3): 265-268.

Aruffo, A., I. Stamenkovic, et al. (1990). "CD44 is the principal cell surface receptor for hyaluronate." *Cell* **61**(7): 1303-1313.
Ashkenazy, H., E. Erez, et al. (2010). "ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids." *Nucleic Acids Res.* **38 Suppl**: W529-W533.
Atha, D. H., J. C. Lormeau, et al. (1985). "Contribution of monosaccharide residues in heparin binding to antithrombin III." *Biochemistry* **24**(23): 6723-6729.
Attwood, M. R., N. Borkakoti, et al. (1996). "Identification and characterisation of an inhibitor of interleukin-8: a receptor based approach." *Bioorg. Med. Chem. Lett.* **6**(15): 1869-1874.
Attwood, M. R., E. A. Conway, et al. (1997). "Peptide based inhibitors of interleukin-8: structural simplification and enhanced potency." *Bioorg. Med. Chem. Lett.* **7**(4): 429-432.
Austin Yongye, Matthew Tessier, et al. (2008). Glycam_0f.dat, www.glycam.com.
Avci, F. Y., N. A. Karst, et al. (2003). "Synthetic oligosaccharides as heparin-mimetics displaying anticoagulant properties." *Curr Pharm Des* **9**(28): 2323-2335.
Babin, V. and C. Sagui (2010). "Conformational free energies of methyl-alpha-L-iduronic and methyl-beta-D-glucuronic acids in water." *J Chem Phys* **132**(10): 104108.
Bae, J., U. R. Desai, et al. (1994). "Interaction of heparin with synthetic antithrombin III peptide analogues." *Biochem. J.* **301**(Pt 1): 121-129.
Baggiolini, M. and I. Clark-Lewis (1992). "Interleukin-8, a chemotactic and inflammatory cytokine." *FEBS Lett.* **307**(1): 97-101.
Baggiolini, M., B. Dewald, et al. (1994). "Interleukin-8 and related chemotactic cytokines--CXC and CC chemokines." *Adv. Immunol.* **55**: 97-179.
Baggiolini, M., A. Walz, et al. (1989). "Neutrophil-activating peptide-1/interleukin 8, a novel cytokine that activates neutrophils." *J. Clin. Invest.* **84**(4): 1045-1049.
Bal Dit Sollier, C., E. Neuhart, et al. (2009). *Anticoagulant activities of EP42675 - synthetic direct inhibitor and indirect factor Xa inhibitor*. XXII congress of the International Society of Thrombosis and Haemostasis, Boston.
Bal Dit Sollier, C., E. Neuhart, et al. (2009). *Pharmacokinetics and pharmacodynamics of EP42675 a new synthetic anticoagulant with a dual mechanism of action*. XII congress of the international society of thrombosis and haemostasis Boston.
Baldwin, E. T., K. A. Franklin, et al. (1990). "Crystallization of human interleukin-8. A protein chemotactic for neutrophils and T-lymphocytes." *J. Biol. Chem.* **265**(12): 6851-6853.
Baldwin, E. T., I. T. Weber, et al. (1991). "Crystal structure of interleukin 8: symbiosis of NMR and crystallography." *Proc. Natl. Acad. Sci. U.S.A.* **88**(2): 502-506.
Bame, K. J. (2001). "Heparanases: endoglycosidases that degrade heparan sulfate proteoglycans." *Glycobiology* **11**(6): 91R-98R.
Bandtlow, C. E. and D. R. Zimmermann (2000). "Proteoglycans in the developing brain: new conceptual insights for old proteins." *Physiol. Rev.* **80**(4): 1267-1290.
Banin, E., Y. Neuberger, et al. (2002). "A novel linear code nomenclature for complex carbohydrates." *Trends Glycosci. Glycotechnol.* **14**(77): 127-138.
Barash, U., V. Cohen-Kaplan, et al. (2010). "A novel human heparanase splice variant, T5, endowed with protumorigenic characteristics." *FASEB J.* **24**(4): 1239-1248.
Barash, U., V. Cohen-Kaplan, et al. (2010). "Proteoglycans in health and disease: new concepts for heparanase function in tumor progression and metastasis." *FEBS J.* **277**(19): 3890-3903.
Barbier-Chassefière, V., S. Garcia-Filipe, et al. (2009). "Matrix therapy in regenerative medicine, a new approach to chronic wound healing." *J Biomed Mater Res A* **90A**(3): 641-647.
Barbosa, I., C. Morin, et al. (2005). "A synthetic glycosaminoglycan mimetic (RGTA) modifies natural glycosaminoglycan species during myogenesis." *J. Cell. Sci.* **118**(1): 253-264.
Barinka, C., A. Prahl, et al. (2008). "Structure of human monocyte chemoattractant protein 4 (MCP-4/CCL13)." *Acta Crystallogr. Sect. D* **64**(3): 273-278.
Bar-Ner, M., A. Eldor, et al. (1987). "Inhibition of heparanase-mediated degradation of extracellular matrix heparan sulfate by non-anticoagulant heparin species." *Blood* **70**(2): 551-557.
Barritault, D. and J. P. Caruelle (2006). "Regenerating agents (RGTAs): a new therapeutic approach." *Ann Pharm Fr* **64**(2): 135-144.
Basche, M., D. L. Gustafson, et al. (2006). "A phase I biological and pharmacologic study of the heparanase inhibitor PI-88 in patients with advanced solid tumors." *Clin. Cancer Res.* **12**(18): 5471-5480.

Basma, M., S. Sundara, et al. (2001). "Solvated ensemble averaging in the calculation of partial atomic charges." *J Comput Chem* **22**(11): 1125-1137.

Basten, J., G. Jauran, et al. (1992). "Biologically active heparin-like fragments with a "non-glycosamino"glycan structure. Part 2 : a tetra-o-methylated pentasaccharide with high affinity for antithrombin III." *Bioorg. Med. Chem. Lett.* **2**(9): 901-904.

Bathe, M., G. C. Rutledge, et al. (2005). "A Coarse-Grained Molecular Model for Glycosaminoglycans: Application to Chondroitin, Chondroitin Sulfate, and Hyaluronic Acid." *Biophys. J.* **88**(6): 3870-3887.

Bauer, K. A. (2002). "Selective inhibition of coagulation factors: Advances in antithrombotic therapy." *Semin. Thromb. Hemost.* **28**(s2): 015-024.

Bayly, C. I., P. Cieplak, et al. (1993). "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model." *J Phys Chem* **97**(40): 10269-10280.

Becker, C. F., J. A. Guimaraes, et al. (2005). "Molecular dynamics and atomic charge calculations in the study of heparin conformation in aqueous solution." *Carbohydr. Res.* **340**(8): 1499-1507.

Bedke, J., P. J. Nelson, et al. (2010). "A novel CXCL8 protein-based antagonist in acute experimental renal allograft damage." *Mol. Immunol.* **47**(5): 1047-1057.

Bellini, R. (2010). "BELLUS Health ends NC-503 diabetes development program following results."

Benedix, A., C. M. Becker, et al. (2009). "Predicting free energy changes using structural ensembles." *Nat. Methods* **6**(1): 3-4.

Bérces, A., D. M. Whitfield, et al. (2001). "Quantitative description of six-membered ring conformations following the IUPAC conformational nomenclature." *Tetrahedron* **57**(3): 477-491.

Berendsen, H. J. C., J. R. Grigera, et al. (1987). "The missing term in effective pair potentials." *J Phys Chem* **91**(24): 6269-6271.

Berendsen, H. J. C., J. P. M. Postma, et al. (1984). "Molecular dynamics with coupling to an external bath." *J Chem Phys* **81**(8): 3684-3690.

Berezin, C., F. Glaser, et al. (2004). "ConSeq: the identification of functionally and structurally important residues in protein sequences." *Bioinformatics* **20**(8): 1322-1324.

Bernfield, M., M. Gotte, et al. (1999). "Functions of cell surface heparan sulfate proteoglycans." *Annu. Rev. Biochem.* **68**: 729-777.

Best, R. B., N.-V. Buchete, et al. (2008). "Are current molecular dynamics force fields too helical?" *Biophys. J.* **95**(1): L07-L09.

Biarnes, X., A. Ardevol, et al. (2007). "The conformational free energy landscape of beta-D-glucopyranose. Implications for substrate preactivation in beta-glucoside hydrolases." *J. Am. Chem. Soc.* **129**(35): 10686-10693.

Bijsterveld, N. R., R. Vink, et al. (2004). "Recombinant factor VIIa reverses the anticoagulant effect of the long-acting pentasaccharide idraparinux in healthy volunteers." *Br. J. Haematol.* **124**(5): 653-658.

Bishop, J. R., M. Schuksz, et al. (2007). "Heparan sulphate proteoglycans fine-tune mammalian physiology." *Nature* **446**(7139): 1030-1037.

Bisio, A., A. Mantegazza, et al. (2007). "High-performance liquid chromatographic/mass spectrometric studies on the susceptibility of heparin species to cleavage by heparanase." *Semin. Thromb. Hemost.* **33**(5): 488-495.

Bitomsky, W. and R. C. Wade (1999). "Docking of glycosaminoglycans to heparin-binding proteins: Validation for aFGF, bFGF, and antithrombin and application to IL-8." *J. Am. Chem. Soc.* **121**(13): 3004-3013.

Bizzarri, C., M. Allegretti, et al. (2003). "Pharmacological inhibition of interleukin-8 (CXCL8) as a new approach for the prevention and treatment of several human diseases." *Curr Med Chem Anti Inflamm Anti Allergy Agents* **2**(1): 67-79.

Bland, C. E., H. Ginsburg, et al. (1982). "Mouse heparin proteoglycan. Synthesis by mast cell-fibroblast monolayers during lymphocyte-dependent mast cell proliferation." *J. Biol. Chem.* **257**(15): 8661-8666.

Blanquaert, F., J. L. Saffar, et al. (1995). "Heparan-like molecules induce the repair of skull defects." *Bone* **17**(6): 499-506.

Boeckmann, B., A. Bairoch, et al. (2003). "The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003." *Nucleic Acids Res.* **31**(1): 365-370.

Bohne-Lang, A., E. Lang, et al. (2001). "LINUCS: linear notation for unique description of carbohydrate sequences." *Carbohydr. Res.* **336**(1): 1-11.
Borsig, L. (2010). "Antimetastatic activities of heparins and modified heparins. Experimental evidence." *Thromb. Res.* **125**(Supplement 2): S66-S71.
Borza, D. B. and W. T. Morgan (1998). "Histidine-proline-rich glycoprotein as a plasma pH sensor. Modulation of its interaction with glycosaminoglycans by pH and metals." *J. Biol. Chem.* **273**(10): 5493-5499.
Boulanger, M. J. and K. C. Garcia (2004). "Shared cytokine signaling receptors: structural insights from the gp130 system." *Adv. Protein Chem.* **68**: 107-146.
Bowie, J. U., R. Luthy, et al. (1991). "A method to identify protein sequences that fold into a known three-dimensional structure." *Science* **253**(5016): 164-170.
Brändén, C.-I. and J. Tooze (1999). *Introduction to protein structure*. New York, Garland Pub.
Brooks, B. R., R. E. Bruccoleri, et al. (1983). "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations." *J Comput Chem* **4**(2): 187-217.
Brown, J., J. R. Brown, et al. (2010). "22. Small molecule inhibitors of glycosaminoglycan biosynthesisas substrate optimization therapy for the mucopolysaccharidoses." *Mol Genet Metab* **99**(2): S12-S12.
Brown, J., J. R. Brown, et al. (2010). *Small molecule inhibitors of glycosaminoglycan biosynthesis as substrate optimization therapy for the mucopolysaccharidoses*. Lysosomal Disease Network WORLD Symposium, Florida, United States
Brown, J. R., B. E. Crawford, et al. (2007). "Glycan antagonists and inhibitors: A fount for drug discovery." *Crit. Rev. Biochem. Mol. Biol.* **42**(6): 481-515.
Brylinski, M. and J. Skolnick (2008). "A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation." *Proc. Natl. Acad. Sci. U.S.A.* **105**(1): 129-134.
Brylinski, M. and J. Skolnick (2009). "FINDSITE: a threading-based approach to ligand homology modeling." *PLoS Comput. Biol.* **5**(6): e1000405.
Brylinski, M. and J. Skolnick (2010). "Comparison of structure-based and threading-based approaches to protein functional annotation." *Proteins* **78**(1): 118-134.
Buller, H. R., A. T. Cohen, et al. (2007). "Extended prophylaxis of venous thromboembolism with idraparinux." *N. Engl. J. Med.* **357**(11): 1105-1112.
Buller, H. R., A. T. Cohen, et al. (2007). "Idraparinux versus standard therapy for venous thromboembolic disease." *N. Engl. J. Med.* **357**(11): 1094-1104.
Buller, H. R., A. T. Cohen, et al. (2004). "A novel long-acting synthetic factor Xa inhibitor (SanOrg34006) to replace warfarin for secondary prevention in deep vein thrombosis. A Phase II evaluation." *J. Thromb. Haemost.* **2**(1): 47-53.
Burrows, S. D., M. L. Doyle, et al. (1994). "Determination of the monomer-dimer equilibrium of interleukin-8 reveals it is a monomer at physiological concentrations." *Biochemistry* **33**(43): 12741-12745.
Busch-Petersen, J. (2006). "Small molecule antagonists of the CXCR2 and CXCR1 chemokine receptors as therapeutic agents for the treatment of inflammatory diseases." *Curr Top Med Chem* **6**(13): 1345-1352.
Bytheway, I. and S. Cochran (2004). "Validation of molecular docking calculations involving FGF-1 and FGF-2." *J Med Chem* **47**(7): 1683-1693.
Bytheway, I., E. Hammond, et al. (2009). The dual angiogenesis/heparanase inhibitor PG545, but not the tyrosine kinase inhibitor sorafenib, inhibits spontaneous metastasis in models of breast and lung cancer. *AACR-NCI-EORTC International Conference: Molecular Targets and Cancer Therapeutics*. Boston, MA, Mol. Cancer Ther. (Meeting Abstract Supplement).
Cabannes, E., A. Caravano, et al. (2009). *Heparan Sulfate Mimetics as anticancer "small-glyco drugs*. 67th Harden Conference, Cambridge, UK.
Caldwell, E. E., V. D. Nadkarni, et al. (1996). "Importance of specific amino acids in protein binding sites for heparin and heparan sulfate." *Int. J. Biochem. Cell Biol.* **28**(2): 203-216.
Canales, A., J. Angulo, et al. (2005). "Conformational flexibility of a synthetic glycosylaminoglycan bound to a fibroblast growth factor. FGF-1 recognizes both the (1)C(4) and (2)S(O) conformations of a bioactive heparin-like hexasaccharide." *J. Am. Chem. Soc.* **127**(16): 5778-5779.
Canales, A., R. Lozano, et al. (2006). "Solution NMR structure of a human FGF-1 monomer, activated by a hexasaccharide heparin-analogue." *FEBS J.* **273**(20): 4716-4727.

Cantu, D., W. Nerinckx, et al. (2008). "Theory and computation show that Asp463 is the catalytic proton donor in human endoplasmic reticulum alpha-(1-->2)-mannosidase I." *Carbohydr. Res.* **343**(13): 2235-2242.
Cao, G., C. D. O'Brien, et al. (2002). "Involvement of human PECAM-1 in angiogenesis and in vitro endothelial cell migration." *Am. J. Physiol., Cell Physiol.* **282**(5): C1181-C1190.
Cao, H. and B. Yu (2005). "Synthesis of a S-linked heparan sulfate trisaccharide as the substrate mimic of heparanase." *Tetrahedron Lett.* **46**(25): 4337-4340.
Capila, I., M. J. Hernaiz, et al. (2001). "Annexin V--heparin oligosaccharide complex suggests heparan sulfate--mediated assembly on cell surfaces." *Structure* **9**(1): 57-64.
Capila, I. and R. J. Linhardt (2002). "Heparin-protein interactions." *Angew. Chem. Int. Ed. Engl.* **41**(3): 391-412.
Cardin, A. D. and H. J. Weintraub (1989). "Molecular modeling of protein-glycosaminoglycan interactions." *Arteriosclerosis* **9**(1): 21-32.
Carré, P. C., R. L. Mortenson, et al. (1991). "Increased expression of the interleukin-8 gene by alveolar macrophages in idiopathic pulmonary fibrosis. A potential mechanism for the recruitment and activation of neutrophils in lung fibrosis." *J. Clin. Invest.* **88**(6): 1802-1810.
Case, D., T. Darden, et al. (2006). AMBER San Francisco, University of California.
Case, D. A. (1994). "Normal mode analysis of protein dynamics." *Curr. Opin. Struct. Biol.* **4**(2): 285-290.
Case, D. A., T. E. Cheatham Iii, et al. (2005). "The Amber biomolecular simulation programs." *J Comput Chem* **26**(16): 1668-1688.
Case, D. A., J. Wang, et al. (2011). *AmberTools 1.5 User's Manual*, The Amber Development Team (Creative Commons Attribution 2.5).
Casu, B., M. Guerrini, et al. (2004). "Structural and conformational aspects of the anticoagulant and anti-thrombotic activity of heparin and dermatan sulfate." *Curr. Pharm. Des.* **10**(9): 939-949.
Casu, B. and U. Lindahl (2001). "Structure and biological interactions of heparin and heparan sulfate." *Adv Carbohydr Chem Biochem* **57**: 159-206.
Casu, B., A. Naggi, et al. (2010). "Heparin-derived heparan sulfate mimics to modulate heparan sulfate-protein interaction in inflammation and cancer." *Matrix Biol.* **29**(6): 442-452.
Casu, B., I. Vlodavsky, et al. (2008). "Non-anticoagulant heparins and inhibition of cancer." *Pathophysiol. Haemost. Thromb.* **36**(3-4): 195-203.
Ceroni, A., A. Dell, et al. (2007). "The GlycanBuilder: a fast, intuitive and flexible software tool for building and displaying glycan structures." *Source Code Biol Med* **2**: 3.
Chakrabarti, S., L. Beaulieu, et al. (2009). "M118, a novel low-molecular weight heparin with decreased polydispersity leads to enhanced anticoagulant activity and thrombotic occlusion in ApoE knockout mice." *J. Thromb. Thrombolysis* **28**(4): 394-400.
Champe, P., R. Harvey, et al. (1994). Glycosaminoglycans and glycoproteins. *Lippincott's illustrated reviews: Biochemistry*. P. Champe and R. Harvey. Baltimore, MA, Lippincott Williams & Wilkins **4:** 528.
Cheatham, T. E., 3rd and M. A. Young (2000). "Molecular dynamics simulation of nucleic acids: successes, limitations, and promise." *Biopolymers* **56**(4): 232-256.
Chen, M. E., H. X. Cang, et al. (2007). NOC. Florida, Inst. Mol. Biol., Florida State University.
Chen, W. Y. and G. Abatangelo (1999). "Functions of hyaluronan in wound repair." *Wound Repair Regen* **7**(2): 79-89.
Chevalier, F., J. Angulo, et al. (2002). "The heparin−Ca2+ interaction: structure of the Ca2+ binding site." *Eur. J. Org. Chem.* **2002**(14): 2367-2376.
Chevalier, F., R. Lucas, et al. (2004). "The heparin-Ca(2+) interaction: the influence of the O-sulfation pattern on binding." *Carbohydr. Res.* **339**(5): 975-983.
Chiara, M., F. Ian, et al. (2009). "Design and rationale of the Evaluation of M118 IN pErcutaNeous Coronary intErvention (EMINENCE) trial." *Am. Heart J.* **158**(5): 726-733.
Choay, J., M. Petitou, et al. (1983). "Structure-activity relationship in heparin: a synthetic pentasaccharide with high affinity for antithrombin III and eliciting high anti-factor Xa activity." *Biochem. Biophys. Res. Commun.* **116**(2): 492-499.
Chollet-Martin, S., P. Montravers, et al. (1993). "High levels of interleukin-8 in the blood and alveolar spaces of patients with pneumonia and adult respiratory distress syndrome." *Infect. Immun.* **61**(11): 4553-4559.

Chow, L., D. Gustafson, et al. (2008). "A phase I pharmacological and biological study of PI-88 and docetaxel in patients with advanced malignancies." *Cancer Chemother. Pharmacol.* **63**(1): 65-74.

Chu, C., J. Duffner, et al. (2009). M-ONC 402, A novel non-anticoagulant heparin, inhibits P-Selectin function and metastatic seeding of tumor cells in mice. *100th Annual Meeting of American Association for Cancer Research (AACR)*. Dever, Colorado.

Chuntharapai, A. and K. J. Kim (1995). "Regulation of the expression of IL-8 receptor A/B by IL-8: possible functions of each receptor." *J. Immunol.* **155**(5): 2587-2594.

Clark-Lewis, I., B. Dewald, et al. (1994). "Structural requirements for interleukin-8 function identified by design of analogs and CXC chemokine hybrids." *J. Biol. Chem.* **269**(23): 16075-16081.

Clinicaltrials.gov (2002). A Phase II/III Study of the Safety and Efficacy of NC-503 in Patients Suffering From Secondary (AA) Amyloidosis (NCT00035334).

Clore, G. M., E. Appella, et al. (1989). "Determination of the secondary structure of interleukin-8 by nuclear magnetic resonance spectroscopy." *J. Biol. Chem.* **264**(32): 18907-18911.

Clore, G. M., E. Appella, et al. (1990). "Three-dimensional structure of interleukin 8 in solution." *Biochemistry* **29**(7): 1689-1696.

Clore, G. M. and A. M. Gronenborn (1991). "Comparison of the solution nuclear magnetic resonance and crystal structures of interleukin-8. Possible implications for the mechanism of receptor binding." *J. Mol. Biol.* **217**(4): 611-620.

Clubb, R. T., J. G. Omichinski, et al. (1994). "Mapping the binding surface of interleukin-8 complexed with an N-terminal fragment of the Type 1 human interleukin-8 receptor." *FEBS Lett.* **338**(1): 93-97.

Cochran, S., C. P. Li, et al. (2005). "An experimental and molecular-modeling study of the binding of linked sulfated tetracyclitols to FGF-1 and FGF-2." *Chembiochem* **6**(10): 1882-1890.

Codée, J. D. C., H. S. Overkleeft, et al. (2004). "The synthesis of well-defined heparin and heparan sulfate fragments." *Drug Discov Today Technol* **1**(3): 317-326.

Combet, C., C. Blanchet, et al. (2000). "NPS@: network protein sequence analysis." *Trends Biochem. Sci.* **25**(3): 147-150.

Conrad, H. (1998). *Heparin-binding proteins*. San Diego, Academic Press.

Coombe, D. R. (2008). "Biological implications of glycosaminoglycan interactions with haemopoietic cytokines." *Immunol. Cell Biol.* **86**(7): 598-607.

Coombe, D. R., S. M. Stevenson, et al. (2008). "Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 2. Biochemical analyses." *Biochemistry* **47**(17): 4863-4875.

Cornell, W., R. Abseher, et al. (2001). "Continuum solvent molecular dynamics study of flexibility in interleukin-8." *J. Mol. Graph. Model.* **19**(1): 136-145.

Cornell, W. D., P. Cieplak, et al. (1995). "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules." *J. Am. Chem. Soc.* **117**(19): 5179-5197.

Cornell, W. D., P. Cieplak, et al. (1993). "Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation." *J. Am. Chem. Soc.* **115**(21): 9620-9631.

Coutinho, P. M., M. K. Dowd, et al. (1997). "Automated docking of glucosyl disaccharides in the glucoamylase active site." *Proteins* **28**(2): 162-173.

Coutinho, P. M., M. K. Dowd, et al. (1997). "Automated docking of monosaccharide substrates and analogues and methyl alpha-acarviosinide in the glucoamylase active site." *Proteins* **27**(2): 235-248.

Coutinho, P. M., M. K. Dowd, et al. (1998). "Automated docking of α-(1,4)- and α-(1,6)-linked llucosyl trisaccharides in the glucoamylase active site." *Ind Eng Chem Res* **37**(6): 2148-2157.

Covello, J. M., S. Bird, et al. (2009). "Cloning and expression analysis of three striped trumpeter (Latris lineata) pro-inflammatory cytokines, TNF-α, IL-1α and IL-8, in response to infection by the ectoparasitic, Chondracanthus goldsmidi." *Fish Shellfish Immunol.* **26**(5): 773-786.

Cremer, D. and J. A. Pople (1975). "General definition of ring puckering coordinates." *J. Am. Chem. Soc.* **97**(6): 1354-1358.

Czjzek, M., A. Ben David, et al. (2005). "Enzyme-substrate complex structures of a GH39 beta-xylosidase from Geobacillus stearothermophilus." *J. Mol. Biol.* **353**(4): 838-846.

da Rocha Pita, S. S., T. V. A. Fernandes, et al. (2008). "Studies of molecular docking between fibroblast growth factor and heparin using generalized simulated annealing." *Int J Quantum Chem* **108**(13): 2608-2614.
Darden, T., D. York, et al. (1993). "Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems." *J Chem Phys* **98**(12): 10089-10092.
Das, S. T., L. Rajagopalan, et al. (2010). "Monomeric and dimeric CXCL8 are both essential for *in vivo* neutrophil recruitment." *PLoS ONE* **5**(7): e11754.
David, R., R. Günther, et al. (2008). "Artificial chemokines: Combining chemistry and molecular biology for the elucidation of interleukin-8 functionality." *J. Am. Chem. Soc.* **130**(46): 15311-15317.
Davies, G. and B. Henrissat (1995). "Structures and mechanisms of glycosyl hydrolases." *Structure* **3**(9): 853-859.
Davies, G. J., K. S. Wilson, et al. (1997). "Nomenclature for sugar-binding subsites in glycosyl hydrolases." *Biochem. J.* **321 (Pt 2)**: 557-559.
Davies, M. N., C. P. Toseland, et al. (2006). "Benchmarking pK(a) prediction." *BMC Biochem.* **7**: 18.
De Kort, M., R. C. Buijsman, et al. (2005). "Synthetic heparin derivatives as new anticoagulant drugs." *Drug Discov. Today* **10**(11): 769-779.
De Kort, M. and C. A. A. Van Boeckel (2010). Antithrombotic dual inhibitors comprising a biotin residue. United States, N.V. ORGANON (Oss, NL).
de Paz, J. L., E. A. Moseman, et al. (2007). "Profiling heparin-chemokine interactions using synthetic tools." *ACS Chem. Biol.* **2**(11): 735-744.
de Paz, J. L., C. Noti, et al. (2006). "Microarrays of synthetic heparin oligosaccharides." *J. Am. Chem. Soc.* **128**(9): 2766-2767.
de Paz, J. L. and P. H. Seeberger (2008). "Deciphering the glycosaminoglycan code with the help of microarrays." *Mol Biosyst* **4**(7): 707-711.
Deaglio, S., M. Morra, et al. (1998). "Human CD38 (ADP-ribosyl cyclase) is a counter-receptor of CD31, an Ig superfamily member." *J. Immunol.* **160**(1): 395-402.
Dealwis, C., E. J. Fernandez, et al. (1998). "Crystal structure of chemically synthesized [N33A] stromal cell-derived factor 1α, a potent ligand for the HIV-1 "fusin" coreceptor." *Proc. Natl. Acad. Sci. U.S.A.* **95**(12): 6941-6946.
Debeche, T., C. Bliard, et al. (2002). "Probing the catalytically essential residues of the alpha-L-arabinofuranosidase from Thermobacillus xylanilyticus." *Protein Eng.* **15**(1): 21-28.
DeLisser, H. M., H. C. Yan, et al. (1993). "Platelet/endothelial cell adhesion molecule-1 (CD31)-mediated cellular aggregation involves cell surface glycosaminoglycans." *J. Biol. Chem.* **268**(21): 16037-16046.
DeMarco, M. L. and R. J. Woods (2008). "Structural glycobiology: A game of snakes and ladders." *Glycobiology* **18**(6): 426-440.
Dember, L. M., P. N. Hawkins, et al. (2007). "Eprodisate for the treatment of renal disease in AA amyloidosis." *N. Engl. J. Med.* **356**(23): 2349-2360.
Dempsey, L. A., T. B. Plummer, et al. (2000). "Heparanase expression in invasive trophoblasts and acute vascular damage." *Glycobiology* **10**(5): 467-475.
Desai, U., R. Swanson, et al. (2000). "Role of arginine 129 in heparin binding and activation of antithrombin." *J. Biol. Chem.* **275**(25): 18976-18984.
Desai, U. R. (2004). "New antithrombin-based anticoagulants." *Med Res Rev* **24**(2): 151-181.
Desai, U. R., M. Petitou, et al. (1998). "Mechanism of heparin activation of antithrombin: Evidence for an induced-fit model of allosteric activation involving two interaction subsites." *Biochemistry* **37**(37): 13033-13041.
Desai, U. R., M. Petitou, et al. (1998). "Mechanism of heparin activation of antithrombin: Role of individual residues of the pentasaccharide activating sequence in the recognition of native and activated states of antithrombin " *J. Biol. Chem.* **273**(13): 7478-7487.
Desai, U. R., H. M. Wang, et al. (1993). "Structure elucidation of a novel acidic tetrasaccharide and hexasaccharide derived from a chemically modified heparin." *Carbohydr. Res.* **241**: 249-259.
Desai, U. R., H. M. Wang, et al. (1993). "Specificity studies on the heparin lyases from Flavobacterium heparinum." *Biochemistry* **32**(32): 8140-8145.
DeVries, M. E., A. A. Kelvin, et al. (2006). "Defining the origins and evolution of the Chemokine/Chemokine receptor system." *J. Immunol.* **176**(1): 401-415.

Diaz-Nido, J., F. Wandosell, et al. (2002). "Glycosaminoglycans and beta-amyloid, prion and tau peptides in neurodegenerative diseases." *Peptides* **23**(7): 1323-1332.
Dietrich, C. P., J. F. Paiva, et al. (1999). "Structural features and anticoagulant activities of a novel natural low molecular weight heparin from the shrimp Penaeus brasiliensis." *Biochim. Biophys. Acta* **1428**(2-3): 273-283.
DiGabriele, A. D., I. Lax, et al. (1998). "Structure of a heparin-linked biologically active dimer of fibroblast growth factor." *Nature* **393**(6687): 812-817.
Dominguez, R., H. Souchon, et al. (1996). "The crystal structure of a family 5 endoglucanase mutant in complexed and uncomplexed forms reveals an induced fit activation mechanism." *J. Mol. Biol.* **257**(5): 1042-1051.
Dong, J., A. K. Kukula, et al. (2000). "Genomic organization and chromosome localization of the newly identified human heparanase gene." *Gene* **253**(2): 171-178.
Draganov, D., T. Wright, et al. (2009). "Pharmacokinetics of M118, unfractionated heparin and enoxaparin sodium in normal and 5/6 nephrectomized uremic rats." *Toxicol. Lett.* **189**(Supplement 1): S113-S113.
Dredge, K., E. Hammond, et al. (2009). "The PG500 series: novel heparan sulfate mimetics as potent angiogenesis and heparanase inhibitors for cancer therapy." *Invest New Drugs* **28**(3): 276-283.
Driguez, H. (2001). "Thiooligosaccharides as tools for structural biology." *Chembiochem* **2**(5): 311-318.
Duan, H., T. Koga, et al. (2001). "Interleukin-8-positive neutrophils in psoriasis." *J. Dermatol. Sci.* **26**(2): 119-124.
Dvir, H., M. Harel, et al. (2003). "X-ray structure of human acid-beta-glucosidase, the defective enzyme in Gaucher disease." *EMBO Rep.* **4**(7): 704-709.
Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." *Nucleic Acids Res.* **32**(5): 1792-1797.
Eigenbrot, C., H. B. Lowman, et al. (1997). "Structural change and receptor binding in a chemokine mutant with a rearranged disulfide: X-ray structure of e38C/C50A IL-8 at 2 Å resolution." *Proteins* **27**(4): 556-566.
Eikelboom, J. W. and J. I. Weitz (2010). "New anticoagulants." *Circulation* **121**(13): 1523-1532.
El-Assal, O. N., A. Yamanoi, et al. (2001). "The clinicopathological significance of heparanase and basic fibroblast growth factor expressions in hepatocellular carcinoma." *Clin. Cancer Res.* **7**(5): 1299-1305.
Eldridge, M. D., C. W. Murray, et al. (1997). "Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes." *J. Comput. Aided Mol. Des.* **11**(5): 425-445.
Elkin, M., N. Ilan, et al. (2001). "Heparanase as mediator of angiogenesis: mode of action." *FASEB J.* **15**(9): 1661-1663.
Eric, F. P., D. G. Thomas, et al. (2004). "UCSF Chimera - A visualization system for exploratory research and analysis." *J Comput Chem* **25**(13): 1605-1612.
Ernst, S., G. Venkataraman, et al. (1998). "Pyranose ring flexibility. Mapping of physical data for iduronate in continuous conformational space." *J. Am. Chem. Soc.* **120**(9): 2099-2107.
Escartin, Q., C. Lallam-Laroye, et al. (2003). "A new approach to treat tissue destruction in periodontitis with chemically modified dextran polymers." *FASEB J.* **17**(6): 644-651.
Esko, J. D. and R. J. Linhardt (2009). Proteins that bind sulfated glycosaminoglycans. *Essentials of glycobiology*. A. Varki, R. D. Cummings, J. D. Eskoet al. New York, CSHL Press.
Esther, K., R. Jordi, et al. (2004). "Comparative evaluation of eight docking tools for docking and virtual screening accuracy." *Proteins* **57**(2): 225-242.
Faaij, R. A., J. Burggraaf, et al. (1999). "A phase I single rising dose study to investigate the safety, tolerance and pharmacokinetics of SANORG 34006 in healthy young male volunteers." *Thromb Haemost*(Abstract 2709): 853-853.
Faaij, R. A., J. Burggraaf, et al. (1999). "A phase I single rising dose study to investigate the safety, tolerance and pharmacokinetics of subcutaneous SANORG 34006 in healthy male and female elderly volunteers." *Thromb Haemost*(Abstract 1547): 490-491.
Fabricius, J., S. B. Engelsen, et al. (1997). "The consistent force field. 5. PEF95SAC: optimized potential energy function for alcohols and carbohydrates." *J Carbohydr Chem* **16**(6): 751 - 772.

Fadda, E. and R. J. Woods (2010). "Molecular simulations of carbohydrates and protein-carbohydrate interactions: motivation, issues and prospects." *Drug Discov. Today* **15**(15-16): 596-609.
Faham, S., R. E. Hileman, et al. (1996). "Heparin structure and interactions with basic fibroblast growth factor." *Science* **271**(5252): 1116-1120.
Fairbanks, M. B., A. M. Mildner, et al. (1999). "Processing of the human heparanase precursor and evidence that the active enzyme is a heterodimer." *J. Biol. Chem.* **274**(42): 29587-29590.
Fallahi, A., B. Kroll, et al. (2005). "Structural model of the amino propeptide of collagen XI alpha1 chain with similarity to the LNS domains." *Protein Sci.* **14**(6): 1526-1537.
Fareed, J., D. A. Hoppensteadt, et al. (2000). "An update on heparins at the beginning of the new millennium." *Semin. Thromb. Hemost.* **26** (Suppl 1): 5-21.
Fareed, J., W. Jeske, et al. (2008). "Neutralization of the anticoagulant and anti-Xa effects of fondaparinux and idraparinux by a novel synthetic antagonist. Pharmacologic implications." *FASEB J.* **22**(1_MeetingAbstracts): 1118.1117-.
Fernandez, C., C. M. Hattan, et al. (2006). "Semi-synthetic heparin derivatives: chemical modifications of heparin beyond chain length, sulfate substitution pattern and N-sulfo/N-acetyl groups." *Carbohydr. Res.* **341**(10): 1253-1265.
Fernandez, E. J. and E. Lolis (2002). "Structure, function, and inhibition of chemokines." *Annu. Rev. Pharmacol. Toxicol.* **42**(1): 469-499.
Fernando, H., C. Chin, et al. (2004). "Dimer dissociation is essential for interleukin-8 (IL-8) binding to CXCR1 receptor." *J. Biol. Chem.* **279**(35): 36175-36178.
Ferro, D. R., A. Provasoli, et al. (1990). "Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences." *Carbohydr. Res.* **195**(2): 157-167.
Ferro, D. R., A. Provasoli, et al. (1986). "Evidence for conformational equilibrium of the sulfated L-iduronate residue in heparin and in synthetic heparin mono- and oligo-saccharides: NMR and force-field studies." *J. Am. Chem. Soc.* **108**(21): 6773-6778.
Ferro, D. R., P. Pumilia, et al. (1995). "Treatment of ionic species in force-field calculations: sulfate and carboxylate groups in carbohydrates." *Int. J. Biol. Macromol.* **17**(3-4): 131-136.
Ferro, D. R., P. Pumilia, et al. (1997). "An improved force field for conformational analysis of sulfated polysaccharides." *J Comput Chem* **18**(3): 351-367.
Ferro, V. and R. Don (2003). "The development of the novel angiogenesis inhibitor PI-88 as an anticancer drug." *Australas Biotechnol* **13**(1): 38-39.
Ferro, V., K. Dredge, et al. (2007). "PI-88 and novel heparan sulfate mimetics inhibit angiogenesis." *Semin. Thromb. Hemost.* **33**(05): 557-568.
Ferro, V., E. Hammond, et al. (2004). "The development of inhibitors of heparanase, a key enzyme involved in tumour metastasis, angiogenesis and inflammation." *Mini Rev Med Chem* **4**(6): 693-702.
Fier, I., M. A. Nedelman, et al. (2007). "A novel, rationally engineered heparin (M118) prevents thrombosis more effectively than unfractionated heparin in a canine model of deep arterial injury." *Journal of the American College of Cardiology* **49**(9): 379A-380A.
Fier, I. D., C. M. Brandquist, et al. (2009). "Lack of pharmacokinetic and pharmacodynamic interactions between M118, a novel low-molecular-weight-heparin and Eptifibatide in healthy subjects." *J Clin Pharmacol* **49**(9): 73.
Fogolari, F., A. Brigo, et al. (2003). "Protocol for MM/PBSA molecular dynamics simulations of proteins." *Biophys. J.* **85**(1): 159-166.
Ford, M. G., T. Weimar, et al. (2003). "Molecular dynamics simulations of galectin-1-oligosaccharide complexes reveal the molecular basis for ligand diversity." *Proteins* **53**(2): 229-240.
Forli, S. and M. Botta (2007). "Lennard-Jones potential and dummy atom settings to overcome the AUTODOCK limitation in treating flexible ring systems." *J. Chem. Inf. Model.* **47**(4): 1481-1492.
Forster, M. and B. Mulloy (2006). "Computational approaches to the identification of heparin-binding sites on the surfaces of proteins." *Biochem. Soc. Trans.* **34**(Pt 3): 431-434.
Forster, M. J. (2005). mdxvu-Molecular Dynamics X11 Viewer.
Forster, M. J. and B. Mulloy (1993). "Molecular dynamics study of iduronate ring conformation." *Biopolymers* **33**(4): 575-588.
Fort, S., A. Varrot, et al. (2001). "Mixed-linkage cellooligosaccharides: a new class of glycoside hydrolase inhibitors." *Chembiochem* **2**(5): 319-325.

Frank, M. and S. Schloissnig (2010). "Bioinformatics and molecular modeling in glycobiology." *Cell Mol. Life Sci.* **67**(16): 2749-2772.
Fraser, P. E., J. T. Nguyen, et al. (1991). "pH-dependent structural transitions of Alzheimer amyloid peptides." *Biophys. J.* **60**(5): 1190-1201.
Freeman, C., L. Liu, et al. (2005). "Use of sulfated linked cyclitols as heparan sulfate mimetics to probe the heparin/heparan sulfate binding specificity of proteins." *J. Biol. Chem.* **280**(10): 8842-8849.
Freeman, C. and C. R. Parish (1998). "Human platelet heparanase: purification, characterization and catalytic activity." *Biochem. J.* **330** (Pt 3): 1341-1350.
Friedmann, Y., I. Vlodavsky, et al. (2000). "Expression of heparanase in normal, dysplastic, and neoplastic human colonic mucosa and stroma. Evidence for its role in colonic tumorigenesis." *Am. J. Pathol.* **157**(4): 1167-1175.
Frisch, M. J., G. W. Trucks, et al. (2004). Gaussian 03. Wallingford, CT, Gaussian, Inc.
Frishman, D. and P. Argos (1996). "Incorporation of non-local interactions in protein secondary structure prediction from the amino acid sequence." *Protein Eng.* **9**(2): 133-142.
Fromm, J. R., R. E. Hileman, et al. (1997). "Pattern and spacing of basic amino acids in heparin binding sites." *Arch. Biochem. Biophys.* **343**(1): 92-100.
Fry, E. E., S. M. Lea, et al. (1999). "The structure and function of a foot-and-mouth disease virus-oligosaccharide receptor complex." *EMBO J.* **18**(3): 543-554.
Fugedi, P. (2003). "The potential of the molecular diversity of heparin and heparan sulfate for drug development." *Mini Rev Med Chem* **3**(7): 659-667.
Fuster, M. M. and J. D. Esko (2005). "The sweet and sour of cancer: Glycans as novel therapeutic targets." *Nat. Rev. Cancer* **5**(7): 526-542.
Fux, L., N. Feibish, et al. (2009). "Structure-function approach identifies a COOH-terminal domain that mediates heparanase signaling." *Cancer Res* **69**(5): 1758-1767.
Fux, L., N. Ilan, et al. (2009). "Heparanase: busy at the cell surface." *Trends Biochem. Sci.* **34**(10): 511-519.
Gandhi, N. S., D. R. Coombe, et al. (2008). "Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 1. Molecular modeling studies." *Biochemistry* **47**(17): 4851-4862.
Gandhi, N. S. and R. L. Mancera (2008). "The structure of glycosaminoglycans and their interactions with proteins." *Chem Biol Drug Des* **72**(6): 455-482.
Gandhi, N. S. and R. L. Mancera (2009). "Free energy calculations of glycosaminoglycan-protein interactions." *Glycobiology* **19**(10): 1103-1115.
Gandhi, N. S. and R. L. Mancera (2010). "Heparin/heparan sulphate-based drugs." *Drug Discov. Today* **15**(23-24): 1058-1069.
Gandhi, N. S. and R. L. Mancera (2011). "Molecular dynamics simulations of CXCL-8 and its interactions with a receptor peptide, heparin fragments, and sulfated linked cyclitols." *J Chem Inf Model* **51**(2): 335-358.
Garcia-Filipe, S., V. Barbier-Chassefiere, et al. (2007). "RGTA OTR4120, a heparan sulfate mimetic, is a possible long-term active agent to heal burned skin." *J Biomed Mater Res A* **80A**(1): 75-84.
Garrett, M. M., S. G. David, et al. (1998). "Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function." *J Comput Chem* **19**(14): 1639-1662.
Gatti, G., B. Casu, et al. (1979). "Studies on the conformation of heparin by 1H and 13C NMR spectroscopy." *Macromolecules* **12**(5): 1001-1007.
Gayle, R. B. d., P. R. Sleath, et al. (1993). "Importance of the amino terminus of the interleukin-8 receptor in ligand interactions." *J. Biol. Chem.* **268**(10): 7283-7289.
Geerts, H. (2004). "NC-531 (Neurochem)." *Curr Opin Investig Drugs* **5**(1): 95 - 100.
Gehlhaar, D. K., G. M. Verkhivker, et al. (1995). "Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming." *Chem. Biol.* **2**(5): 317-324.
Gerber, N., H. Lowman, et al. (2000). "Receptor-binding conformation of the "ELR" motif of IL-8: X-ray structure of the L5C/H33C variant at 2.35 Å resolution." *Proteins* **38**(4): 361-367.
Gerke, V., C. E. Creutz, et al. (2005). "Annexins: linking Ca2+ signalling to membrane dynamics." *Nat. Rev. Mol. Cell Biol.* **6**(6): 449-461.

Gervais, F., C. Morissette, et al. (2003). "Proteoglycans and amyloidogenic proteins in peripheral amyloidosis." *Curr Med Chem Immunol Endocr Metab Agents* **3**(4): 361-370.
Gervais, F., J. Paquette, et al. (2007). "Targeting soluble Aβ peptide with Tramiprosate for the treatment of brain amyloidosis." *Neurobiol. Aging* **28**(4): 537-547.
Gesslbauer, B. and A. J. Kungl (2006). "Glycomic approaches toward drug development: Therapeutically exploring the glycosaminoglycanome." *Curr. Opin. Mol. Ther.* **8**(6): 521 - 528.
Gilat, D., R. Hershkoviz, et al. (1995). "Molecular behavior adapts to context: heparanase functions as an extracellular matrix-degrading enzyme or as a T cell adhesion molecule, depending on the local pH." *J. Exp. Med.* **181**(5): 1929-1934.
Gilson, M. K. and B. Honig (1988). "Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis." *Proteins* **4**(1): 7-18.
Gimbrone, M. A., Jr., M. S. Obin, et al. (1989). "Endothelial interleukin-8: a novel inhibitor of leukocyte-endothelial interactions." *Science* **246**(4937): 1601-1603.
Goger, B., Y. Halden, et al. (2002). "Different affinities of glycosaminoglycan oligosaccharides for monomeric and dimeric interleukin-8: A model for chemokine regulation at inflammatory sites." *Biochemistry* **41**(5): 1640-1646.
Gohji, K., M. Okamoto, et al. (2001). "Heparanase protein and gene expression in bladder cancer." *J. Urol.* **166**(4): 1286-1290.
Gohlke, H. and D. A. Case (2004). "Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf." *J Comput Chem* **25**(2): 238-250.
Goldenberg, O., E. Erez, et al. (2009). "The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures." *Nucleic Acids Res.* **37**(Database issue): D323-D327.
Gonzalez-Iglesias, R., M. A. Pajares, et al. (2002). "Prion protein interaction with glycosaminoglycan occurs with the formation of oligomeric complexes stabilized by Cu(II) bridges." *J. Mol. Biol.* **319**(2): 527-540.
Gonzalez-Stawinski, G. V., W. Parker, et al. (1999). "Partial sequence of human platelet heparitinase and evidence of its ability to polymerize." *Biochim. Biophys. Acta* **1429**(2): 431-438.
Goodford, P. J. (1985). "A computational procedure for determining energetically favorable binding sites on biologically important macromolecules." *J. Med. Chem.* **28**(7): 849-857.
Goodman, R. B., J. W. Forstrom, et al. (1991). "Identification of two neutrophil chemotactic peptides produced by porcine alveolar macrophages." *J. Biol. Chem.* **266**(13): 8455-8463.
Goodman, R. B., D. C. Foster, et al. (1992). "Molecular cloning of porcine alveolar macrophage-derived neutrophil chemotactic factors I and II; identification of porcine IL-8 and another intercrine-alpha protein." *Biochemistry* **31**(43): 10483-10490.
Gordon, J. C., J. B. Myers, et al. (2005). "H++: a server for estimating pKas and adding missing hydrogens to macromolecules." *Nucleic Acids Res.* **33**(Web Server issue): W368-W371.
Goshen, R., A. A. Hochberg, et al. (1996). "Purification and characterization of placental heparanase and its expression by cultured cytotrophoblasts." *Mol. Hum. Reprod.* **2**(9): 679-684.
Gozalbes, R., S. Mosulen, et al. (2009). "Development and NMR validation of minimal pharmacophore hypotheses for the generation of fragment libraries enriched in heparanase inhibitors." *J. Comput. Aided Mol. Des.* **23**(8): 555-569.
Graham, L. D. and P. A. Underwood (1996). "Comparison of the heparanase enzymes from mouse melanoma cells, mouse macrophages, and human platelets." *Biochem. Mol. Biol. Int.* **39**(3): 563-571.
Grasberger, B. L., A. M. Gronenborn, et al. (1993). "Analysis of the backbone dynamics of interleukin-8 by 15N relaxation measurements." *J. Mol. Biol.* **230**(2): 364-372.
Green, D. F. (2008). "Optimized parameters for continuum solvation calculations with carbohydrates." *J Phys Chem B* **112**(16): 5238-5249.
Gregory, H., J. Young, et al. (1988). "Structure determination of a human lymphocyte derived neutrophil activating peptide (LYNAP)." *Biochem. Biophys. Res. Commun.* **151**(2): 883-890.
Grootenhuis, P. D. J. and C. A. G. Haasnoot (1993). "A CHARMm based force field for carbohydrates using the CHEAT approach: carbohydrate hydroxyl groups represented by extended atoms." *Mol Simul* **10**(2): 75 - 95.

Grootenhuis, P. D. J. and C. A. A. Van Boeckel (1991). "Constructing a molecular model of the interaction between antithrombin III and a potent heparin analog." *J. Am. Chem. Soc.* **113**(7): 2743-2747.

Guermeur, Y., C. Geourjon, et al. (1999). "Improved performance in protein secondary structure prediction by inhomogeneous score combination." *Bioinformatics* **15**(5): 413-421.

Guerrini, M., S. Guglieri, et al. (2006). "Conformational transitions induced in heparin octasaccharides by binding with antithrombin III." *Biochem. J.* **399**(2): 191-198.

Gunnarsson, G. T. and U. R. Desai (2002). "Designing small, nonsugar activators of antithrombin using hydropathic interaction analyses." *J. Med. Chem.* **45**(6): 1233-1243.

Gunnarsson, G. T. and U. R. Desai (2002). "Interaction of designed sulfated flavanoids with antithrombin: lessons on the design of organic activators." *J. Med. Chem.* **45**(20): 4460-4470.

Guo, N. H., H. C. Krutzsch, et al. (1992). "Heparin- and sulfatide-binding peptides from the type I repeats of human thrombospondin promote melanoma cell adhesion." *Proc. Natl. Acad. Sci. U.S.A.* **89**(7): 3040-3044.

Haasnoot, C. A. G., F. A. A. M. de Leeuw, et al. (1980). "The relationship between proton-proton NMR coupling constants and substituent electronegativities--I : An empirical generalization of the karplus equation." *Tetrahedron* **36**(19): 2783-2792.

Hallgren, J., S. Backstrom, et al. (2004). "Histidines are critical for heparin-dependent activation of mast cell tryptase." *J. Immunol.* **173**(3): 1868-1875.

Hammond, E., I. Bytheway, et al. (2009). The dual angiogenesis/heparanase inhibitor PG545 inhibits solid tumor progression in models of breast, prostate and liver cancer: A comparative assessment of once versus twice weekly administration schedules *AACR-NCI-EORTC International Conference: Molecular Targets and Cancer Therapeutics*. Boston, MA, Mol. Cancer Ther. (Meeting Abstract Supplement).

Hammond, M. E., V. Shyamala, et al. (1996). "Receptor recognition and specificity of interleukin-8 is determined by residues that cluster near a surface-accessible hydrophobic pocket." *J. Biol. Chem.* **271**(14): 8228-8235.

Harenberg, J. (2009). "Development of idraparinux and idrabiotaparinux for anticoagulant therapy." *Thromb. Haemost.* **102**(5): 811-815.

Harenberg, J., I. Jörg, et al. (2008). "Anticoagulant effects of Idraparinux after termination of therapy for prevention of recurrent venous thromboembolism: observations from the van Gogh trials." *Eur. J. Clin. Pharmacol.* **64**(6): 555-563.

Harenberg, J., Y. Vukojevic, et al. (2008). "Long elimination half-life of idraparinux may explain major bleeding and recurrent events of patients from the van Gogh trials." *J. Thromb. Haemost.* **6**(5): 890-892.

Harenberg, J., M. Wehling, et al. (2009). "The anticoagulant Idraparinux: Is the extensive half life of 60 days the cause of bleeding complications." *British Journal of Clinical Pharmacology* **68**(Suppl. 1): 21-21.

Hay, D. W. P. and H. M. Sarau (2001). "Interleukin-8 receptor antagonists in pulmonary diseases." *Curr Opin Pharmacol* **1**(3): 242-247.

Hebert, C. A., F. W. Luscinskas, et al. (1990). "Endothelial and leukocyte forms of IL-8. Conversion by thrombin and interactions with neutrophils." *J. Immunol.* **145**(9): 3033-3040.

Hébert, C. A., R. V. Vitangcol, et al. (1991). "Scanning mutagenesis of interleukin-8 identifies a cluster of residues required for receptor binding." *J. Biol. Chem.* **266**(28): 18989-18994.

Heldin, P., E. Karousou, et al. (2008). "Importance of hyaluronan-CD44 interactions in inflammation and tumorigenesis." *Connect. Tissue Res.* **49**(3): 215-218.

Hendrickx, P. M., F. Corzana, et al. (2010). "The use of time-averaged 3JHH restrained molecular dynamics (tar-MD) simulations for the conformational analysis of five-membered ring systems: methodology and applications." *J Comput Chem* **31**(3): 561-572.

Henrissat, B. and A. Bairoch (1993). "New families in the classification of glycosyl hydrolases based on amino acid sequence similarities." *Biochem. J.* **293** ((Pt 3)): 781-788.

Henrissat, B. and A. Bairoch (1996). "Updating the sequence-based classification of glycosyl hydrolases." *Biochem. J.* **316** ((Pt 2)): 695-696.

Henry, B. L., J. F. Wayne, et al. (1997). "Monomeric variants of IL-8: Effects of side chain substitutions and solution conditions upon dimer formation." *Protein Sci.* **6**(3): 598-608.

Herbert, J. M., J. P. Herault, et al. (1998). "Biochemical and pharmacological properties of SANORG 34006, a potent and long-acting synthetic pentasaccharide." *Blood* **91**(11): 4197-4205.

Herges, T. and W. Wenzel (2004). "An all-atom force field for tertiary structure prediction of helical proteins." *Biophys. J.* **87**(5): 3100-3109.
Hileman, R. E., J. R. Fromm, et al. (1998). "Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins." *Bioessays* **20**(2): 156-167.
Hill, A. D. and P. J. Reilly (2007). "Puckering coordinates of monocyclic rings by triangular decomposition." *J Chem Inf Model* **47**(3): 1031-1035.
Hill, A. D. and P. J. Reilly (2008). "Computational analysis of glycoside hydrolase family 1 specificities." *Biopolymers* **89**(11): 1021-1031.
Hill, A. D. and P. J. Reilly (2008). "A Gibbs free energy correlation for automated docking of carbohydrates." *J Comput Chem* **29**(7): 1131-1141.
Hjelm, R. and S. Schedin-Weiss (2007). "High affinity interaction between a synthetic, highly negatively charged pentasaccharide and alpha- or beta-antithrombin is predominantly due to nonionic interactions." *Biochemistry* **46**(11): 3378-3384.
Hoch, R. C., I. U. Schraufstatter, et al. (1996). "In vivo, in vitro, and molecular aspects of interleukin-8 and the interleukin-8 receptors." *J. Lab. Clin. Med.* **128**(2): 134-145.
Holmes, W. E., J. Lee, et al. (1991). "Structure and functional expression of a human interleukin-8 receptor." *Science* **253**(5025): 1278-1280.
Holt, C. E. and B. J. Dickson (2005). "Sugar codes for axons?" *Neuron* **46**(2): 169-172.
Hong, X., F. Jiang, et al. (2008). "Increased chemotactic migration and growth in heparanase-overexpressing human U251n glioma cells." *J. Exp. Clin. Cancer Res.* **27**: 23.
Honig, B. and A. Nicholls (1995). "Classical electrostatics in biology and chemistry." *Science* **268**(5214): 1144-1149.
Hoover, D. M., J. Shaw, et al. (2000). "The crystal structure of Met-Rantes: comparison with native Rantes and AOP-Rantes." *Protein Pept. Lett.* **7**(2): 73-82.
Hoppensteadt, D., W. Jeske, et al. (2008). "AVE5026: A new hemisynthetic ultra low molecular weight heparin (ULMWH) with enriched anti-Xa activity and enhanced antithrombotic activity for management of cancer associated thrombosis." *J. Clin. Oncol. (Meeting Abstracts)* **26**(15_suppl): 14653.
Hotta, K., K. Hayashi, et al. (1990). "Coding region structure of interleukin-8 gene of human lung giant cell carcinoma LU65C cells that produce LUCT/interleukin-8: homogeneity in interleukin-8 genes." *Immunol. Lett.* **24**(3): 165-169.
Hou, T., J. Wang, et al. (2011). "Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations." *J. Chem. Inf. Model.* **51**(1): 69-82.
Hovel, K., D. Shallom, et al. (2003). "Crystal structure and snapshots along the reaction pathway of a family 51 alpha-L-arabinofuranosidase." *EMBO J.* **22**(19): 4922-4932.
Hovingh, P. and A. Linker (1982). "An unusual heparan sulfate isolated from lobsters (Homarus americanus)." *J. Biol. Chem.* **257**(16): 9840-9844.
Hovingh, P. and A. Linker (1993). "Glycosaminoglycans in Anodonta californiensis, a freshwater mussel." *Biol. Bull.* **185**(2): 263-276.
Hricovini, M. (2006). "B3LYP/6-311++G** study of structure and spin-spin coupling constant in methyl 2-O-sulfo-alpha-L-iduronate." *Carbohydr. Res.* **341**(15): 2575-2580.
Hricovini, M. and F. Bizik (2007). "Relationship between structure and three-bond proton-proton coupling constants in glycosaminoglycans." *Carbohydr. Res.* **342**(6): 779-783.
Hricovini, M., M. Guerrini, et al. (1999). "Structure of heparin-derived tetrasaccharide complexed to the plasma protein antithrombin derived from NOEs, J-couplings and chemical shifts." *Eur. J. Biochem.* **261**(3): 789-801.
Hricovini, M., M. Guerrini, et al. (2002). "Active conformations of glycosaminoglycans. NMR determination of the conformation of heparin sequences complexed with antithrombin and fibroblast growth factors in solution." *Semin. Thromb. Hemost.* **28**(4): 325-334.
Hricovini, M., M. Guerrini, et al. (2001). "Conformation of heparin pentasaccharide bound to antithrombin III." *Biochem. J.* **359**(Pt 2): 265-272.
Hricovíni, M., P. M. Nieto, et al. (2003). NMR of sulfated oligo- and polysaccharides. *NMR Spectroscopy of Glycoconjugates*. J. Jiménez-Barbero and T. Peters. Weinheim, FRG, Wiley-VCH Verlag GmbH & Co. KGaA**:** 189-229.

Hsu, S. T., C. Peter, et al. (2005). "Entropy calculation of HIV-1 Env gp120, its receptor CD4, and their complex: an analysis of configurational entropy changes upon complexation." *Biophys. J.* **88**(1): 15-24.

Huang, L. and R. J. Kerns (2006). "Diversity-oriented chemical modification of heparin: Identification of charge-reduced N-acyl heparin derivatives having increased selectivity for heparin-binding proteins." *Bioorg. Med. Chem.* **14**(7): 2300-2313.

Huey, R., G. M. Morris, et al. (2007). "A semiempirical free energy force field with charge-based desolvation." *J Comput Chem* **28**(6): 1145-1152.

Huige, C. J. M. and C. Altona (1995). "Force field parameters for sulfates and sulfamates based on ab initio calculations: Extensions of AMBER and CHARMm fields." *J Comput Chem* **16**(1): 56-79.

Hulett, M. D., C. Freeman, et al. (1999). "Cloning of mammalian heparanase, an important enzyme in tumor invasion and metastasis." *Nat. Med.* **5**(7): 803-809.

Hulett, M. D., J. R. Hornby, et al. (2000). "Identification of active-site residues of the pro-metastatic endoglycosidase heparanase." *Biochemistry* **39**(51): 15659-15667.

Ihrcke, N. S., W. Parker, et al. (1998). "Regulation of platelet heparanase during inflammation: role of pH and proteinases." *J. Cell. Physiol.* **175**(3): 255-267.

Iizasa, H. and K. Matsushima (2001). IL-8. *Cytokine reference: a compendium of cytokines and other mediators of host defense*. J. J. Oppenheim, M. Feldmann, S. K. Durumet al. New York, U.S.A., Academic Press. **I:** 1061-1068.

Imberty, A., E. Bettler, et al. (1999). Building sugars: the sweet part of structural biology. *Perspectives in Structural Biology.* M. Vijayan, N. Yathindra and A. S. Kolaskar. Hyderabad, India, Indian Academy of Sciences and Universities Press**:** 392–409.

Imberty, A., H. Lortat-Jacob, et al. (2007). "Structural view of glycosaminoglycan-protein interactions." *Carbohydr. Res.* **342**(3-4): 430-439.

Inoue, Y., M. Endo, et al. (2003). "Molecular cloning and sequencing of the silver chimaera (Chimaera phantasma) interleukin-8 cDNA." *Fish Shellfish Immunol.* **15**(3): 269-274.

Iozzo, R. V., I. R. Cohen, et al. (1994). "The biology of perlecan: the multifaceted heparan sulphate proteoglycan of basement membranes and pericellular matrices." *Biochem. J.* **302** (Pt 3): 625-639.

Iozzo, R. V. and J. D. San Antonio (2001). "Heparan sulfate proteoglycans: heavy hitters in the angiogenesis arena." *J. Clin. Invest.* **108**(3): 349-355.

Irimura, T., M. Nakajima, et al. (1986). "Chemically modified heparins as inhibitors of heparan sulfate specific endo-beta-glucuronidase (heparanase) of metastatic melanoma cells." *Biochemistry* **25**(18): 5322-5328.

Ishida, K., G. Hirai, et al. (2004). "Structure-based design of a selective heparanase inhibitor as an antimetastatic agent." *Mol. Cancer Ther.* **3**(9): 1069-1077.

Ishida, K., M. K. Wierzba, et al. (2004). "Novel heparan sulfate mimetic compounds as antitumor agents." *Chem. Biol.* **11**(3): 367-377.

Isorna, P., J. Polaina, et al. (2007). "Crystal structures of Paenibacillus polymyxa beta-glucosidase B complexes reveal the molecular basis of substrate specificity and give new insights into the catalytic machinery of family I glycosidases." *J. Mol. Biol.* **371**(5): 1204-1218.

Itano, N. (2008). "Simple primary structure, complex turnover regulation and multiple roles of hyaluronan." *J. Biochem.* **144**(2): 131-137.

Jackson, R. L., S. J. Busch, et al. (1991). "Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes." *Physiol. Rev.* **71**(2): 481-539.

Jain, S., W. B. Drendel, et al. (1996). "Structure of human beta-glucuronidase reveals candidate lysosomal targeting and active-site motifs." *Nat. Struct. Biol.* **3**(4): 375-381.

Jairajpuri, M. A., A. Lu, et al. (2003). "Antithrombin III phenylalanines 122 and 121 contribute to its high affinity for heparin and its conformational activation." *J. Biol. Chem.* **278**(18): 15941-15950.

Jalali-Heravi, M., M. Asadollahi-Baboli, et al. (2008). "QSAR study of heparanase inhibitors activity using artificial neural networks and Levenberg-Marquardt algorithm." *Eur J Med Chem* **43**(3): 548-556.

Jeffrey, G. A. and J. H. Yates (1979). "Stereographic representation of the cremer-pople ring-puckering parameters for pyranoid rings." *Carbohydr. Res.* **74**(1): 319-322.

Jeske, W., A. Brubaker, et al. (2009). Novel antagonists for low molecular weight heparin and heparin-like drugs. *American Society of Hematology*. New Orleans, LA.

Jeske, W., A. Brubaker, et al. (2007). *In vitro* characterization of the neutralization of unfractionated heparin and low molecular weight heparin by novel salicylamide derivatives. *American Society of Hematology*. Atlanta, Georgia.

Jin, L. (2007). Ph.D., University of Edinburgh.

Jin, L., J. P. Abrahams, et al. (1997). "The anticoagulant activation of antithrombin by heparin." *Proc. Natl. Acad. Sci. U.S.A.* **94**(26): 14683-14688.

Jin, L., P. E. Barran, et al. (2005). "Conformation of glycosaminoglycans by ion mobility mass spectrometry and molecular modelling." *Phys Chem Chem Phys* **7**(19): 3464-3471.

Jin, L., M. Hricovini, et al. (2009). "Residual dipolar coupling investigation of a heparin tetrasaccharide confirms the limited effect of flexibility of the iduronic acid on the molecular shape of heparin." *Glycobiology* **19**(11): 1185-1196.

JoAnne, M., F. Trudy, et al. (1999). "Interactions of Alzheimer amyloid-β peptides with glycosaminoglycans." *Eur. J. Biochem.* **266**(3): 1101-1110.

Johnson, D. J., W. Li, et al. (2006). "Antithrombin-S195A factor Xa-heparin structure reveals the allosteric mechanism of antithrombin activation." *EMBO J.* **25**(9): 2029-2037.

Johnson, Z., C. A. Power, et al. (2004). "Chemokine inhibition--why, when, where, which and how?" *Biochem. Soc. Trans.* **32**(Pt 2): 366-377.

Johnson, Z., A. E. Proudfoot, et al. (2005). "Interaction of chemokines and glycosaminoglycans: a new twist in the regulation of chemokine function with opportunities for therapeutic intervention." *Cytokine Growth Factor Rev.* **16**(6): 625-636.

Johnstone, K. D., T. Karoli, et al. (2010). "Synthesis and biological evaluation of polysulfated oligosaccharide glycosides as inhibitors of angiogenesis and tumor growth." *J. Med. Chem.* **53**(4): 1686-1699.

Jones, D. T., W. R. Taylor, et al. (1992). "The rapid generation of mutation data matrices from protein sequences." *Comput. Appl. Biosci.* **8**(3): 275-282.

Jorgensen, W. L., J. Chandrasekhar, et al. (1983). "Comparison of simple potential functions for simulating liquid water." *J Chem Phys* **79**: 926-935.

Juhl, P. B., P. Trodler, et al. (2009). "Modelling substrate specificity and enantioselectivity for lipases and esterases by substrate-imprinted docking." *BMC Struct. Biol.* **9**: 39.

Kagiampakis, I., H. Jin, et al. (2008). "Conservation of unfavorable sequence motifs that contribute to the chemokine quaternary state." *Biochemistry* **47**(40): 10637-10648.

Kanagarajadurai, K. and R. Sowdhamini (2008). "Sequence and structural analyses of interleukin-8-like chemokine superfamily." *In Silico Biol.* **8**(3): 307-330.

Karoli, T., L. Liu, et al. (2005). "Synthesis, biological activity, and preliminary pharmacokinetic evaluation of analogues of a phosphosulfomannan angiogenesis inhibitor (PI-88)." *J. Med. Chem.* **48**(26): 8229-8236.

Kassam, G., A. Manro, et al. (1997). "Characterization of the heparin binding properties of annexin II tetramer." *J. Biol. Chem.* **272**(24): 15093-15100.

Katoh, K. and H. Toh (2008). "Recent developments in the MAFFT multiple sequence alignment program." *Brief. Bioinformatics* **9**(4): 286-298.

Kelley, L. A. and M. J. Sternberg (2009). "Protein structure prediction on the Web: a case study using the Phyre server." *Nat Protoc* **4**(3): 363-371.

Kellogg, G., G. Joshi, et al. (1992). "New tools for modeling and understanding hydrophobicity and hydrophobic interactions." *Med Chem Res* **1**: 444-453.

Kellogg, G. E., S. F. Semus, et al. (1991). "HINT: a new method of empirical hydrophobic field calculation for CoMFA." *J. Comput. Aided Mol. Des.* **5**(6): 545-552.

Kern, A., K. Schmidt, et al. (2003). "Identification of a heparin-binding motif on adeno-associated virus type 2 capsids." *J. Virol.* **77**(20): 11072-11081.

Kerzmann, A., J. Fuhrmann, et al. (2008). "BALLDock/SLICK: a new method for protein-carbohydrate docking." *J. Chem. Inf. Model.* **48**(8): 1616-1625.

Kerzmann, A., D. Neumann, et al. (2006). "SLICK--scoring and energy functions for protein-carbohydrate interactions." *J. Chem. Inf. Model.* **46**(4): 1635-1642.

Keserü, G. and I. Kolossváry (1999). *Molecular mechanics and conformational analysis in drug design*. Oxford, Blackwell Science.

Khammari Chebbi, C., K. Kichenin, et al. (2008). "Pilot study of a new matrix therapy agent (RGTA OTR4120®) in treatment-resistant corneal ulcers and corneal dystrophy " *J Fr Ophtalmol* **31**(5): 465-471.

Khan, S., E. Rodriguez, et al. (2011). "The solution structure of heparan sulphate differs from that of heparin: implications for function." *J Biol Chem*.

Khasraw, M., N. Pavlakis, et al. (2009). Multicentre phase I/II study of PI-88, a heparanase inhibitor in combination with docetaxel in patients with metastatic castrate-resistant prostate cancer, Ann. Oncol.

Kim, S. H. and K. L. Kiick (2007). "Heparin-mimetic sulfated peptides with modulated affinities for heparin-binding peptides and growth factors." *Peptides* **28**(11): 2125-2136.

King, R. D. and M. J. Sternberg (1996). "Identification and application of the concepts important for accurate and reliable protein secondary structure prediction." *Protein Sci.* **5**(11): 2298-2310.

Kirschner, K. N. and R. J. Woods (2001). "Solvent interactions determine carbohydrate conformation." *Proc. Natl. Acad. Sci. U.S.A.* **98**(19): 10541-10545.

Kirschner, K. N., A. B. Yongye, et al. (2008). "GLYCAM06: A generalizable biomolecular force field. Carbohydrates." *J Comput Chem* **29**(4): 622-655.

Kishimoto, T. K., Y. W. Qi, et al. (2009). "M118–A rationally engineered low-molecular-weight heparin designed specifically for the treatment of acute coronary syndromes." *Thromb. Haemost.* **102**: 900-906.

Kisilevsky, R. (2000). "The relation of proteoglycans, serum amyloid P and Apo E to amyloidosis current status, 2000." *Amyloid* **7**(1): 23 - 25.

Kisilevsky, R., J. B. Ancsin, et al. (2007). "Heparan sulfate as a therapeutic target in amyloidogenesis: prospects and possible complications." *Amyloid* **14**(1): 21-32.

Kisilevsky, R., L. J. Lemieux, et al. (1995). "Arresting amyloidosis in vivo using small-molecule anionic sulphonates or sulphates: Implications for Alzheimer's disease." *Nat. Med.* **1**(2): 143-148.

Kitago, Y., S. Karita, et al. (2007). "Crystal structure of Cel44A, a glycoside hydrolase family 44 endoglucanase from Clostridium thermocellum." *J. Biol. Chem.* **282**(49): 35703-35711.

Klement, P. and J. Rak (2006). "Emerging anticoagulants: mechanism of action and future potential." *Vnitr Lek* **52**(Suppl 1): 119-122.

Koch, A. E., P. J. Polverini, et al. (1992). "Interleukin-8 as a macrophage-derived mediator of angiogenesis." *Science* **258**(5089): 1798-1801.

Kokenyesi, R. and M. Bernfield (1994). "Core protein structure and sequence determine the site and presence of heparan sulfate and chondroitin sulfate on syndecan-1." *J. Biol. Chem.* **269**(16): 12304-12309.

Koliopanos, A., H. Friess, et al. (2001). "Heparanase expression in primary and metastatic pancreatic cancer." *Cancer Res.* **61**(12): 4655-4659.

Kollman, P. A., I. Massova, et al. (2000). "Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models." *Acc. Chem. Res.* **33**(12): 889-897.

Kongsted, J. and U. Ryde (2009). "An improved method to predict the entropy term with the MM/PBSA approach." *J. Comput. Aided Mol. Des.* **23**(2): 63-71.

Kongsted, J., U. Ryde, et al. (2007). "Prediction and rationalization of the pH dependence of the activity and stability of family 11 xylanases." *Biochemistry* **46**(47): 13581-13592.

Kony, D., W. Damm, et al. (2002). "An improved OPLS-AA force field for carbohydrates." *J Comput Chem* **23**(15): 1416-1429.

Koropatkin, N. M. and T. J. Smith (2010). "SusG: a unique cell-membrane-associated alpha-amylase from a prominent human gut symbiont targets complex starch molecules." *Structure* **18**(2): 200-215.

Krieger, E., E. Geretti, et al. (2004). "A structural and dynamic model for the interaction of interleukin-8 and glycosaminoglycans: support from isothermal fluorescence titrations." *Proteins* **54**(4): 768-775.

Krissinel, E. and K. Henrick (2004). "Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions." *Acta Crystallogr. Sect. D* **60**(12 Part 1): 2256-2268.

Krissinel, E. and K. Henrick. (2009). "Protein structure comparison service SSM at European Bioinformatics Institute. http://www.ebi.ac.uk/msd-srv/ssm." Retrieved November 15, 2009.

Kudchadkar, R., R. Gonzalez, et al. (2008). "PI-88: a novel inhibitor of angiogenesis." *Expert Opin Investig Drugs* **17**(11): 1769-1776.
Kuschert, G. S., F. Coulin, et al. (1999). "Glycosaminoglycans interact selectively with chemokines and modulate receptor binding and cellular responses." *Biochemistry* **38**(39): 12959-12968.
Kuschert, G. S., A. J. Hoogewerf, et al. (1998). "Identification of a glycosaminoglycan binding surface on human interleukin-8." *Biochemistry* **37**(32): 11193-11201.
Kussie, P. H., J. D. Hulmes, et al. (1999). "Cloning and functional expression of a human heparanase gene." *Biochem. Biophys. Res. Commun.* **261**(1): 183-187.
Kuttel, M., J. W. Brady, et al. (2002). "Carbohydrate solution simulations: producing a force field with experimentally consistent primary alcohol rotational frequencies and populations." *J Comput Chem* **23**(13): 1236-1243.
Kuziej, J., E. Litinas, et al. (2009). "Neutralization of hemorrhagic and antithrombotic activities of heparins by a novel salicylamide derivative." *FASEB J.* **23**(1_MeetingAbstracts): 569.566-.
Laederach, A., M. K. Dowd, et al. (1999). "Automated docking of maltose, 2-deoxymaltose, and maltotetraose into the soybean β-amylase active site." *Proteins* **37**(2): 166-175.
Laederach, A. and P. J. Reilly (2003). "Specific empirical free energy function for automated docking of carbohydrates to proteins." *J Comput Chem* **24**(14): 1748-1757.
Laguri, C., F. Arenzana-Seisdedos, et al. (2008). "Relationships between glycosaminoglycan and receptor binding sites in chemokines--the CXCL12 example." *Carbohydr. Res.* **343**(12): 2018-2023.
Lai, N. S., S. Simizu, et al. (2008). "Requirement of the conserved, hydrophobic C-terminus region for the activation of heparanase." *Exp. Cell Res.* **314**(15): 2834-2845.
Laitinen, T., J. Rouvinen, et al. (2003). "MM-PBSA free energy analysis of endo-1,4-xylanase II (XynII)-substrate complexes: binding of the reactive sugar in a skew boat and chair conformation." *Org. Biomol. Chem.* **1**(20): 3535-3540.
Lam, K., V. S. Rao, et al. (1998). "Molecular modeling studies on binding of bFGF to heparin and its receptor FGFR1." *J. Biomol. Struct. Dyn.* **15**(6): 1009-1027.
Landau, M., I. Mayrose, et al. (2005). "ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures." *Nucleic Acids Res.* **33**(Web Server issue): W299-W302.
LaRosa, G. J., K. M. Thomas, et al. (1992). "Amino terminus of the interleukin-8 receptor is a major determinant of receptor subtype specificity." *J. Biol. Chem.* **267**(35): 25402-25406.
Larsen, C. G., A. O. Anderson, et al. (1989). "The neutrophil-activating protein (NAP-1) is also chemotactic for T lymphocytes." *Science* **243**(4897): 1464-1466.
Lassen, M. R., O. E. Dahl, et al. (2009). "AVE5026, a new hemisynthetic ultra-low-molecular-weight heparin for the prevention of venous thromboembolism in patients after total knee replacement surgery - TREK: a dose-ranging study." *J. Thromb. Haemost.* **7**(4): 566-572.
Lau, E. K., S. Allen, et al. (2004). "Chemokine-receptor interactions: GPCRs, glycosaminoglycans and viral chemokine binding proteins." *Adv. Protein Chem.* **68**: 351-391.
Laurent, T. C., U. B. Laurent, et al. (1996). "The structure and function of hyaluronan: An overview." *Immunol. Cell Biol.* **74**(2): A1-A7.
Lawrence, R., H. Lu, et al. (2008). "Disaccharide structure code for the easy representation of constituent oligosaccharides from glycosaminoglycans." *Nat. Methods* **5**(4): 291-292.
Ledoux, D., D. Papy-Garcia, et al. (2000). "Human plasmin enzymatic activity is inhibited by chemically modified dextrans." *J. Biol. Chem.* **275**(38): 29383-29390.
Lee, D., S. Lee, et al. (2009). "Antiangiogenic activity of orally absorbable heparin derivative in different types of cancer cells." *Pharm Res* **26**(12): 2667-2676.
Lee, E.-Y., H.-H. Park, et al. (2001). "Cloning and sequence analysis of the interleukin-8 gene from flounder (Paralichthys olivaceous)." *Gene* **274**(1-2): 237-243.
Lee, J., R. Horuk, et al. (1992). "Characterization of two high affinity human interleukin-8 receptors." *J. Biol. Chem.* **267**(23): 16283-16287.
Leong, S. R., R. C. Kabakoff, et al. (1994). "Complete mutagenesis of the extracellular domain of interleukin-8 (IL- 8) type A receptor identifies charged residues mediating IL-8 binding and signal transduction." *J. Biol. Chem.* **269**(30): 19343-19348.
Leong, S. R., H. B. Lowman, et al. (1997). "IL-8 single-chain homodimers and heterodimers: interactions with chemokine receptors CXCR1, CXCR2, and DARC." *Protein Sci.* **6**(3): 609-617.

Lever, R., A. Smailbegovic, et al. (2001). "Role of glycosaminoglycans in inflammation." *Inflammopharmacology* **9**(1): 165-169.
Leveugle, B., A. Scanameo, et al. (1994). "Binding of heparan sulfate glycosaminoglycan to [beta]-amyloid peptide: inhibition by potentially therapeutic polysulfated compounds." *NeuroReport* **5**(11): 1389-1392.
Levy-Adam, F., G. Abboud-Jarrous, et al. (2005). "Identification and characterization of heparin/heparan sulfate binding domains of the endoglycosidase heparanase." *J. Biol. Chem.* **280**(21): 20457-20466.
Levy-Adam, F., S. Feld, et al. (2010). "Heparanase 2 interacts with heparan sulfate with high affinity and inhibits heparanase activity." *J. Biol. Chem.* **285**(36): 28010-28019.
Levy-Adam, F., N. Ilan, et al. (2010). "Tumorigenic and adhesive properties of heparanase." *Semin. Cancer Biol.*. **20**(3): 153-160.
Lewis, K., W. Robinson, et al. (2008). "A phase II study of the heparanase inhibitor PI-88 in patients with advanced melanoma." *Invest New Drugs* **26**(1): 89-94.
Ley, K., C. Laudanna, et al. (2007). "Getting to the site of inflammation: the leukocyte adhesion cascade updated." *Nat. Rev. Immunol.* **7**(9): 678-689.
Li, F. and J. R. Gordon (2001). "IL-8(3-73)K11R is a high affinity agonist of the neutrophil CXCR1 and CXCR2." *Biochem. Biophys. Res. Commun.* **286**(3): 595-600.
Li, F., X. Zhang, et al. (2002). "CXCL8(3-73)K11R/G31P antagonizes ligand binding to the neutrophil CXCR1 and CXCR2 receptors and cellular responses to CXCL8/IL-8." *Biochem. Biophys. Res. Commun.* **293**(3): 939-944.
Li, W., D. J. Johnson, et al. (2004). "Structure of the antithrombin-thrombin-heparin ternary complex reveals the antithrombotic mechanism of heparin." *Nat. Struct. Mol. Biol.* **11**(9): 857-862.
Lieth, C.-W. v. d., M. Frank, et al. (2009). *Bioinformatics for Glycobiology and Glycomics : An Introduction*, John Wiley.
Lindahl, U. (2007). "Heparan sulfate-protein interactions--a concept for drug design?" *Thromb. Haemost.* **98**(1): 109-115.
Lindahl, U. and L. Kjellen (1991). "Heparin or heparan sulfate--what is the difference?" *Thromb. Haemost.* **66**(1): 44-48.
Lindahl, U., M. Kusche-Gullberg, et al. (1998). "Regulated diversity of heparan sulfate." *J. Biol. Chem.* **273**(39): 24979-24982.
Lindorff-Larsen, K., S. Piana, et al. (2010). "Improved side-chain torsion potentials for the Amber ff99SB protein force field." *Proteins* **78**(8): 1950-1958.
Linhardt, R. J., S. A. Ampofo, et al. (1992). "Isolation and characterization of human heparin." *Biochemistry* **31**(49): 12441-12445.
Linhardt, R. J. and N. S. Gunay (1999). "Production and chemical processing of low molecular weight heparins." *Semin. Thromb. Hemost.* **25** (Suppl 3): 5-16.
Linhardt, R. J., J. E. Turnbull, et al. (1990). "Examination of the substrate specificity of heparin and heparan sulfate lyases." *Biochemistry* **29**(10): 2611-2617.
Lins, R. D. and P. H. Hunenberger (2005). "A new GROMOS force field for hexopyranose-based carbohydrates." *J Comput Chem* **26**(13): 1400-1412.
Liu, D., Z. Shriver, et al. (2002). "Dynamic regulation of tumor growth and metastasis by heparan sulfate glycosaminoglycans." *Semin. Thromb. Hemost.* **28**(1): 67-78.
Liu, L., I. Bytheway, et al. (2008). "Design, synthesis, FGF-1 binding, and molecular modeling studies of conformationally flexible heparin mimetic disaccharides." *Bioorg. Med. Chem. Lett.* **18**(1): 344-349.
Liu, Y., S. Yang, et al. (2005). "Molecular evolution of CXCR1, a G protein-coupled receptor involved in signal transduction of neutrophils." *J. Mol. Evol.* **61**(5): 691-696.
Liu, Z. and Y. Zhang (2009). "Molecular dynamics simulations and MM-PBSA calculations of the lectin from snowdrop (Galanthus nivalis)." *J Mol Model* **15**(12): 1501-1507.
Lohse, D. L. and R. J. Linhardt (1992). "Purification and characterization of heparin lyases from Flavobacterium heparinum." *J Biol Chem* **267**(34): 24347-24355.
Lolkema, M. P., M. Lockley, et al. (2010). M402, a novel heparan sulfate mimetic, synergizes with Gemcitabine to improve survival and reduce metastasis and epithelial-to-mesenchymal transition (EMT) in a genetically engineered mouse model for pancreatic cancer. *American Association for Cancer Research (AACR)*. Washington, DC.

Lortat-Jacob, H., A. Grosdidier, et al. (2002). "Structural diversity of heparan sulfate binding domains in chemokines." *Proc. Natl. Acad. Sci. U.S.A.* **99**(3): 1229-1234.
Lortat-Jacob, H., J. E. Turnbull, et al. (1995). "Molecular organization of the interferon gamma-binding domain in heparan sulphate." *Biochem. J.* **310** (Pt 2): 497-505.
Lowman, H. B., W. J. Fairbrother, et al. (1997). "Monomeric variants of IL-8: Effects of side chain substitutions and solution conditions upon dimer formation." *Protein Sci.* **6**(3): 598-608.
Lu, Z.-H., Z.-x. Wang, et al. (1995). "The promiscuous chemokine binding profile of the duffy antigen/receptor for chemokines is primarily localized to sequences in the amino-terminal domain." *J. Biol. Chem.* **270**(44): 26239-26245.
Lubkowski, J., G. Bujacz, et al. (1997). "The structure of MCP-1 in two crystal forms provides a rare example of variable quaternary interactions." *Nat. Struct. Biol.* **4**(1): 64-69.
Ludwig, R. J. (2009). "Therapeutic use of heparin beyond anticoagulation." *Curr Drug Discov Technol* **6**(4): 281-289.
Luo, Z., D. J. Butcher, et al. (1997). "Molecular modeling of interleukin-8 receptor beta and analysis of the receptor-ligand interaction." *Protein Eng.* **10**(9): 1039-1045.
Luthy, R., J. U. Bowie, et al. (1992). "Assessment of protein models with three-dimensional profiles." *Nature* **356**(6364): 83-85.
Lynch 3rd, J. P., T. J. Standiford, et al. (1992). "Neutrophilic alveolitis in idiopathic pulmonary fibrosis. The role of interleukin-8." *Am. Rev. Respir. Dis.* **145**(6): 1433-1439.
M. Kuttel, M. (2011). "The conformational free energy of carbohydrates." *Mini Rev Org Chem* **8**(3): 256-262.
Ma, B. and R. Nussinov (2006). "Simulations as analytical tools to understand protein aggregation and predict amyloid conformation." *Curr. Opin. Chem. Biol.* **10**(5): 445-452.
Ma, Q. and J. Fareed (2004). "Idraparinux sodium." *IDrugs* **7**(11): 1028-1034.
Malik, A. and S. Ahmad (2007). "Sequence and structural features of carbohydrate binding in proteins and assessment of predictability using a neural network." *BMC Struct. Biol.* **7**: 1.
Malik, Z. A. and B. F. Tack (2006). "Structure of human MIP-3[alpha] chemokine." *Acta Cryst. Section F* **62**(7): 631-634.
Manenti, L., P. Tansinda, et al. (2008). "Eprodisate in amyloid A amyloidosis: a novel therapeutic approach?" *Expert Opin Pharmacother* **9**(12): 2175-2180.
Maple, J. R. (1998). "Derivation of class II force fields: V. Quantum force field for amides, peptides, and related compounds." *J Comput Chem* **19**(4): 430-458.
Marchetti, D., S. Liu, et al. (1997). "Heparanase and a synthetic peptide of heparan sulfate-interacting protein recognize common sites on cell surface and extracellular matrix heparan sulfate." *J. Biol. Chem.* **272**(25): 15891-15897.
Marcum, J. A. (2000). "The origin of the dispute over the discovery of heparin." *J Hist Med Allied Sci* **55**(1): 37-66.
Margalit, H., N. Fischer, et al. (1993). "Comparative analysis of structurally defined heparin binding sequences reveals a distinct spatial distribution of basic residues." *J. Biol. Chem.* **268**(26): 19228-19231.
Mark, R. M., R. A. Harold, et al. (2003). "Gaussian docking functions." *Biopolymers* **68**(1): 76-90.
Martelly, I., D. Singabraya, et al. (2010). "Glycosaminoglycan mimetics trigger IP3-dependent intracellular calcium release in myoblasts." *Matrix Biol.* **29**(4): 317-329.
Massova, I. and P. A. Kollman (1999). "Computational alanine scanning to probe protein−protein interactions: A novel approach to evaluate binding free energies." *J. Am. Chem. Soc.* **121**(36): 8133-8143.
Matsumoto, R., A. Sali, et al. (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**(33): 19524-19531.
Matsushima, K., K. Morishita, et al. (1988). "Molecular cloning of a human monocyte-derived neutrophil chemotactic factor (MDNCF) and the induction of MDNCF mRNA by interleukin 1 and tumor necrosis factor." *J. Exp. Med.* **167**(6): 1883-1893.
McCoy, A. J., X. Y. Pei, et al. (2003). "Structure of β-antithrombin and the effect of glycosylation on antithrombin's heparin affinity and activity." *J Mol Biol* **326**(3): 823-833.
McDowell, S. E., N. Spackova, et al. (2007). "Molecular dynamics simulations of RNA: an in silico single molecule approach." *Biopolymers* **85**(2): 169-184.

McKenzie, E., K. Tyson, et al. (2000). "Cloning and expression profiling of Hpa2, a novel mammalian heparanase family member." *Biochem. Biophys. Res. Commun.* **276**(3): 1170-1177.
McKenzie, E. A. (2007). "Heparanase: a target for drug discovery in cancer and inflammation." *Br. J. Pharmacol.* **151**(1): 1-14.
Meddahi, A., C. Alexakis, et al. (2002). "Heparin-like polymer improved healing of gastric and colic ulceration." *J Biomed Mater Res A* **60**(3): 497-501.
Meddahi, A., J. Benoit, et al. (1996). "Heparin-like polymers derived from dextran enhance colonic anastomosis resistance to leakage." *J. Biomed. Mater. Res.* **31**(3): 293-297.
Meddahi, A., F. Blanquaert, et al. (1994). "New approaches to tissue regeneration and repair." *Pathol. Res. Pract.* **190**(9-10): 923-928.
Meddahi, A., F. Brée, et al. (2002). "Pharmacological studies of $RGTA_{11}$, a heparan sulfate mimetic polymer, efficient on muscle regeneration." *J. Biomed. Mater. Res.* **62**(4): 525-531.
Meddahi, A., H. Lemdjabar, et al. (1995). "Inhibition by dextran derivatives of FGF-2 plasmin-mediated degradation." *Biochimie* **77**(9): 703-706.
Meddahi, A., H. Lemdjabar, et al. (1996). "FGF protection and inhibition of human neutrophil elastase by carboxymethyl benzylamide sulfonate dextran derivatives." *Int. J. Biol. Macromol.* **18**(1-2): 141-145.
Medeiros, G. F., A. Mendes, et al. (2000). "Distribution of sulfated glycosaminoglycans in the animal kingdom: widespread occurrence of heparin-like compounds in invertebrates." *Biochim. Biophys. Acta* **1475**(3): 287-294.
Meliciani, I., K. Klenin, et al. (2009). "Probing hot spots on protein-protein interfaces with all-atom free-energy simulation." *J Chem Phys* **131**(3): 034114.
Mertz, B., A. D. Hill, et al. (2007). "Automated docking to explore subsite binding by glycoside hydrolase family 6 cellobiohydrolases and endoglucanases." *Biopolymers* **87**(4): 249-260.
Miao, H.-Q., H. Liu, et al. (2006). "Development of heparanase inhibitors for anti-Cancer therapy." *Curr. Med. Chem.* **13**(18): 2101-2111.
Middleton, J., A. M. Patterson, et al. (2002). "Leukocyte extravasation: chemokine transport and presentation by the endothelium." *Blood* **100**(12): 3853-3860.
Mikami, S., K. Ohashi, et al. (2001). "Loss of syndecan-1 and increased expression of heparanase in invasive esophageal carcinomas." *Jpn. J. Cancer Res.*. **92**(10): 1062-1073.
Mikami, S., M. Oya, et al. (2008). "Expression of heparanase in renal cell carcinomas: implications for tumor invasion and prognosis." *Clin. Cancer Res.* **14**(19): 6055-6061.
Mikhailov, D., R. J. Linhardt, et al. (1997). "NMR solution conformation of heparin-derived hexasaccharide." *Biochem. J.* **328**(Pt 1): 51-61.
Mikhailov, D., K. H. Mayo, et al. (1996). "NMR solution conformation of heparin-derived tetrasaccharide." *Biochem. J.* **318**(Pt 1): 93-102.
Miller, E. J., A. B. Cohen, et al. (1992). "Elevated levels of NAP-1/interleukin-8 are present in the airspaces of patients with the adult respiratory distress syndrome and are associated with increased mortality." *Am. Rev. Respir. Dis.* **146**(2): 427-432.
Millward, M., A. Hamilton, et al. (2007). "Final results of a phase I study of daily PI-88 as a single agent and in combination with dacarbazine (D) in patients with metastatic melanoma." *J. Clin. Oncol. (Meeting Abstracts)* **25**(18_suppl): 8532.
Minar, E. and T. P. Investigators (2004). "A novel long-acting synthetic factor Xa inhibitor (SanOrg34006) to replace warfarin for secondary prevention in deep vein thrombosis. A Phase II evaluation (vol 2, pg 47, 2004)." *J. Thromb. Haemost.* **2**(3): 540-540.
Minke, W. E., D. J. Diller, et al. (1999). "The role of waters in docking strategies with incremental flexibility for carbohydrate derivatives: heat-labile enterotoxin, a multivalent test case." *J. Med. Chem.* **42**(10): 1778-1788.
Mishra, N. K., Z. Kriz, et al. (2010). "Recognition of selected monosaccharides by Pseudomonas aeruginosa Lectin II analyzed by molecular dynamics and free energy calculations." *Carbohydr. Res.* **345**(10): 1432-1441.
Mitchell, J. B. O., R. A. Laskowski, et al. (1999). "BLEEP—potential of mean force describing protein–ligand interactions: I. Generating potential." *J Comput Chem* **20**(11): 1165-1176.
Modi, W. S. and Z. Q. Chen (1998). "Localization of the human CXC chemokine subfamily on the long arm of chromosome 4 using radiation hybrids." *Genomics* **47**(1): 136-139.
Modi, W. S., M. Dean, et al. (1989). "Chromosome mapping and RFLP analyses of monocyte-derived neutrophil chemotactic factor (MDNCF/IL-8)." *Cytogenet. Cell Genet.* **51**: 1046.

Modi, W. S., M. Dean, et al. (1990). "Monocyte-derived neutrophil chemotactic factor (MDNCF/IL-8) resides in a gene cluster along with several other members of the platelet factor 4 gene superfamily." *Hum. Genet.* **84**(2): 185-187.

Monien, B. H., K. I. Cheang, et al. (2005). "Mechanism of poly(acrylic acid) acceleration of antithrombin inhibition of thrombin: implications for the design of novel heparin mimics." *J. Med. Chem.* **48**(16): 5360-5368.

Monien, B. H., B. L. Henry, et al. (2006). "Novel chemo-enzymatic oligomers of cinnamic acids as direct and indirect inhibitors of coagulation proteinases." *Bioorg. Med. Chem.* **14**(23): 7988-7998.

Morris, G. M., D. S. Goodsell, et al. (1998). "Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function." *J Comput Chem* **19**(14): 1639-1662.

Morris, G. M., R. Huey, et al. (2009). "AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility." *J Comput Chem* **30**(16): 2785-2791.

Moser, B., B. Dewald, et al. (1993). "Interleukin-8 antagonists generated by N-terminal modification." *J. Biol. Chem.* **268**(10): 7125-7128.

Mosulen, S., L. Orti, et al. (2011). "Production of heparanase constructs suitable for nuclear magnetic resonance and drug discovery studies." *Biopolymers* **95**(2): 151-160.

Moura-Tamames, S. A., M. J. Ramos, et al. (2009). "Modelling beta-1,3-exoglucanase-saccharide interactions: structure of the enzyme-substrate complex and enzyme binding to the cell wall." *J. Mol. Graph. Model.* **27**(8): 908-920.

Mousa, S. A., R. Linhardt, et al. (2006). "Anti-metastatic effect of a non-anticoagulant low-molecular-weight heparin versus the standard low-molecular-weight heparin, enoxaparin." *Thromb Haemost* **96**(6): 816-821.

Mukherjee, S., A. Roy, et al. (2010). PSFloger: an on-line service system for identifying protein structural and functional analogs. Ann Arbor, University of Michigan.

Mulakala, C. and P. J. Reilly (2002). "Understanding protein structure-function relationships in Family 47 alpha-1,2-mannosidases through computational docking of ligands." *Proteins* **49**(1): 125-134.

Muller, W. A., S. A. Weigl, et al. (1993). "PECAM-1 is required for transendothelial migration of leukocytes." *J. Exp. Med.* **178**(2): 449-460.

Mulloy, B., D. T. Crane, et al. (1996). "The interaction between heparin and polylysine: a circular dichroism and molecular modelling study." *Braz. J. Med. Biol. Res.* **29**(6): 721-729.

Mulloy, B. and M. J. Forster (2000). "Conformation and dynamics of heparin and heparan sulfate." *Glycobiology* **10**(11): 1147-1156.

Mulloy, B. and M. J. Forster (2008). "Application of drug discovery software to the identification of heparin-binding sites on protein surfaces: a computational survey of the 4-helix cytokines." *Mol Simul* **34**(4): 481 - 489.

Mulloy, B., M. J. Forster, et al. (1993). "N.M.R. and molecular-modelling studies of the solution conformation of heparin." *Biochem. J.* **293**(Pt 3): 849-858.

Mulloy, B., M. J. Forster, et al. (1994). "The effect of variation of substitution on the solution conformation of heparin: a spectroscopic and molecular modelling study." *Carbohydr. Res.* **255**: 1-26.

Murphy, J. W., Y. Cho, et al. (2007). "Structural and Functional Basis of CXCL12 (Stromal Cell-derived Factor-1{alpha}) Binding to Heparin." *J. Biol. Chem.* **282**(13): 10018-10027.

Murphy, K. J., N. McLay, et al. (2008). "Structural studies of heparan sulfate hexasaccharides: new insights into iduronate conformational behavior." *J. Am. Chem. Soc.* **130**(37): 12435-12444.

Murphy, P. (1997). "Neutrophil receptors for interleukin-8 and related CXC chemokines." *Semin. Hematol.* **34**(4): 311-318.

Murphy, P. M. and H. L. Tiffany (1991). "Cloning of complementary DNA encoding a functional human interleukin-8 receptor." *Science* **253**(5025): 1280-1283.

Nader, H. B., S. F. Chavante, et al. (1999). "Heparan sulfates and heparins: similar compounds performing the same functions in vertebrates and invertebrates?" *Braz. J. Med. Biol. Res.* **32**(5): 529-538.

Naggi, A. (2005). Glycol-splitting as a device for modulating inhibition of growth factors and heparanase by heparin and heparin derivatives. *Chemistry and Biology of Heparin and*

Heparan Sulfate. G. G. Hari, J. L. Robert and A. H. Charles. Amsterdam, Elsevier Science**:** 461-481.

Naggi, A., B. Casu, et al. (2005). "Modulation of the heparanase-inhibiting activity of heparin through selective desulfation, graded N-acetylation, and glycol splitting." *J. Biol. Chem.* **280**(13): 12103-12113.

Nahmany, A., F. Strino, et al. (2005). "The use of a genetic algorithm search for molecular mechanics (MM3)-based conformational analysis of oligosaccharides." *Carbohydr. Res.* **340**(5): 1059-1064.

Nakajima, M., T. Irimura, et al. (1984). "Metastatic melanoma cell heparanase. Characterization of heparan sulfate degradation fragments produced by B16 melanoma endoglucuronidase." *J. Biol. Chem.* **259**(4): 2283-2290.

Nakajima, M., T. Irimura, et al. (1988). "Heparanases and tumor metastasis." *J. Cell. Biochem.* **36**(2): 157-167.

Nam, K. H., W. H. Lee, et al. (2010). "Structural characterization of the bifunctional glucanase-xylanase CelM2 reveals the metal effect and substrate-binding moiety." *Biochem. Biophys. Res. Commun.* **391**(4): 1726-1730.

Nap, R. J. and I. Szleifer (2008). "Structure and interactions of aggrecans: statistical thermodynamic approach." *Biophys. J.* **95**(10): 4570-4583.

Nasser, M. W., S. K. Raghuwanshi, et al. (2009). "Differential activation and regulation of CXCR1 and CXCR2 by CXCL8 monomer and dimer." *J. Immunol.* **183**(5): 3425-3432.

Nasser, N. J., A. Avivi, et al. (2007). "Cloning, expression, and characterization of an alternatively spliced variant of human heparanase." *Biochem. Biophys. Res. Commun.* **354**(1): 33-38.

Nasser, N. J., E. Nevo, et al. (2005). "Adaptive evolution of heparanase in hypoxia-tolerant Spalax: gene cloning and identification of a unique splice variant." *Proc. Natl. Acad. Sci. U.S.A.* **102**(42): 15161-15166.

Nelson, D. and M. Cox (2004). Carbohydrates and glycobiology. *Lehninger Principles of Biochemistry*. New York, W. H. Freeman**:** 238–271.

Nesmelova, I. V., Y. Sham, et al. (2005). "Platelet factor 4 and interleukin-8 CXC chemokine heterodimer formation modulates function at the quaternary structural level." *J. Biol. Chem.* **280**(6): 4948-4958.

Nesmelova, I. V., Y. Sham, et al. (2008). "CXC and CC chemokines form mixed heterodimers: Association free energies from molecular dynamics simulations and experimental coorelations." *J. Biol. Chem.* **283**(35): 24155-24166.

Neugroschl, J. and M. Sano (2010). "Current treatment and recent clinical research in Alzheimer's disease." *Mt. Sinai J. Med.* **77**(1): 3-16.

Newman, P. J. (1997). "The biology of PECAM-1." *J. Clin. Invest.* **99**(1): 3-8.

Nguyen, D. T. and D. A. Case (1985). "On finding stationary states on large-molecule potential energy surfaces." *J Phys Chem* **89**(19): 4020-4026.

Nicolaes, G. A., K. W. Sorensen, et al. (2004). "Altered inactivation pathway of factor Va by activated protein C in the presence of heparin." *Eur. J. Biochem.* **271**(13): 2724-2736.

Nielsen, J. E. and J. A. McCammon (2003). "Calculating pKa values in enzyme active sites." *Protein Sci.* **12**(9): 1894-1901.

Nielsen, J. E. and J. A. McCammon (2003). "On the evaluation and optimization of protein X-ray structures for pKa calculations." *Protein Sci.* **12**(2): 313-326.

Norgan, A. P., P. K. Coffman, et al. (2011). "Multilevel Parallelization of AutoDock 4.2." *J Cheminform* **3**(1): 12.

Norgauer, J., J. Krutmann, et al. (1994). "Actin polymerization, calcium-transients, and phospholipid metabolism in human neutrophils after stimulation with interleukin-8 and N-formyl peptide." *J. Invest. Dermatol.* **102**(3): 310-314.

O'Donovan, N., M. Galvin, et al. (1999). "Physical mapping of the CXC chemokine locus on human chromosome 4." *Cytogenet. Cell Genet.* **84**(1-2): 39-42.

Ohnishi, Y., T. Senda, et al. (2000). "Crystal Structure of Recombinant Native SDF-1α with Additional Mutagenesis Studies: An Attempt at a More Comprehensive Interpretation of Accumulated Structure-Activity Relationship Data." *J. Interferon Cytokine Res.* **20**(8): 691-700.

Okada, Y., S. Yamada, et al. (2002). "Structural recognition by recombinant human heparanase that plays critical roles in tumor metastasis. Hierarchical sulfate groups with different effects and the essential target disulfated trisaccharide sequence." *J. Biol. Chem.* **277**(45): 42488-42495.

Olson, S. T. and Y. J. Chuang (2002). "Heparin activates antithrombin anticoagulant function by generating new interaction sites (exosites) for blood clotting proteinases." *Trends Cardiovasc. Med.* **12**(8): 331-338.

Olsson, M., C. Søndergaard, et al. (2010). "PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions." *J Chem Theory Comput* **0**(0).

Oosta, G. M., W. T. Gardner, et al. (1981). "Multiple functional domains of the heparin molecule." *Proc. Natl. Acad. Sci. U.S.A.* **78**(2): 829-833.

Oppenheim, J., M. Feldmann, et al., Eds. (2001). *Cytokine reference: A compendium of cytokines and other mediators of host defense ligands*. San Diego, Academic Press.

Osborn, H. M. I., P. G. Evans, et al. (2004). "Carbohydrate-based therapeutics." *J. Pharm. Pharmacol.* **56**(6): 691-702.

Ototani, N., M. Kikuchi, et al. (1981). "Comparative studies on the structures of highly active and relatively inactive forms of whale heparin." *J. Biochem.* **90**(1): 241-246.

Ott, K. H. and B. Meyer (1996). "Parametrization of GROMOS force field for oligosaccharides and assessment of efficiency of molecular dynamics simulations." *J Comput Chem* **17**(8): 1068-1084.

Paes, G., L. K. Skov, et al. (2008). "The structure of the complex between a branched pentasaccharide and Thermobacillus xylanilyticus GH-51 arabinofuranosidase reveals xylan-binding determinants and induced fit." *Biochemistry* **47**(28): 7441-7451.

Paolini, J. F., D. Willard, et al. (1994). "The chemokines IL-8, monocyte chemoattractant protein-1, and I-309 are monomers at physiologically relevant concentrations " *J. Immunol.* **153**(6): 2704-2717.

Papy-Garcia, D., V. Barbier-Chassefiere, et al. (2005). "Nondegradative sulfation of polysaccharides. Synthesis and structure characterization of biologically active heparan sulfate mimetics." *Macromolecules* **38**(11): 4647-4654.

Papy-Garcia, D., I. Barbosa, et al. (2002). "Glycosaminoglycan mimetics (RGTA) modulate adult skeletal muscle satellite cell proliferation *in vitro*." *J. Biomed. Mater. Res.* **62**(1): 46-55.

Papy-Garcia, D., M. Christophe, et al. (2011). "Glycosaminoglycans, protein aggregation and neurodegeneration." *Curr. Protein Pept. Sci.* **12**(3): 258-268.

Parish, C. R. (2005). "Heparan sulfate and inflammation." *Nat. Immunol.* **6**(9): 861-862.

Parish, C. R. (2006). "The role of heparan sulphate in inflammation." *Nat. Rev. Immunol.* **6**(9): 633-643.

Parish, C. R., C. Freeman, et al. (1999). "Identification of sulfated oligosaccharide-based inhibitors of tumor growth and metastasis using novel in vitro assays for angiogenesis and heparanase activity." *Cancer Res* **59**(14): 3433-3441.

Parish, C. R., C. Freeman, et al. (2001). "Heparanase: a key enzyme involved in cell invasion." *Biochim. Biophys. Acta* **1471**(3): M99-M108.

Parrish, R. F. and W. R. Fair (1981). "Selective binding of zinc ions to heparin rather than to other glycosaminoglycans." *Biochem. J.* **193**(2): 407-410.

Pejler, G., A. Danielsson, et al. (1987). "Structure and antithrombin-binding properties of heparin isolated from the clams Anomalocardia brasiliana and Tivela mactroides." *J. Biol. Chem.* **262**(24): 11413-11421.

Pellegrini, L., D. F. Burke, et al. (2000). "Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin." *Nature* **407**(6807): 1029-1034.

Penkett, C. J., C. Redfield, et al. (1997). "NMR analysis of main-chain conformational preferences in an unfolded fibronectin-binding protein." *J. Mol. Biol.* **274**(2): 152-159.

Peterson, S. B. and J. Liu (2010). "Unraveling the specificity of heparanase utilizing synthetic substrates." *J. Biol. Chem.* **285**(19): 14504-14513.

Petitou, M., P. Duchaussoy, et al. (1998). "First synthetic carbohydrates with the full anticoagulant properties of heparin." *Angew. Chem. Int. Ed. Engl.* **37**(21): 3009-3014.

Petitou, M., V. Nancy-Portebois, et al. (2009). "From heparin to EP217609: The long way to a new pentasaccharide-based neutralisable anticoagulant with an unprecedented pharmacological profile." *Thromb. Haemost.* **102**(5): 804-810.

Petitou, M. and C. A. A. van Boeckel (2004). "A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next?" *Angew. Chem. Int. Ed. Engl.* **43**(24): 3118-3133.
Pettersen, E. F., T. D. Goddard, et al. (2004). "UCSF Chimera--a visualization system for exploratory research and analysis." *J Comput Chem* **25**(13): 1605-1612.
Piali, L., P. Hammel, et al. (1995). "CD31/PECAM-1 is a ligand for alpha v beta 3 integrin involved in adhesion of leukocytes to endothelium." *J. Cell Biol.* **130**(2): 451-460.
Pigache, A., P. Cieplak, et al. (2004). Automatic and highly reproducible RESP and ESP charge derivation: application to the development of programs RED and X RED. *227th ACS National Meeting*. Anaheim, CA, USA.
Pikas, D. S., J. P. Li, et al. (1998). "Substrate specificity of heparanases from human hepatoma and platelets." *J. Biol. Chem.* **273**(30): 18770-18777.
Pol-Fachin, L. and H. Verli (2008). "Depiction of the forces participating in the 2-O-sulfo-alpha-L-iduronic acid conformational preference in heparin sequences in aqueous solutions." *Carbohydr. Res.* **343**(9): 1435-1445.
Pol-Fachin, L. and H. Verli (2011). "Assessment of glycoproteins dynamics from computer simulations." *Mini Rev Org Chem* **8**(3): 229-238.
Prager, E., R. Sunder-Plassmann, et al. (1996). "Interaction of CD31 with a heterophilic counterreceptor involved in downregulation of human T cell responses." *J. Exp. Med.* **184**(1): 41-50.
Prandoni, P., D. Tormene, et al. (2008). "Idraparinux: review of its clinical efficacy and safety for prevention and treatment of thromboembolic disorders." *Expert Opin Investig Drugs* **17**(5): 773-777.
Pupko, T., R. E. Bell, et al. (2002). "Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues." *Bioinformatics* **18**(suppl_1): S71-77.
Rademacher, T. W., R. B. Parekh, et al. (1988). "Glycobiology." *Annu. Rev. Biochem.* **57**: 785-838.
Rafii, M. and P. Aisen (2009). "Recent developments in Alzheimer's disease therapeutics." *BMC Med* **7**(1): 7.
Ragazzi, M., D. R. Ferro, et al. (1990). "Conformation of the pentasaccharide corresponding to the binding site of heparin for antithrombin III." *Carbohydr. Res.* **195**(2): 169-185.
Ragazzi, M., D. R. Ferro, et al. (1986). "A force-field study of the conformational characteristics of the iduronate ring." *J Comput Chem* **7**(2): 105-112.
Ragazzi, M., D. R. Ferro, et al. (1993). "Conformation of the unsaturated uronic acid residues of glycosamtnoglycan disaccharides." *J Carbohydr Chem* **12**(4): 523 - 535.
Raghuraman, A., P. D. Mosier, et al. (2006). "Finding a needle in a haystack: development of a combinatorial virtual screening approach for identifying high specificity heparin/heparan sulfate sequence(s)." *J. Med. Chem.* **49**(12): 3553-3562.
Raghuraman, A., M. Riaz, et al. (2007). "Rapid and efficient microwave-assisted synthesis of highly sulfated organic scaffolds." *Tetrahedron Lett.* **48**(38): 6754-6758.
Raghuraman, A., V. Tiwari, et al. (2005). "Structural characterization of a serendipitously discovered bioactive macromolecule, lignin sulfate." *Biomacromolecules* **6**(5): 2822-2832.
Rajagopalan, L. and K. Rajarathnam (2004). "Ligand selectivity and affinity of chemokine receptor CXCR1: Role of N-terminal domain." *J. Biol. Chem.* **279**(29): 30000-30008.
Rajarathnam, K., I. Clark-Lewis, et al. (1996). "1H NMR evidence that Glu-38 interacts with the N-terminal functional domain in interleukin-8." *FEBS Lett* **399**(1-2): 43-46.
Rajarathnam, K., I. Clark-Lewis, et al. (1994). "1H NMR studies of interleukin 8 analogs: characterization of the domains essential for function." *Biochemistry* **33**(21): 6623-6630.
Rajarathnam, K., I. Clark-Lewis, et al. (1995). "1H NMR solution structure of an active monomeric interleukin-8." *Biochemistry* **34**(40): 12983-12990.
Rajarathnam, K., C. M. Kay, et al. (1997). "Characterization of quaternary structure of interleukin-8 and functional implications." *Methods Enzymol.* **287**: 89-105.
Rajarathnam, K., G. N. Prado, et al. (2006). "Probing receptor binding activity of interleukin-8 dimer using a disulfide trap." *Biochemistry* **45**(25): 7882-7888.
Rajarathnam, K., B. D. Sykes, et al. (1999). "Disulfide bridges in interleukin-8 probed using non-natural disulfide analogues: Dissociation of roles in structure from function." *Biochemistry* **38**(24): 7653-7658.

Rajarathnam, K., B. D. Sykes, et al. (1994). "Neutrophil activation by monomeric interleukin-8." *Science* **264**(5155): 90-92.
Rambaut, A. (2006). FigTree, a graphical viewer of phylogenetic trees and as a program for producing publication-ready figures (http://tree.bio.ed.ac.uk/software/figtree/).
Rao, V. S. R., P. K. Qasba, et al. (1998). *Conformation of carbohydrates*. Amsterdam, Harwood Academic Pub.
Rastelli, G., A. Del Rio, et al. (2010). "Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA." *J Comput Chem* **31**(4): 797-810.
Ravindran, A., P. R. B. Joseph, et al. (2009). "Structural basis for differential binding of the Interleukin-8 monomer and dimer to the CXCR1 N-Domain: Role of coupled interactions and dynamics." *Biochemistry* **48**(37): 8795-8805.
Reeves, E. P., M. Williamson, et al. (2010). "IL-8 dictates glycosaminoglycan binding and stability of IL-18 in cystic fibrosis." *J. Immunol.* **184**: 1642-1652.
Rek, A., H. Potzinger, et al. (2010). "Turning the glycan switch to change CXCL8 into a potent anti-inflammatory biological drug." *Proc. Natl. Acad. Sci. U.S.A.* **submitted for publication**.
Remko, M. and M. Hricovíni (2007). "Theoretical study of structure and properties of hexuronic acid and D-glucosamine structural units of glycosaminoglycans " *Struct Chem* **18**(5): 537-547.
Remko, M., M. Swart, et al. (2007). "Conformational behavior of basic monomeric building units of glycosaminoglycans: isolated systems and solvent effect." *J Phys Chem B* **111**(9): 2313-2321.
Revill, P., N. Serradell, et al. (2006). "Eprodisate Sodium." *Drugs Future* **31**(7): 576-578.
Ricard-Blum, S., O. Feraud, et al. (2004). "Characterization of endostatin binding to heparin and heparan sulfate by surface plasmon resonance and molecular modeling: role of divalent cations." *J. Biol. Chem.* **279**(4): 2927-2936.
Rider, C. C. (1997). "The potential for heparin and its derivatives in the therapy and prevention of HIV-1 infection." *Glycoconj. J.* **14**(5): 639-642.
Rider, C. C., D. R. Coombe, et al. (1994). "Anti-HIV-1 activity of chemically modified heparins: correlation between binding to the V3 loop of gp120 and inhibition of cellular HIV-1 infection in vitro." *Biochemistry* **33**(22): 6974-6980.
Ronald, E. H., R. F. Jonathan, et al. (1998). "Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins." *BioEssays* **20**(2): 156-167.
Rosen, J., L. Miguet, et al. (2009). "Shape: automatic conformation prediction of carbohydrates using a genetic algorithm." *J Cheminform* **1**(1): 16.
Rosenberg, R. D., N. W. Shworak, et al. (1997). "Heparan sulfate proteoglycans of the cardiovascular system. Specific structures emerge but how is synthesis regulated?" *J. Clin. Invest.* **99**(9): 2062-2070.
Rosenthal, M. A., D. Rischin, et al. (2002). "Treatment with the novel anti-angiogenic agent PI-88 is associated with immune-mediated thrombocytopenia." *Ann. Oncol.* **13**(5): 770-776.
Rost, B. and C. Sander (1993). "Prediction of protein secondary structure at better than 70% accuracy." *J. Mol. Biol.* **232**(2): 584-599.
Rost, B. and C. Sander (1994). "Combining evolutionary information and neural networks to predict protein secondary structure." *Proteins* **19**(1): 55-72.
Rostand, K. S. and J. D. Esko (1997). "Microbial adherence to and invasion through proteoglycans." *Infect. Immun.* **65**(1): 1-8.
Rouet, V., Y. Hamma-Kourbali, et al. (2005). "A synthetic glycosaminoglycan mimetic minds vascular endothelial growth factor and modulates angiogenesis." *J. Biol. Chem.* **280**(38): 32792-32800.
Rouet, V., A. Meddahi-Pellé, et al. (2006). "Heparin-like synthetic polymers, named RGTAs, mimic biological effects of heparin *in vitro*." *J Biomed Mater Res A* **78A**(4): 792-797.
Roy, A., A. Kucukural, et al. (2010). "I-TASSER: a unified platform for automated protein structure and function prediction." *Nat Protoc* **5**(4): 725-738.
Rusnati, M., D. Coltrini, et al. (1994). "Distinct role of 2-O-, N-, and 6-O-sulfate groups of heparin in the formation of the ternary complex with basic fibroblast growth factor and soluble FGF receptor-1." *Biochem. Biophys. Res. Commun.* **203**(1): 450-458.
Rusnati, M., D. Coltrini, et al. (1997). "Interaction of HIV-1 Tat protein with heparin. Role of the backbone structure, sulfation, and size." *J. Biol. Chem.* **272**(17): 11313-11320.

Ryckaert, J.-P., G. Ciccotti, et al. (1977). "Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes." *J Comput Phys* **23**(3): 327-341.

Sadir, R., F. Baleux, et al. (2001). "Characterization of the stromal cell-derived factor-1alpha-heparin complex." *J. Biol. Chem.* **276**(11): 8288-8296.

Sali, A. and T. L. Blundell (1993). "Comparative protein modelling by satisfaction of spatial restraints." *J. Mol. Biol.* **234**(3): 779-815.

Samama, M. and G. Gerotziafas (2010). "Newer anticoagulants in 2009." *J. Thromb. Thrombolysis* **29**(1): 92-104.

Samsonov, S. A., J. Teyra, et al. (2011). "Docking glycosaminoglycans to proteins: analysis of solvent inclusion." *J. Comput. Aided Mol. Des.* **25**(5): 477-489.

Sanderson, R. D. (2001). "Heparan sulfate proteoglycans in invasion and metastasis." *Semin. Cell Dev. Biol.* **12**(2): 89-98.

Sanner, M. F., A. J. Olson, et al. (1996). "Reduced surface: an efficient way to compute molecular surfaces." *Biopolymers* **38**(3): 305-320.

Santa-Maria, I., F. Hernandez, et al. (2007). "Tramiprosate, a drug of potential interest for the treatment of Alzheimer's disease, promotes an abnormal aggregation of tau." *Mol Neurodegener* **2**(1): 17.

Sapay, N., E. Cabannes, et al. (2011). "Molecular modeling of the interaction between heparan sulfate and cellular growth factors: bringing pieces together." *Glycobiology*.

Sapay, N., É. Cabannes, et al. (2011). "Molecular model of human heparanase with proposed binding mode of a heparan sulfate oligosaccharide and catalytic amino acids." *Biopolymers*: n/a-n/a.

Sasisekharan, R., R. Raman, et al. (2006). "Glycomics approach to structure-function relationships of glycosaminoglycans." *Annu Rev Biomed Eng* **8**: 181-231.

Sasisekharan, R., Z. Shriver, et al. (2002). "Roles of heparan-sulphate glycosaminoglycans in cancer." *Nat. Rev. Cancer* **2**(7): 521-528.

Sasisekharan, R. and G. Venkataraman (2000). "Heparin and heparan sulfate: biosynthesis, structure and function." *Curr Opin Chem Biol* **4**(6): 626-631.

Sato, M., K. Amemiya, et al. (2008). "Subcellular localization of human heparanase and its alternative splice variant in COS-7 cells." *Cell Biochem. Funct.* **26**(6): 676-683.

Sattelle, B. M. and A. Almond (2010). "Less is more when simulating unsulfated glycosaminoglycan 3D-structure: comparison of GLYCAM06/TIP3P, PM3-CARB1/TIP3P, and SCC-DFTB-D/TIP3P predictions with experiment." *J Comput Chem* **31**(16): 2932-2947.

Sattelle, B. M., S. U. Hansen, et al. (2010). "Free energy landscapes of iduronic acid and related monosaccharides." *J. Am. Chem. Soc.* **132**(38): 13132-13134.

Saumier, D., P. Aisen, et al. (2009). "Lessons learned in the use of volumetric MRI in therapeutic trials in Alzheimer's disease: The Alzhemedtm (Tramiprosate) experience." *J Nutr Health Aging* **13**(4): 370-372.

Savi, P., J. P. Herault, et al. (2008). "Reversible biotinylated oligosaccharides: a new approach for a better management of anticoagulant therapy." *J. Thromb. Haemost.* **6**(10): 1697-1706.

Sawitzky, D. (1996). "Protein-glycosaminoglycan interactions: infectiological aspects." *Med. Microbiol. Immunol.* **184**(4): 155-161.

Schedin-Weiss, S., U. R. Desai, et al. (2002). "Importance of lysine 125 for heparin binding and activation of antithrombin." *Biochemistry* **41**(15): 4779-4788.

Schlessinger, J., A. N. Plotnikov, et al. (2000). "Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization." *Mol. Cell* **6**(3): 743-750.

Schlitter, J. (1993). "Estimation of absolute and relative entropies of macromolecules using the covariance matrix." *Chem Phys Lett* **215**(6): 617-621.

Schmid, J. and C. Weissmann (1987). "Induction of mRNA for a serine protease and a beta-thromboglobulin-like protein in mitogen-stimulated human leukocytes." *J. Immunol.* **139**(1): 250-256.

Schmidt, A., A. Schlacher, et al. (1998). "Structure of the xylanase from Penicillium simplicissimum." *Protein Sci.* **7**(10): 2081-2088.

Schonberger, O., L. Horonchik, et al. (2003). "Novel heparan mimetics potently inhibit the scrapie prion protein and its endocytosis." *Biochem. Biophys. Res. Commun.* **312**(2): 473-479.

Schroder, J. M. and E. Christophers (1989). "Secretion of novel and homologous neutrophil-activating peptides by LPS-stimulated human endothelial cells." *J. Immunol.* **142**(1): 244-251.
Schroder, J. M., U. Mrowietz, et al. (1988). "Identification of different charged species of a human monocyte derived neutrophil activating peptide (MONAP)." *Biochem. Biophys. Res. Commun.* **152**(1): 277-284.
Schroder, J. M., U. Mrowietz, et al. (1987). "Purification and partial biochemical characterization of a human monocyte-derived, neutrophil-activating peptide that lacks interleukin 1 activity." *J. Immunol.* **139**(10): 3474-3483.
Schug, A., T. Herges, et al. (2003). "Reproducible protein folding with the stochastic tunneling method." *Phys. Rev. Lett.* **91**(15): 158102.
Schuksz, M., M. M. Fuster, et al. (2008). "Surfen, a small molecule antagonist of heparan sulfate." *Proc. Natl. Acad. Sci. U.S.A.* **105**(35): 13075-13080.
Schuttelkopf, A. W. and D. M. van Aalten (2004). "PRODRG: a tool for high-throughput crystallography of protein-ligand complexes." *Acta Crystallogr. D Biol. Crystallogr.* **60**(Pt 8): 1355-1363.
Seeberger, P. H. and D. B. Werz (2007). "Synthesis and medical applications of oligosaccharides." *Nature* **446**(7139): 1046-1051.
Sega, M., E. Autieri, et al. (2009). "On the calculation of puckering free energy surfaces." *J Chem Phys* **130**(22): 225102.
Sekido, N., N. Mukaida, et al. (1993). "Prevention of lung reperfusion injury in rabbits by a monoclonal antibody against interleukin-8." *Nature* **365**(6447): 654-657.
Seo, M., N. Castillo, et al. (2007). "Approach for the simulation and modeling of flexible rings: application to the α-d-arabinofuranoside ring, a key constituent of polysaccharides from mycobacterium tuberculosis." *J Chem Theory Comput* **4**(1): 184-191.
Sepulcre, M. P., E. Sarropoulou, et al. (2007). "Vibrio anguillarum evades the immune response of the bony fish sea bass (Dicentrarchus labrax L.) through the inhibition of leukocyte respiratory burst and down-regulation of apoptotic caspases." *Mol. Immunol.* **44**(15): 3751-3757.
Serina, G., J. F. Mirjoleta, et al. (2010). *Antitumor activity of EP80061, a small-glyco drug in preclinical studies*. AACR meeting, Washington DC.
Shao, C., F. Zhang, et al. (2006). "Crystallographic analysis of calcium-dependent heparin binding to annexin A2." *J. Biol. Chem.* **281**(42): 31689-31695.
Sharon, N. (1986). "IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). Nomenclature of glycoproteins, glycopeptides and peptidoglycans. Recommendations 1985." *Eur. J. Biochem.* **159**(1): 1-6.
Sharp, K. A., A. Nicholls, et al. (1998). DelPhi. New York, Dept. of Biochemistry and Molecular Biophysics, Columbia University.
Shaw, J. P., Z. Johnson, et al. (2004). "The X-Ray Structure of RANTES: Heparin-Derived Disaccharides Allows the Rational Design of Chemokine Inhibitors." *Structure* **12**(11): 2081-2093.
Shen, M. Y. and A. Sali (2006). "Statistical potential for assessment and prediction of protein structures." *Protein Sci.* **15**(11): 2507-2524.
Shi, J., T. L. Blundell, et al. (2001). "FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties." *J. Mol. Biol.* **310**(1): 243-257.
Shieh, M. T., D. WuDunn, et al. (1992). "Cell surface receptors for herpes simplex virus are heparan sulfate proteoglycans." *J. Cell Biol.* **116**(5): 1273-1281.
Shields, D. C. (2000). "Gene conversion among chemokine receptors." *Gene* **246**(1-2): 239-245.
Shionyu-Mitsuyama, C., T. Shirai, et al. (2003). "An empirical approach for structure-based prediction of carbohydrate-binding sites on proteins." *Protein Eng.* **16**(7): 467-478.
Showalter, S. A. and R. Brüschweiler (2007). "Quantitative molecular ensemble interpretation of NMR dipolar couplings without restraints." *J. Am. Chem. Soc* **129**(14): 4158-4159.
Showalter, S. A., E. Johnson, et al. (2007). "Toward quantitative interpretation of methyl side-chain dynamics from NMR by molecular dynamics simulations." *J. Am. Chem. Soc* **129**(46): 14146-14147.
Shriver, Z., S. Raguram, et al. (2004). "Glycomics: a pathway to a class of new and improved therapeutics." *Nat Rev Drug Discov* **3**(10): 863-873.

Shukla, D. and P. G. Spear (2001). "Herpesviruses and heparan sulfate: an intimate relationship in aid of viral entry." *J. Clin. Invest.* **108**(4): 503-510.
Sibille, N., A. Sillen, et al. (2006). "Structural impact of heparin binding to full-length Tau as studied by NMR spectroscopy." *Biochemistry* **45**(41): 12560-12572.
Sick, C., K. Schneider, et al. (2000). "Novel chicken CXC and CC chemokines." *Cytokine* **12**(3): 181-186.
Silvian, L., P. Jin, et al. (2006). "Artemin crystal structure reveals insights into heparan sulfate binding." *Biochemistry* **45**(22): 6801-6812.
Simizu, S., K. Ishida, et al. (2004). "Heparanase as a molecular target of cancer chemotherapy." *Cancer Sci.* **95**(7): 553-558.
Simizu, S., K. Ishida, et al. (2004). "Secretion of heparanase protein is regulated by glycosylation in human tumor cell lines." *J. Biol. Chem.* **279**(4): 2697-2703.
Simizu, S., K. Ishida, et al. (2003). "Expression of heparanase in human tumor cell lines and human head and neck tumors." *Cancer Lett.* **193**(1): 83-89.
Simizu, S., T. Suzuki, et al. (2007). "Involvement of disulfide bond formation in the activation of heparanase." *Cancer Res* **67**(16): 7841-7849.
Singh, N. and A. Warshel (2010). "Absolute binding free energy calculations: on the accuracy of computational scoring of protein-ligand interactions." *Proteins* **78**(7): 1705-1723.
Sitkoff, D., K. A. Sharp, et al. (1994). "Accurate calculation of hydration free energies using macroscopic solvent models." *J Phys Chem* **98**(7): 1978-1988.
Skelton, N. J., C. Quan, et al. (1999). "Structure of a CXC chemokine-receptor fragment in complex with interleukin-8." *Structure* **7**(2): 157-168.
Skolnick, J. and M. Brylinski (2009). "FINDSITE: a combined evolution/structure-based approach to protein function prediction." *Brief. Bioinformatics* **10**(4): 378-391.
Snow, A. D., J. Willmer, et al. (1987). "Sulfated glycosaminoglycans: a common constituent of all amyloids?" *Lab. Invest.* **56**(1): 120-123.
Sobel, M., D. F. Soler, et al. (1992). "Localization and characterization of a heparin binding domain peptide of human von Willebrand factor." *J. Biol. Chem.* **267**(13): 8857-8862.
Sobieraj-Teague, M., M. O'Donnell, et al. (2009). "New anticoagulants for atrial fibrillation." *Semin. Thromb. Hemost.* **35**(5): 515-524.
Soding, J., A. Biegert, et al. (2005). "The HHpred interactive server for protein homology detection and structure prediction." *Nucleic Acids Res.* **33**(Web Server issue): W244-W248.
Spear, P. G. (2004). "Herpes simplex virus: receptors and ligands for cell entry." *Cell. Microbiol.* **6**(5): 401-410.
Spear, P. G., R. J. Eisenberg, et al. (2000). "Three classes of cell surface receptors for alphaherpesvirus entry." *Virology* **275**(1): 1-8.
Spear, P. G., M. T. Shieh, et al. (1992). "Heparan sulfate glycosaminoglycans as primary cell surface receptors for herpes simplex virus." *Adv. Exp. Med. Biol.* **313**: 341-353.
Spillmann, D. and U. Lindahl (1994). "Glycosaminoglycan-protein interactions: a question of specificity." *Curr. Opin. Struct. Biol.* **4**(5): 677-682.
Spillmann, D., D. Witt, et al. (1998). "Defining the interleukin-8-binding domain of heparan sulfate." *J. Biol. Chem.* **273**(25): 15487-15493.
Spiwok, V. and A. D. French (2011). "Modelling the effect of solvents on carbohydrates." *Mini Rev Org Chem* **8**(3): 249-255.
Spiwok, V., B. Kralova, et al. (2010). "Modelling of beta-D-glucopyranose ring distortion in different force fields: a metadynamics study." *Carbohydr. Res.* **345**(4): 530-537.
Spiwok, V. and I. Tvaroska (2009). "Metadynamics modelling of the solvent effect on primary hydroxyl rotamer equilibria in hexopyranosides." *Carbohydr. Res.* **344**(12): 1575-1581.
Srinivasan, J., T. E. Cheatham, et al. (1998). "Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate - DNA helices." *J. Am. Chem. Soc.* **120**(37): 9401-9409.
Srinivasan, S. R., B. Radhakrishnamurthy, et al. (1975). "Studies on the interaction of heparin with serum lipoproteins in the presence of Ca2+, Mg2+, and Mn2+." *Arch. Biochem. Biophys.* **170**(1): 334-340.
Stahl, M. and M. Rarey (2001). "Detailed analysis of scoring functions for virtual screening." *J. Med. Chem.* **44**(7): 1035-1042.
Stern, R. (2008). "Association between cancer and "acid mucopolysaccharides": an old concept comes of age, finally." *Semin. Cancer Biol.* **18**(4): 238-243.

Steven, R. L., A. H. Caroline, et al. (1997). "IL-8 single-chain homodimers and heterodimers: Interactions with the chemokine receptors CXCR1, CXCR2, and DARC." *Protein Sci.* **6**(3): 609-617.

Still, W. C., A. Tempczyk, et al. (1990). "Semianalytical treatment of solvation for molecular mechanics and dynamics." *J. Am. Chem. Soc.* **112**(16): 6127-6129.

Strieter, R. M., P. J. Polverini, et al. (1995). "Role of C-X-C chemokines as regulators of angiogenesis in lung-cancer." *J. Leukoc. Biol.* **57**(5): 752-762.

Stringer, S. E., M. J. Forster, et al. (2002). "Characterization of the binding site on heparan sulfate for macrophage inflammatory protein-1alpha." *Blood* **100**(5): 1543-1550.

Stringer, S. E. and J. T. Gallagher (1997). "Specific binding of the chemokine platelet factor 4 to heparan sulfate." *J. Biol. Chem.* **272**(33): 20508-20514.

Strino, F., J. H. Lii, et al. (2010). "Selenoglycosides in silico: ab initio-derived reparameterization of MM4, conformational analysis using histo-blood group ABH antigens and lectin docking as indication for potential of bioactivity." *J. Comput. Aided Mol. Des.* **24**(12): 1009-1021.

Suhre, K. and Y.-H. Sanejouand (2004). "ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement." *Nucleic Acids Res.* **32**(Web Server issue): W610-W614.

Suzuki, H., G. N. Prado, et al. (1994). "The N terminus of interleukin-8 (IL-8) receptor confers high affinity binding to human IL-8." *J. Biol. Chem.* **269**(28): 18263-18266.

Suzuki, K., H. Miyasaka, et al. (1989). "Purification and partial primary sequence of a chemotactic protein for polymorphonuclear leukocytes derived from human lung giant cell carcinoma LU65C cells." *J. Exp. Med.* **169**(6): 1895-1901.

Swanson, J. M., R. H. Henchman, et al. (2004). "Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy." *Biophys. J.* **86**(1 Pt 1): 67-74.

Taha, H. A., N. Castillo, et al. (2009). "Conformational studies of methyl β-d-arabinofuranoside using the AMBER/GLYCAM approach." *J Chem Theory Comput* **5**(2): 430-438.

Takaoka, T., K. Mori, et al. (2007). "Prediction of the structure of complexes comprised of proteins and glycosaminoglycans using docking simulation and cluster analysis." *J Chem Theory Comput* **3**(6): 2347-2356.

Tan, K., M. Duquette, et al. (2006). "The structures of the thrombospondin-1 N-terminal domain and its complex with a synthetic pentameric heparin." *Structure* **14**(1): 33-42.

Tang, Y. T. and G. R. Marshall (2011). "PHOENIX: a scoring function for affinity prediction derived using high-resolution crystal structures and calorimetry measurements." *J Chem Inf Model* **51**(2): 214-228.

Tardieu, M., C. Gamby, et al. (1992). "Derivatized dextrans mimic heparin as stabilizers, potentiators, and protectors of acidic or basic FGF." *J. Cell. Physiol.* **150**(1): 194-203.

Taroni, C., S. Jones, et al. (2000). "Analysis and prediction of carbohydrate binding sites." *Protein Eng.* **13**(2): 89-98.

Taylor, E. J., N. L. Smith, et al. (2006). "Structural insight into the ligand specificity of a thermostable family 51 arabinofuranosidase, Araf51, from Clostridium thermocellum." *Biochem. J.* **395**(1): 31-37.

Teeter, M. M. and D. A. Case (1990). "Harmonic and quasiharmonic descriptions of crambin." *J Phys Chem* **94**(21): 8091-8097.

The AMADEUS Investigators, B. MG, et al. (2008). "Comparison of idraparinux with vitamin K antagonists for prevention of thromboembolism in patients with atrial fibrillation: a randomised, open-label, non-inferiority trial." *The Lancet* **371**(9609): 315-321.

The PERSIST Investigators (2002). "A novel long-acting synthetic factor Xa inhibitor (idraparinux sodium) to replace warfarin for secondary prevention in deep vein thrombosis. A phase II evaluation." *Blood* **100**(11): 301.

Thomas, A. H. (1996). "Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94." *J Comput Chem* **17**(5-6): 490-519.

Thomas, A. H. (1996). "Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions." *J Comput Chem* **17**(5-6): 520-552.

Thomas, A. H. (1996). "Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94." *J Comput Chem* **17**(5-6): 553-586.

Thomas, A. H. (1996). "Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules." *J Comput Chem* **17**(5-6): 616-641.
Thomas, A. H. and B. N. Robert (1996). "Merck molecular force field. IV. conformational energies and geometries for MMFF94." *J Comput Chem* **17**(5-6): 587-615.
Thompson, L. D., M. W. Pantoliano, et al. (1994). "Energetic characterization of the basic fibroblast growth factor-heparin interaction: identification of the heparin binding domain." *Biochemistry* **33**(13): 3831-3840.
Thunberg, L., G. Backstrom, et al. (1982). "Enzymatic depolymerization of heparin-related polysaccharides. Substrate specificities of mouse mastocytoma and human platelet endo-beta-D-glucuronidases." *J. Biol. Chem.* **257**(17): 10278-10282.
Tidor, B. and M. Karplus (1994). "The contribution of vibrational entropy to molecular association. The dimerization of insulin." *J Mol Biol* **238**(3): 405-414.
Timar, J., K. Lapis, et al. (2002). "Proteoglycans and tumor progression: Janus-faced molecules with contradictory functions in cancer." *Semin. Cancer Biol.* **12**(3): 173-186.
Tom, D., Y. Darrin, et al. (1993). "Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems." *J Chem Phys* **98**(12): 10089-10092.
Tong, M., B. Tuk, et al. (2009). "Stimulated neovascularization, inflammation resolution and collagen maturation in healing rat cutaneous wounds by a heparan sulfate glycosaminoglycan mimetic, OTR4120." *Wound Repair Regen* **17**(6): 840-852.
Toole, B. P. (1997). "Hyaluronan in morphogenesis." *J. Intern. Med.* **242**(1): 35-40.
Toole, B. P., T. N. Wight, et al. (2002). "Hyaluronan-cell interactions in cancer and vascular disease." *J. Biol. Chem.* **277**(7): 4593-4596.
Toyoshima, M. and M. Nakajima (1999). "Human heparanase. Purification, characterization, cloning, and expression." *J. Biol. Chem.* **274**(34): 24153-24160.
Tremblay, P., P. Aisen, et al. (2005). "Functional GAG mimetics as an approach for the treatment of amyloid diseases." *Alzheimers Dement* **1**(1, Supplement 1): S2-S2.
Tsui, V. and D. A. Case (2000). "Theory and applications of the generalized Born solvation model in macromolecular simulations." *Biopolymers* **56**(4): 275-291.
Turley, E. A., P. W. Noble, et al. (2002). "Signaling properties of hyaluronan receptors." *J. Biol. Chem.* **277**(7): 4589-4592.
Turpie, A. G. G. (2004). "Fondaparinux: a Factor Xa inhibitor for antithrombotic therapy." *Expert Opin Pharmacother* **5**(6): 1373-1384.
Tvaroska, I., M. Hricovíni, et al. (1989). "An attempt to derive a new Karplus-type equation of vicinal proton-carbon coupling constants for C---O---C---H segments of bonded atoms." *Carbohydr. Res.* **189**: 359-362.
Urbanc, B., L. Cruz, et al. (2004). "Molecular dynamics simulation of amyloid beta dimer formation." *Biophys. J.* **87**(4): 2310-2321.
van Boeckel, C. A. A. and M. Petitou (1993). "The unique antithrombin III binding domain of heparin: A lead to new synthetic antithrombotics." *Angew. Chem. Int. Ed. Engl.* **32**(12): 1671-1690.
van Boeckel, C. A. A., S. F. van Aelst, et al. (1987). "Conformational analysis of synthetic heparin-like oligosaccharides containing α-L-idopyranosyluronic acid." *Recl. Trav. Chim. Pays Bas* **106**(1): 19-29.
Van Damme, J., J. Van Beeumen, et al. (1989). "Purification of granulocyte chemotactic peptide/interleukin-8 reveals N-terminal sequence heterogeneity similar to that of beta-thromboglobulin." *Eur. J. Biochem.* **181**(2): 337-344.
Van der Spoel, D., E. Lindahl, et al. (2006). Gromacs User Manual Nijenborgh, Department of Biophysical Chemistry, University of Groningen. . **4:** 9747.
Van Gunsteren, W., S. Billeter, et al. (1996). Biomolecular simulation: the GROMOS96 manual and user guide, Zürich, Switzerland: Hochschulverlag AG an der ETH Zürich.
van Horssen, J., P. Wesseling, et al. (2003). "Heparan sulphate proteoglycans in Alzheimer's disease and amyloid-related disorders." *Lancet Neurol* **2**(8): 482-492.
Vanquelef, E., S. Simon, et al. (2011). "R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments." *Nucleic Acids Res.* **39**(Web Server issue): W511-517.
Varki, A., R. Cummings, et al., Eds. (2009). *Essenstials of glycobiology*. Cold Spring Harbor (NY), Cold Spring Harbor Laboratory Press.

Varki, A., R. C. J. Esko, et al., Eds. (1999). *Essentials of glycobiology*. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.

Varrot, A., M. Schulein, et al. (2001). "Atomic resolution structure of endoglucanase Cel5A in complex with methyl 4,4II,4III,4IV-tetrathio-alpha-cellopentoside highlights the alternative binding modes targeted by substrate mimics." *Acta Crystallogr. D Biol. Crystallogr.* **57**(Pt 11): 1739-1742.

Veldkamp, C. T., F. C. Peterson, et al. (2005). "The monomer-dimer equilibrium of stromal cell-derived factor-1 (CXCL 12) is altered by pH, phosphate, sulfate, and heparin." *Protein Sci.* **14**(4): 1071-1081.

Verli, H. and J. A. Guimaraes (2004). "Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution." *Carbohydr. Res.* **339**(2): 281-290.

Verli, H. and J. A. Guimaraes (2005). "Insights into the induced fit mechanism in antithrombin-heparin interaction using molecular dynamics simulations." *J. Mol. Graph. Model.* **24**(3): 203-212.

Veyrat-Follet, C., N. Vivier, et al. (2009). "The pharmacokinetics of idraparinux, a long-acting indirect factor Xa inhibitor: population pharmacokinetic analysis from Phase III clinical trials." *J. Thromb. Haemost.* **7**(4): 559-565.

Viktor, H., A. Robert, et al. (2006). "Comparison of multiple Amber force fields and development of improved protein backbone parameters." *Proteins* **65**(3): 712-725.

Viskov, C., M. Just, et al. (2009). "Description of the chemical and pharmacological characteristics of a new hemisynthetic ultra-low-molecular-weight heparin, AVE5026." *J. Thromb. Haemost.* **7**(7): 1143-1151.

Vlodavsky, I., M. Elkin, et al. (2008). "Heparanase: one molecule with multiple functions in cancer progression." *Connect. Tissue Res.* **49**(3): 207-210.

Vlodavsky, I. and Y. Friedmann (2001). "Molecular properties and involvement of heparanase in cancer metastasis and angiogenesis." *J. Clin. Invest.* **108**(3): 341-347.

Vlodavsky, I., Y. Friedmann, et al. (1999). "Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis." *Nat. Med.* **5**(7): 793-802.

Vlodavsky, I., N. Ilan, et al. (2007). "Heparanase, heparin and the coagulation system in cancer progression." *Thromb. Res.* **120 Suppl 2**: S112-S120.

Vlodavsky, I., N. Ilan, et al. (2007). "Heparanase: Structure, biological functions, and inhibition by heparin-derived mimetics of heparan sulfate." *Curr Pharm Des* **13**(20): 2057-2073.

Vlodavsky, I., M. Mohsen, et al. (1994). "Inhibition of tumor metastasis by heparanase inhibiting species of heparin." *Invasion Metastasis* **14**(1-6): 290-302.

Volovyk, Z., D. Monroe, et al. (2009). "A rationally designed heparin, M118, has anticoagulant activity similar to unfractionated heparin and different from Lovenox in a cell-based model of thrombin generation." *J. Thromb. Thrombolysis* **28**(2): 132-139.

Volpi, N. (2006). "Therapeutic applications of glycosaminoglycans." *Curr. Med. Chem.* **13**(15): 1799-1810.

Vyas, A. A., J. J. Pan, et al. (1997). "Analysis of binding of cobra cardiotoxins to heparin reveals a new beta-sheet heparin-binding structural motif." *J. Biol. Chem.* **272**(15): 9661-9670.

Wadstrom, T. and A. S. A. Ljungh (1999). "Glycosaminoglycan-binding microbial proteins in tissue adhesion and invasion: key events in microbial pathogenicity." *J. Med. Microbiol.* **48**(3): 223-233.

Wall, D., S. Douglas, et al. (2001). "Characterisation of the anticoagulant properties of a range of structurally diverse sulfated oligosaccharides." *Thromb. Res.* **103**(4): 325-335.

Walter, M. R. (2004). "Structural analysis of IL-10 and Type I interferon family members and their complexes with receptor." *Adv. Protein Chem.* **68**: 171-223.

Walz, A., P. Peveri, et al. (1987). "Purification and amino acid sequencing of NAF, a novel neutrophil-activating factor produced by monocytes." *Biochem. Biophys. Res. Commun.* **149**(2): 755-761.

Wang, J., R. Wolf, et al. (2004). "Development and testing of a general amber force field." *J Comput Chem* **25**(9): 1157-1174.

Warda, M., E. M. Gouda, et al. (2003). "Isolation and characterization of raw heparin from dromedary intestine: evaluation of a new source of pharmaceutical heparin." *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* **136**(4): 357-365.

Warda, M., W. Mao, et al. (2003). "Turkey intestine as a commercial source of heparin? Comparative structural studies of intestinal avian and mammalian glycosaminoglycans." *Comp. Biochem. Physiol. B, Biochem. Mol. Biol.* **134**(1): 189-197.

Wardrop, D. and D. Keeling (2008). "The story of the discovery of heparin and warfarin." *Br. J. Haematol.* **141**(6): 757-763.

Warner, C. D., J. A. Hoy, et al. (2010). "Tertiary structure and characterization of a glycoside hydrolase family 44 endoglucanase from Clostridium acetobutylicum." *Appl. Environ. Microbiol.* **76**(1): 338-346.

Waszkowycz, B., D. E. Clark, et al. (2011). "Outstanding challenges in protein–ligand docking and structure-based virtual screening." *Wiley Interdisciplinary Reviews: Computational Molecular Science* 1(2): 229-259.

Webb, L. M., M. U. Ehrengruber, et al. (1993). "Binding to heparan sulfate or heparin enhances neutrophil responses to interleukin 8." *Proc. Natl. Acad. Sci. U.S.A.* **90**(15): 7158-7162.

Westerduin, P., C. A. A. van Boeckel, et al. (1994). "Feasible synthesis and biological properties of six 'non-glycosamino' glycan analogues of the antithrombin III binding heparin pentasaccharide." *Bioorg. Med. Chem.* **2**(11): 1267-1280.

Wettreich, A., A. Sebollela, et al. (1999). "Acidic pH modulates the interaction between human granulocyte-macrophage colony-stimulating factor and glycosaminoglycans." *J. Biol. Chem.* **274**(44): 31468-31475.

Whitfield, D. M., J. Choay, et al. (1992). "Heavy metal binding to heparin disaccharides. I. Iduronic acid is the main binding site." *Biopolymers* **32**(6): 585-596.

Whitfield, D. M. and B. Sarkar (1992). "Heavy metal binding to heparin disaccharides. II. First evidence for zinc chelation." *Biopolymers* **32**(6): 597-619.

Wickstrom, L., A. Okur, et al. (2009). "Evaluating the performance of the ff99SB force field based on NMR scalar coupling data." *Biophys. J.* **97**(3): 853-856.

Wickstrom, L., A. Okur, et al. (2006). "The unfolded state of the villin headpiece helical subdomain: Computational studies of the role of locally stabilized structure." *J. Mol. Biol.* **360**(5): 1094-1107.

Wilken, J., D. Hoover, et al. (1999). "Total chemical synthesis and high-resolution crystal structure of the potent anti-HIV protein AOP-RANTES." *Chem. Biol.* **6**(1): 43-51.

Williams, G., N. Borkakoti, et al. (1996). "Mutagenesis studies of interleukin-8. Identification of a second epitope involved in receptor binding." *J. Biol. Chem.* **271**(16): 9579-9586.

Withers, S. and S. Williams. (2010). "Glycoside hydrolases." Retrieved December 15th, 2010, from http://www.cazypedia.org/.

Wolfe, S. (1972). "Gauche effect. Stereochemical consequences of adjacent electron pairs and polar bonds." *Acc. Chem. Res.* **5**(3): 102-111.

Wolfgang, K. and S. Christian (1983). "Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features." *Biopolymers* **22**(12): 2577-2637.

Woods, A., J. B. McCarthy, et al. (1993). "A synthetic peptide from the COOH-terminal heparin-binding domain of fibronectin promotes focal adhesion formation." *Mol. Biol. Cell* **4**(6): 605-613.

Woods, R. J. (2005-2011). "GLYCAM Web (http://glycam.ccrc.uga.edu/ccrc/pages/3dspt.html)." from http://www.glycam.com.

Woods, R. J. and R. Chappelle (2000). "Restrained electrostatic potential atomic partial charges for condensed-phase simulations of carbohydrates." *J Mol Struct* **527**(1-3): 149-156.

Woods, R. J., R. A. Dwek, et al. (1995). "Molecular mechanical and molecular dynamic simulations of glycoproteins and oligosaccharides. 1. GLYCAM_93 parameter development." *J Phys Chem* **99**(11): 3832-3846.

Woods, R. J., K. N. Kirschner, et al. (2011). "Sulfating GLYCAM residues." Retrieved 27th July, 2011, from http://glycam.ccrc.uga.edu/ccrc/pages/sulfation.html.

Woods, R. J. and M. B. Tessier (2010). "Computational glycoscience: characterizing the spatial and temporal properties of glycans and glycan-protein complexes." *Curr. Opin. Struct. Biol.* **20**(5): 575-583.

Wright, T. M. (2006). "Tramiprosate." *Drugs Today* **42**(5): 291-298.

Wu, Y. F., J. H. Shien, et al. (2008). "Structural and functional homology among chicken, duck, goose, turkey and pigeon interleukin-8 proteins." *Vet. Immunol. Immunopathol.* **125**(3-4): 205-215.

Wu, Z. L. and M. Lech (2005). "Characterizing the non-reducing end structure of heparan sulfate." *J. Biol. Chem.* **280**(40): 33749-33755.

Xu-song, Z. and X. Bing-ren (2009). "Discontinued drugs in 2008: cardiovascular drugs." *Expert Opin Investig Drugs* **18**(7): 875-885.

Yamaguchi, M., T. Ohmori, et al. (2008). "Oligo(tyrosine sulfate)s as heparin pentasaccharide mimic: evaluation by surface noncovalent affinity mass spectrometry." *Bioorg. Med. Chem.* **16**(6): 3342-3351.

Yamauchi, H., P. Desgranges, et al. (2000). "New agents for the treatment of infarcted myocardium." *FASEB J.* **14**(14): 2133-2134.

Yang, C. Y., H. Sun, et al. (2009). "Importance of ligand reorganization free energy in protein-ligand binding-affinity prediction." *J. Am. Chem. Soc.* **131**(38): 13709-13721.

Yang, J. K., H. J. Yoon, et al. (2004). "Crystal structure of beta-D-xylosidase from Thermoanaerobacterium saccharolyticum, a family 39 glycoside hydrolase." *J. Mol. Biol.* **335**(1): 155-165.

Yap, K. L., J. B. Ames, et al. (2002). "Vector geometry mapping. A method to characterize the conformation of helix-loop-helix calcium-binding proteins." *Methods Mol. Biol.* **173**: 317-324.

Yates, E. A., F. Santini, et al. (1996). "1H and 13C NMR spectral assignments of the major sequences of twelve systematically modified heparin derivatives." *Carbohydr. Res.* **294**: 15-27.

Yip, G. W., M. Smollich, et al. (2006). "Therapeutic value of glycosaminoglycans in cancer." *Mol. Cancer Ther.* **5**(9): 2139-2148.

Yongye, A. B., J. Gonzalez-Outeirino, et al. (2008). "The conformational properties of methyl alpha-(2,8)-di/trisialosides and their N-acyl analogues: implications for anti-Neisseria meningitidis B vaccine design." *Biochemistry* **47**(47): 12493-12514.

Young, E. (2008). "The anti-inflammatory effects of heparin and related compounds." *Thromb. Res.* **122**(6): 743-752.

Yu, G., N. S. Gunay, et al. (2002). "Preparation and anticoagulant activity of the phosphosulfomannan PI-88." *Eur J Med Chem* **37**(10): 783-791.

Yuriev, E., M. Agostino, et al. (2011). "Challenges and advances in computational docking: 2009 in review." *J. Mol. Recognit.* **24**(2): 149-164.

Zehnder, J. L., M. Shatsky, et al. (1995). "Involvement of CD31 in lymphocyte-mediated immune responses: importance of the membrane-proximal immunoglobulin domain and identification of an inhibiting CD31 peptide." *Blood* **85**(5): 1282-1288.

Zhang, J., G. Riverst, et al. (2001). "Identification of inhibitors of heparin-growth factor interactions from combinatorial libraries of four-component condensation reactions." *Bioorg. Med. Chem.* **9**(4): 825-836.

Zhang, Y. (2008). "I-TASSER server for protein 3D structure prediction." *BMC Bioinformatics* **9**: 40.

Zhang, Y. (2009). "I-TASSER: fully automated protein structure prediction in CASP8." *Proteins* **77 Suppl 9**: 100-113.

Zhang, Y. J., B. J. Rutledge, et al. (1994). "Structure/activity analysis of human monocyte chemoattractant protein-1 (MCP-1) by mutagenesis. Identification of a mutated protein that inhibits MCP-1-mediated monocyte chemotaxis." *J. Biol. Chem.* **269**(22): 15918-15924.

Zhang, Z., S. A. McCallum, et al. (2008). "Solution structures of chemoenzymatically synthesized heparin and its precursors." *J. Am. Chem. Soc.* **130**(39): 12998-13007.

Zhao, J., D. J. Nelson, et al. (2006). "Potential influence of Asp in the Ca2+ coordination position 5 of parvalbumin on the calcium-binding affinity: a computational study." *J. Inorg. Biochem.* **100**(11): 1879-1887.

Zhou, H., N. Dussault, et al. (2009). M-ONC 402-a non anticoagulant low molecular weight heparin inhibits tumor metastasis. *100th Annual Meeting of American Association for Cancer Research (AACR)*. Denver, Colorado.

Zhou, H., S. Roy, et al. (2010). M402 –A novel heparan sulfate proteoglycan mimetic targeting tumor-host interactions. *American Association for Cancer Research (AACR)*. Washington, DC.

Zhou, Z., M. Bates, et al. (2006). "Structure modeling, ligand binding, and binding affinity calculation (LR-MM-PBSA) of human heparanase for inhibition and drug design." *Proteins* **65**(3): 580-592.

Zlotnik, A. and O. Yoshie (2000). "Chemokines: A new classification system and their role in immunity." *Immunity* **12**(2): 121-127.

Zlotnik, A., O. Yoshie, et al. (2006). "The chemokine and chemokine receptor superfamilies and their molecular evolution." *Genome Biol.* **7**(12): 243.

Zuijdendorp, H. M., X. Smit, et al. (2008). "Significant reduction in neural adhesions after administration of the regenerating agent OTR4120, a synthetic glycosaminoglycan mimetic, after peripheral nerve injury in rats." *J. Neurosurg.* **109**(5): 967-973.

Zwanzig, R. W. (1954). "High-temperature equation of state by a perturbation method. I. Nonpolar gases." *J Chem Phys* **22**(8): 1420-1426.

APPENDIX A

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Sep 10, 2010

This is a License Agreement between Neha S Gandhi ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2503511076102
License date	Sep 07, 2010
Licensed content publisher	John Wiley and Sons
Licensed content publication	Chemical Biology & Drug Design
Licensed content title	The Structure of Glycosaminoglycans and their Interactions with Proteins
Licensed content author	Neha S. Gandhi,Ricardo L. Mancera
Licensed content date	Dec 1, 2008
Start page	455
End page	482
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Order reference number	
Total	0.00 USD

APPENDIX B

This is a License Agreement between Neha S Gandhi ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Neha S Gandhi
Customer address	School of Biomedical Sciences, Bentley, other 6845
License number	2715780443898
License date	Jul 25, 2011
Licensed content publisher	Elsevier
Licensed content publication	Drug Discovery Today
Licensed content title	Heparin/heparan sulphate-based drugs
Licensed content author	Neha S. Gandhi, Ricardo L. Mancera
Licensed content date	December 2010
Licensed content volume number	15
Licensed content issue number	23-24
Number of pages	12
Start Page	1058
End Page	1069
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Molecular modelling of the interactions of complex carbohydrates with proteins
Expected completion date	Aug 2011
Estimated size (number of pages)	225
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.0 USD / 0.0 GBP
Total	0.00 USD

APPENDIX C

AMERICAN CHEMICAL SOCIETY LICENSE
TERMS AND CONDITIONS

Mar 04, 2011

This is a License Agreement between Neha S Gandhi ("You") and American Chemical Society ("American Chemical Society") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by American Chemical Society, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	621771362498
License Date	Mar 04, 2011
Licensed content publisher	American Chemical Society
Licensed content publication	Journal of Chemical Information and Modeling
Licensed content title	Molecular Dynamics Simulations of CXCL-8 and Its Interactions with a Receptor Peptide, Heparin Fragments, and Sulfated Linked Cyclitols
Licensed content author	Neha S. Gandhi et al.
Licensed content date	Feb 1, 2011
Volume number	51
Issue number	2
Type of Use	Thesis/Dissertation
Requestor type	Not specified
Format	Print
Portion	Full article
Author of this ACS article	Yes
Title of the thesis / dissertation	Molecular modelling of the interactions of complex carbohydrates with proteins
Expected completion date	May 2011
Estimated size(pages)	200

APPENDIX D

OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS

August 29, 2011

This is a License Agreement between Neha S Gandhi ("You") and Oxford University Press ("Oxford University Press") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2738050323199
License date	Aug 29, 2011
Licensed content publisher	Oxford University Press
Licensed content publication	Glycobiology
Licensed content title	
Licensed content author	Neha S. Gandhi, Ricardo L. Mancera
Licensed content date	first published online July 11, 2011
Type of Use	Thesis/Dissertation
Institution name	Curtin University
Title of your work	Molecular modelling of the interactions of complex carbohydrates with proteins
Publisher of your work	n/a
Expected publication date	Aug 2011
Permissions cost	0.00 USD
Value added tax	0.00 USD
Total	0.00 USD

APPENDIX E

OXFORD UNIVERSITY PRESS LICENSE

TERMS AND CONDITIONS

Sep 10, 2010

This is a License Agreement between Neha S Gandhi ("You") and Oxford University Press ("Oxford University Press") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2503520090431
License date	Sep 07, 2010
Licensed content publisher	Oxford University Press
Licensed content publication	Glycobiology
Licensed content title	Free energy calculations of glycosaminoglycan-protein interactions
Licensed content author	Neha S Gandhi, et. al.
Licensed content date	October 2009
Type of Use	Thesis/Dissertation
Institution name	Curtin University
Title of your work	Molecular modelling of the interactions of complex carbohydrates with proteins.
Publisher of your work	n/a
Expected publication date	Feb 2011
Permissions cost	0.00 USD
Value added tax	0.00 USD
Total	0.00 USD
Total	0.00 USD

APPENDIX F

ELSEVIER LICENSE
TERMS AND CONDITIONS

Sep 10, 2010

This is a License Agreement between Neha S Gandhi ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Neha S Gandhi
Customer address	School of Biomedical Sciences, Bentley, other 6845
License number	2503520364899
License date	Sep 07, 2010
Licensed content publisher	Elsevier
Licensed content publication	Carbohydrate Research
Licensed content title	Can current force fields reproduce ring puckering in 2-*O*-sulfo-α-l-iduronic acid? A molecular dynamics simulation study
Licensed content author	Neha S. Gandhi, Ricardo L. Mancera
Licensed content date	30 March 2010
Licensed content volume number	345
Licensed content issue number	5
Number of pages	7
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	

Title of your thesis/dissertation	Molecular modelling of the interactions of complex carbohydrates with proteins.
Expected completion date	Feb 2011
Estimated size (number of pages)	200
Elsevier VAT number	GB 494 6272 12

APPENDIX G

ELSEVIER LICENSE
TERMS AND CONDITIONS

Oct 31, 2010

This is a License Agreement between Neha S Gandhi ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Neha S Gandhi
Customer address	School of Biomedical Sciences, Bentley, other 6845
License number	2539660711248
License date	Oct 31, 2010
Licensed content publisher	Elsevier
Licensed content publication	Molecular Genetics and Metabolism
Licensed content title	22. Small molecule inhibitors of glycosaminoglycan biosynthesisas substrate optimization therapy for the mucopolysaccharidoses
Licensed content author	Jill Brown, J.R. Brown, R. Carroll, C. Glass, B.E. Crawford
Licensed content date	February 2010
Licensed content volume number	99
Licensed content issue number	2
Number of pages	1
Type of Use	reuse in a journal/magazine
Requestor type	author of new work
Intended publisher of new work	Elsevier
Elsevier Company	Elsevier
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No

Order reference number	
Title of the article	Heparin/Heparan Sulphate drugs
Publication new article is in	Drug Discovery Today
Publisher of the new article	Elsevier
Author of new article	Neha S. Gandhi and Ricardo L. Mancera
Expected publication date	Nov 2010
Estimated size of new article (number of pages)	12
Elsevier VAT number	GB 494 6272 12

“Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.”