USING MOLECULAR DYNAMICS SIMULATIONS TO UNDERSTAND MECHANISMS OF NEOVASCULARIZATION IN AGE-RELATED MACULAR DEGENERATION

A thesis presented to the graduate faculty of
New England College of Optometry in partial fulfillment
of the requirements for the degree of Master of Science

Jennifer A. Williams

July 2015

©Jennifer A. Williams
All rights reserved

The author hereby grants New England College of Optometry permission to reproduce and to distribute publicly paper and electronic copies of the thesis proposal document in whole or in part.

USING MOLECULAR DYNAMICS SIMULATIONS TO UNDERSTAND MECHANISMS OF NEOVASCULARIZATION IN AGE-RELATED MACULAR DEGENERATION

Jennifer A. Williams

This manuscript has been read and accepted by the Thesis Examination Committee in satisfaction of the thesis requirement for the degree of Master of Science.

July 29, 2015
Date

Jean L. Spencer
Jean L. Spencer, Ph.D.
Graduate Faculty Advisor

Steven B. Koevary, Ph.D.
Thesis Committee Member

Débora Nickla, Ph.D.
Thesis Committee Member

Matthew Nugent, Ph.D.
Thesis Committee Member

8/04/2015
Date

Nancy J. Coletta
Nancy J. Coletta, O.D., Ph.D.
Director of Graduate Studies

New England College of Optometry

ABSTRACT

USING MOLECULAR DYNAMICS SIMULATIONS TO UNDERSTAND MECHANISMS OF NEOVASCULARIZATION IN AGE-RELATED MACULAR DEGENERATION

Jennifer A. Williams

New England College of Optometry, 2015

Age-related macular degeneration (AMD) is a leading cause of blindness in developed countries. It exists in two forms, dry (without new blood vessels) and wet (with new blood vessels), and destroys the region in the back of the eye responsible for central vision. The wet form often leads to significant visual impairment and legal blindness. A curative treatment for wet AMD remains elusive because the molecular mechanism behind the development of this disease is not well understood.

The growth of new blood vessels in the back of the eye is stimulated by overproduction of vascular endothelial growth factor (VEGF). The activity of VEGF is enhanced by binding to the protein fibronectin in the extracellular matrix (ECM). The availability of binding sites on fibronectin is regulated by prior binding of heparan sulfate (HS), which "opens" fibronectin and exposes hidden sites. The presence of an increased number of sites on fibronectin for VEGF binding promotes neovascularization.

The goal of this project was to better understand the mechanism of interaction between HS and fibronectin in order to discover improved targeted treatments for wet AMD. The hypothesis was that HS molecules with appropriate sulfation patterns can modulate fibronectin to enhance VEGF binding and subsequent blood vessel growth. The approach to uncover this mechanism was based entirely on computational methods using molecular visualization software combined with molecular modeling, molecular

dynamics, and molecular docking software.

Molecules of 12 to 36 sugars were constructed of heparin using fully sulfated disaccharides and of HS using patterns of fully sulfated and nonsulfated disaccharides. Molecular dynamics (MD) simulations at physiological conditions showed that each molecule was flexible, with flexibility increasing with the presence of more highly sulfated disaccharides. Additional MD simulations of each heparin/HS molecule with fibronectin (heparin-binding domains 12-14) assessed fibronectin "opening" by measuring the increase in the angle between domains 12 and 13. HS had a greater effect on increasing this angle than heparin, and the impact was dependent on HS sulfation pattern. The enhanced capability of HS to modulate fibronectin "opening" was attributed to an increased association with fibronectin through non-ionic interactions (van der Waals and hydrogen bonding). Based on this finding, a novel HS structure was created to optimize fibronectin "opening." Molecular docking studies were performed with fully sulfated and nonsulfated hexasaccharide probes to establish an optimal HS sulfation pattern and alignment for binding to fibronectin. Results suggested a 30-sugar structure with nonsulfated ends and a nonsulfated middle connected by fully sulfated spacers. The novel HS molecule was constructed and run in simulation with fibronectin, and results are still to be determined.

Future analysis of the intermolecular interactions during fibronectin "opening" by optimized HS structures will reveal specific components that play critical roles in this molecular mechanism. Approaches for disrupting this mechanism and controlling VEGF activity will provide an innovative basis for the development of new therapeutic options for a more targeted treatment of neovascular AMD.

ACKNOWLEDGEMENTS

This project would not have been possible without the exceptional guidance of Dr. Jean Spencer. Her knowledge of the subject matter, excellent writing skills, and strong background in research techniques truly propelled this project and lifted it to its full potential. Dr. Spencer was present since day one of this project in the summer of 2012, and I appreciate all of her support over the following three years. Dr. Spencer supported me even when it seemed that this project would not pan out, and when it seemed we could do anything and write any grant. Dr. Spencer is the type of advisor that I hope every student encounters once in their life, someone who supports them no matter what and truly inspires independent thinking. I cannot thank Dr. Spencer enough for all of her efforts and years of work that went into this research project.

Dr. Matt Nugent was also present since the day this project was developed, and he stuck by the research until fruition. Dr. Nugent was always seeking out new publications to share, and was constantly encouraging and promoting our research whenever he had the chance. Additionally, Dr. Nugent offered up his own lab space for over a year of this project, so that I had any equipment that I needed and the space necessary to perform my research. This project would not have been possible without Dr. Nugent's generosity.

Dr. Steven Koevary and Dr. Debora Nickla were kind enough to participate on my thesis committee at the New England College of Optometry. Your guidance and time spent on this project are much appreciated, and I could not have completed it without you.

Funding for this project was provided by the New England College of Optometry (NECO). NECO went above and beyond to provide funding not only for summer

research stipends, but also during the academic year in order to encourage continued research. Dr. Barry Fisch played an integral role in this project's success, and to him I am very grateful.

The Fight for Sight research organization also provided funding for over three months of research supporting this project. I am extremely thankful for Fight for Sight's generosity and their appreciation for this research that Dr. Spencer and I have worked so hard on for over three years.

Stephanie Adams and Amy Burke provided much needed moral support as fellow classmates, future fellow doctors of optometry, and friends. The countless hours we spent talking about the project, reviewing posters and PowerPoint presentations is something I will not forget and I am very grateful for their friendship. They will be two of my favorite colleagues throughout our careers.

I cannot write this acknowledgment section without recognizing Dr. Nancy Coletta, the Director of Graduate Studies, for all of her support and guidance. Dr. Coletta encouraged my research even when my program was taking unexpected twists and turns. I am grateful that NECO has professors like Dr. Nancy Coletta, because she is a great role model for anyone seeking a career in teaching, research, or optometry.

Thank you NECO community, Dr. Clifford Scott, Dr. Bina Patel and the wonderful class of 2015.

Finally, my thanks need to be extended to my family. Mom and Dad, thank you for listening to me talk about my project for hours on end at times, and for always supporting me no matter what I decided to do. Sarah, thank you for your support and encouragement.

TABLE OF CONTENTS

Abstract.....Page iii

Acknowledgements..... Page v

Table of Contents.….....Page vii

List of Figures.…..…..... .Page ix

Introduction.....Page 1

Basis of Scientific Question.....Page 1

Computational Approach.....Page 3

Computational Techniques……..Page 4

Project Goal .....Page 6

Methods.....Page 8

Biochemical Nomenclature.....Page 8

Software and Equipment.....Page 14

Molecular Modeling.....Page 14

Molecular Dynamics.....Page 16

Molecular Docking.....Page 18

Analysis.....Page 19

Results.....Page 21

Molecular Modeling: Heparin.....Page 21

Molecular Modeling: Heparan Sulfate.....Page 24

Molecular Modeling: Fibronectin….....Page 26

Molecular Dynamics: Heparin and Heparan Sulfate Systems.....Page 27

Molecular Dynamics: Fibronectin Systems.....Page 32

Molecular Docking.....Page 38

Final Optimized Structure and Simulation….....Page 43

Discussion and Conclusions.....Page 47

Molecular Modeling.....Page 47

Molecular Dynamics: All Systems.....Page 48

Molecular Dynamics: Heparin and Heparan Sulfate Systems.....Page 49

Molecular Dynamics: Fibronectin Systems.....Page 50

Molecular Docking.....Page 52

Final Optimized Structure and Simulation.....Page 52

General Summary and Future Work.....Page 54

Bibliography.....Page 56

Appendix I: Additional Tables.....Page 62

Appendix II: Abbreviations.....Page 64

Appendix III: Additional Figures.....Page 65

LIST OF FIGURES AND TABLES

Figure 1. Isomers of Hexose Sugar in Pyranose Ring Formation.....Page 9

Figure 2. Pyranose Ring Conformation.....Page 10

Figure 3. Disaccharide Units.....Page 11

Table 1. Nomenclature System.....Page 12

Figure 4. GAG Orientation….....Page 13

Table 2. Dihedral Angle Measurements.....Page 22

Figure 5. Molecular Modeling Construction.....Page 23

Figure 6. GAG Structures.....Page 26

Figure 7. Processed Fibronectin…….....Page 27

Figure 8. Thermodynamic Stability of MD Simulation.....Page 28

Figure 9. Trajectory Snapshots…….....Page 29

Figure 10. Shape Descriptors.....Page 30

Figure 11. Molecule Flexibility (θ_m vs. D_m).....Page 31

Table 3. Summary of Flexibility……...…Page 32

Figure 12. Molecular Dynamics System Alignment.....Page 33

Figure 13. MD Simulation Solvation Box….....Page 35

Figure 14. Change in Joint Angle…….....Page 36

Table 4. Summary of Changes in Joint Angle……..Page 37

Figure 15. Non-ionic Interactions..…….....Page 37

Figure 16. Hexasaccharide Probes….....Page 38

Figure 17. Ligand Output Structures….....Page 40

Figure 18. Fibronectin Docking Regions.....Page 41

Figure 19. Optimal Docking Structures.....Page 42

Table 5. GScores……...Page 43

Figure 20. Optimized Heparan Sulfate-30 Structure…….....Page 45

Figure 21. Heparan Sulfate-30/Fibronectin Alignment…….....Page 46

Table A-I. Summary of Molecular Dynamics Simulations….....Page 62

Table B-I. Summary of Docking Simulations….....Page 63

Table A-III. Molecular Dynamics Set Up Screen 1…….....Page 65

Table B-III. Molecular Dynamics Set Up Screen 2…….....Page 66

Table C-III. LigPrep Set Up Screen…….....Page 67

Table D-III. Docking Set Up Screen…….....Page 68

Table E-III. Glide Set Up Screen….....Page 69

INTRODUCTION

Basis of Scientific Question

Age-related macular degeneration (AMD) is a common eye condition in patients over 65 years old, with 1.6 million individuals affected in the United States (Friedman et al., 2004). AMD involves a gradual destruction of the macula, loss of central vision, and an inability to perform tasks requiring visual discrimination. The disease exists in two forms—dry or atrophic and wet or neovascular. Although wet AMD affects only a small percentage (about 10%) of people with this condition, it represents the more advanced form of AMD and accounts for the majority of severe vision loss (Chakravarthy et al., 2010). Wet AMD is defined by the presence of a neovascular membrane in the choroid beneath the macula of the retina. Current treatments of wet AMD include intravitreal injections of anti-VEGF (vascular endothelial growth factor) antibodies (Brown and Regillo, 2007). These injections are given on a monthly or bimonthly basis (although newer studies recommend less frequent injections) and are not curative to the disease (Curtis et al., 2012; Schmidt-Erfurth et al., 2011). Anti-VEGF drugs are also used in other ocular conditions such as diabetic retinopathy with macular edema and retinopathy of prematurity (Mintz-Hittner et al., 2011; Nicholson and Schachat, 2010).

In wet AMD the neovascular membrane represents the body's attempt to compensate for cumulative oxidative damages associated with aging and other risk factors (Zarbin, 2004). These factors compromise the normal structure of the extracellular matrix (ECM), which leads to a hypoxic state of inadequate oxygen and subsequent release of VEGF by retinal pigment epithelial (RPE) cells (Birch and Liang, 2007; Schlingemann, 2004). VEGF stimulates the growth of new blood vessels by binding to

receptors on endothelial cells and other cell types (Ferrara et al., 2003; Rahimi, 2006). This activity can be enhanced by interaction with components of the ECM such as the multi-domain protein fibronectin (Goerges and Nugent, 2004; Mitsi et al., 2006). VEGF binds with high affinity to fibronectin type III domains 12-14 (FNIII 12-14) (Wijelath et al., 2006). The FNIII 12-14 region is also the strongest heparin/heparan sulfate-binding domain (Hep2 domain) within fibronectin (Barkalow and Schwarzbauer, 1991; Carpentier et al., 2014). Because the domains in FNIII 12-14 lack disulfide bonds, the conformation of this binding region is highly sensitive to topographical stresses produced by bound molecules (Ambesi and McKeown-Longo, 2014).

VEGF binding to fibronectin is modulated by heparan sulfate proteoglycan, a class of macromolecules found throughout the body on cell surfaces and within the ECM (Bishop et al., 2007; Esko and Selleck, 2002). Heparan sulfate proteoglycan consists of heparan sulfate (HS) chains attached to a protein core. HS is a linear polysaccharide of repeating disaccharides (25-200) with varying degrees of sulfation arranged in specific patterns of highly sulfated domains and nonsulfated domains (Bernfield et al., 1999; Gallagher, 2001). Heparin, another linear polysaccharide found in mast cells, is often considered a highly sulfated form of HS (Liu et al., 2012), although its chain length is smaller (15-20 disaccharides) (Conrad, 1998). Hypoxic conditions modify the expression and structure of HS to a more abundant "active" form (Sandwall et al., 2009). Interaction of active HS with fibronectin causes a conformational change in FNIII 12-14 from "closed" to "open," exposing a hidden site for increased VEGF binding and enhanced activity (Mitsi et al., 2008). Disruption of this mechanism may be a new treatment strategy for preventing choroidal neovascularization in wet AMD.

Computational Approach

Heparin and HS obtained from natural sources are heterogeneous mixtures of closely related polymers of varying size and structure (Conrad, 1998). The ability to obtain pure materials for use in experiments aimed at studying structure-activity relationships between heparin/HS and proteins is a major challenge (Dulaney et al., 2012). The most reliable strategy for sourcing well-defined heparin and HS oligosaccharides is by chemical synthesis (Zulueta et al., 2013). However, commercial synthetic products have major limitations in both chain size and available sulfation patterns (see http://www.v-labs.com). As a result, laboratory experiments are often hampered by the inability to vary structural parameters that are critical to the interaction.

The use of computational methods to study the heparin/HS-fibronectin interaction that "opens" the protein is an ideal technique for gaining insight into a mechanism which requires a relatively long chain (greater than 22 sugars) with some sulfation (Mitsi et al., 2006). The ability to model heparin/HS chains to precise specifications, followed by dynamic interactions with fibronectin, is a viable alternative for exploratory studies.

In the last 15 years, computational methods have become an increasingly common way of exploring the mechanisms behind natural and disease processes within the human body (Hansson et al., 2002). Many research groups have employed these techniques quite extensively in the investigation of heparin/HS interactions with proteins. A partial protein list includes annexin A2, antithrombin (AT), endostatin, fibroblast growth factor-2 (FGF-2), fibronectin, interleukin-8 (IL-8), neutrophil elastase, platelet endothelial cell adhesion molecule-1 (PECAM-1), stromal cell-derived factor-1α (SDF-1α), and β-tryptase (Carpentier et al., 2014; Gandhi et al., 2009, 2011; Krieger et al., 2004; Ricard-Blum et

al., 2004; Sapay et al., 2001; Spencer et al., 2006; Verli et al., 2005; Wang et al., 2015).

In the majority of these computational studies, the heparin/HS molecules were modeled as small chains having only 2-12 sugars. Of the two studies with larger molecules (24 sugars), the chains were assembled partially on the protein surface in a pattern of sulfated and nonsulfated hexasaccharides (Gandhi et al., 2011; Krieger et al., 2004). Working with unbound, long heparin/HS chains has been pointedly avoided because of the many difficulties associated with the system size, chain flexibility, and negative charges (Sapay et al., 2001). However, longer heparin/HS chains are necessary for realistic interactions with multimeric proteins, particularly if sulfation patterns play an important role in binding (Robinson et al., 2006; Stringer et al., 2002).

Computational Techniques

The term "molecular modeling" often includes all theoretical and computational techniques that provide insight into the behavior of molecular systems (Leach, 2001). For this project, molecular modeling is used more specifically to mean those computational techniques for building molecular structures. Molecular modeling allows researchers and drug developers to create novel molecules as well as modify existing biomolecules from archives such as the Protein Data Bank (PDB). A graphical user interface (GUI) enables visualization of the structure during addition and removal of charges, bonds, atoms, or fragments. The potential energy of the structure is calculated using a molecular mechanics force field with a solvation model, and subsequent energy minimization leads to the lowest energy (or most stable) conformation of the molecule. Examples of molecular modeling software are CHARMM (Harvard University, Cambridge, MA)

MacroModel (Schrödinger, New York, NY), Tinker (Washington University, St. Louis, MO), and YASARA (YASARA Biosciences, Vienna, AU). Many of these software packages also include other capabilities such as molecular dynamics simulations.

Molecular dynamics (MD) simulations are computational techniques for studying the time-dependent behavior of a molecular system. Atoms and molecules are allowed to interact for a period of time by integrating Newton's laws of motion to generate successive configurations of the system. The resulting configurations, or trajectory, show how the positions and velocities of the components vary with time (Leach, 2001). Statistical mechanics is used to convert the atomic positions and velocities to average values of the macroscopic properties of the system (temperature, pressure, energy, and heat capacity). Molecular dynamics simulations are useful for observing the complex dynamics of a system over a wide range of time scales, from fast internal motions (picoseconds) to slow conformational changes (nanoseconds). A convenient GUI allows the user to input specific setup parameters, including simulation time, initial conditions (temperature, pressure, and volume), solvent effects, and ensemble type (e.g., isothermal-isobaric). Some of the major molecular dynamics programs are Amber (University of California, San Francisco, CA), CHARMM (Harvard University, Cambridge, MA), Desmond (D. E. Shaw Research, New York, NY), GROMACS (University of Groningen, Groningen, Netherlands), and NAMD (University of Illinois at Urbana-Champaign, Urbana, IL) (Liao et al., 2011).

Molecular docking is a technique that optimizes the intermolecular complex formed between two molecules through positioning and orientation (Lengauer et al., 1996). Docking is often used in the pharmaceutical industry to design new drugs and to optimize

drug interaction and efficacy (Jorgensen, 2004). The optimal position of one smaller molecule relative to a larger molecule can be obtained by positioning the smaller molecule, or ligand, in different orientations with respect to the larger molecule, or receptor. The generation of these ligand-receptor poses (search algorithm) and the subsequent evaluation of the plausible structures (scoring function) are the basic components of docking. The structure of the receptor is required for docking, and the coordinates are often downloaded from the PDB. The structure of the ligand may also be downloaded from a library or archive, but in many cases it is built with modeling software. Some of the most popular docking tools are AutoDock (The Scripps Research Institute, La Jolla, CA), DOCK (University of California, San Francisco, CA), FlexX (BioSolveIT, Sankt Augustin, Germany), Glide (Schrödinger, New York, NY), GOLD (Cambridge Crystallographic Data Centre, Cambridge, UK), ICM (Molsoft, San Diego, CA), and Surflex-Dock (Tripos, St. Louis, MO) (Liao et al., 2011).

Project Goal

The goal of this project is to better understand the mechanism of HS-fibronectin interaction in order to discover improved targeted treatments for wet AMD. We hypothesize that HS molecules with appropriate sulfation patterns modulate fibronectin to enhance VEGF binding and subsequent blood vessel growth. Because defined HS preparations are difficult or impossible to obtain for experimental investigation, we will undertake a computational approach for investigating HS-fibronectin interactions at the molecular level. The four aims of this research are to **(1)** construct large heparin and HS molecules *in silico,* **(2)** interact the constructed molecules with fibronectin using

molecular dynamics, **(3)** optimize the HS structure and alignment with fibronectin using molecular docking, and **(4)** investigate the effects of optimized HS on fibronectin "opening."

METHODS

Biochemical Nomenclature

Glycosaminoglycans (GAGs) are linear polysaccharides, or sugar chains, with repeating disaccharide units. The two GAGs of interest for this project were heparin and heparan sulfate (HS). The sugar-ring units within these molecular structures are named and organized in a particular way by the International Union of Pure and Applied Chemistry (IUPAC) and the International Union of Biochemistry and Molecular Biology (IUBMB) (McNaught, 1996).

The individual hexose sugars are in the form of pyranose rings, which are six-membered rings with five carbon atoms and one oxygen atom. The carbon atoms within this ring are numbered from one to five, in a clockwise fashion, starting with the anomeric carbon atom (C1) next to the oxygen atom within the ring (Figure 1A). A hydroxymethyl (-CH_2OH) group extending from the fifth carbon atom (C5) in the ring completes the sugar with carbon atom C6. Oxygen and hydrogen atoms, which are directly bonded to the carbon atoms, take on the number of the carbon atom. For the oxygen atom (O5) in the ring, the designated number implies the carbon atom in the open chain structure of the sugar.

The pyranose ring (or closed chain structure of the sugar) forms L- and D-isomers. The L-isomer has the –CH_2OH group situated below the ring at C5, while the D-isomer has the –CH2OH group above the ring at C5 (Figure 1B, C). The pyranose ring can also exist in two isoforms, α and β**.** The α-isomer has the hydroxyl (–OH) group at C1 on the opposite side of the ring as the -CH_2OH group at C5, while the β-isomer has the -OH group on the same side of the ring as the –CH_2OH group (Figure 1D, E).

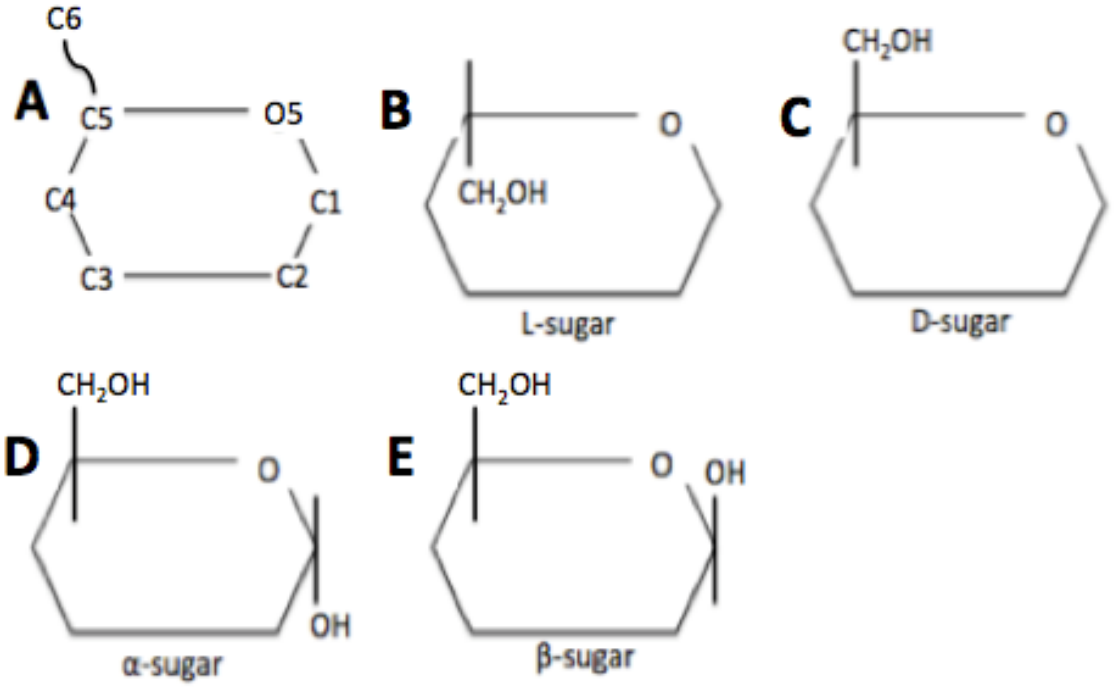


Figure 1. Isomers of hexose sugar in pyranose ring formation. Diagrams show **(A)** simple pyranose ring with numbered carbon atoms for hexose sugar, **(B)** L-isomer of hexose sugar, **(C)** D-isomer of hexose sugar, **(D)** α-isomer of D-hexose sugar, and **(E)** β-isomer of D-hexose sugar. The curved line in **(A)** indicates that the C6 group can be either above or below the ring.

Pyranose rings are puckered in reality, resulting in 38 possible conformations, but certain conformations provide more stable energy states for the sugars than others. Ring conformations include the boat (B), chair (C), and skew, or skew-boat, (S) conformations (Figure 2). The atoms above the plane of the sugar (determined by four atoms in the ring) are written as a superscript before the conformer abbreviation (B, C, or S), and the atoms below the plane are written as a subscript after the conformer abbreviation. For example, the $^{4}C_{1}$ conformation is shown in Figure 2B.

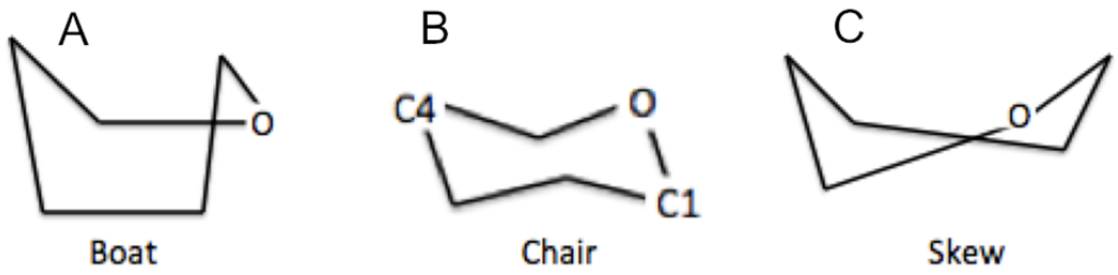


Figure 2. Selected conformations of pyranose ring. Diagrams illustrate **(A)** ^{1,4}B boat, **(B)** $^{4}C_1$ chair, and **(C)** $^{1}S_5$ skew (or skew-boat) conformations.

The repeating disaccharide units within the heparin and HS structures consist of a hexuronic acid residue and an α-D-glucosamine residue. The hexuronic acid can be either β-D-glucuronic acid or α-L-iduronic acid with possible 2-*O*-sulfation. The glucosamine has an amino group in place of the hydroxyl group at C2 on the ring, and the amino group can be acetylated or sulfated. The glucosamine can also have other locations of sulfation (mainly 6-*O*-sulfation). Disaccharides were identified using a four-character structural code (X#X#) in which the first two characters represent the hexuronic acid and the last two characters the glucosamine (Lawrence et al., 2008). For the hexuronic acid, the first character is the stereochemistry (G = glucuronic; I = iduronic), and the second is the sulfation (0 = none; 2 = 2-*O*-sulfation). For the glucosamine, the first character is the *N* substituent (A = *N*-acetylation; S = *N*-sulfation), and the second is the sulfation (0 = none; 6 = 6-*O*-sulfation). Based on this nomenclature, heparin was represented by the sequence $[I2S6]_i$, where the index *i* is the number of disaccharides in the chain. Likewise, HS was designated by various sequences of highly sulfated [I2S6] and nonsulfated [G0A0] domains; for example, $[I2S6]_i[G0A0]_j[I2S6]_k$, where *i, j, k* are the number of disaccharides in each domain. Images of the hexuronic acid residues and glucosamine residues are seen in Figure 3.

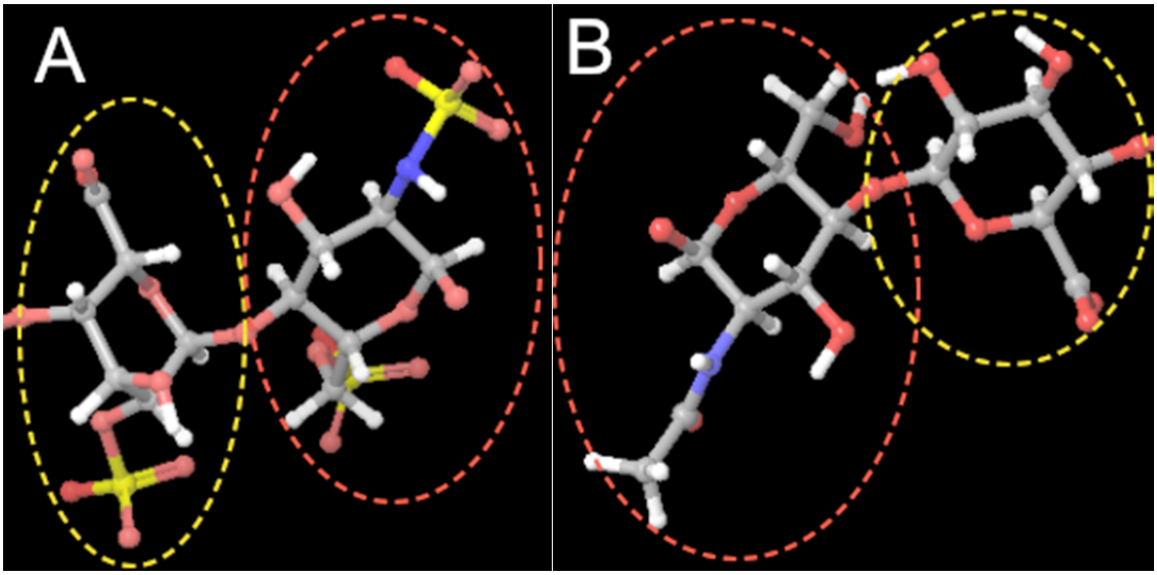


Figure 3. Disaccharide units representing fully sulfated disaccharide and nonsulfated disaccharide. The models show **(A)** a fully sulfated disaccharide [I2S6] from heparin consisting of iduronic acid (yellow dashed circle) and *N*-sulfated glucosamine (red dashed circle) and **(B)** a nonsulfated disaccharide [G0A0] from HS consisting of glucuronic acid (yellow dashed circle) and *N*-acetylated glucosamine (red dashed circle).

We often abbreviated the structures of heparin and HS for simplicity while working on this project. The abbreviated naming system first specified whether the GAG was heparin (Hep) or heparan sulfate (HS), followed by the number of monosaccharides or sugars within the molecular chain (often called the degree of polymerization, or dp). For example, the 12-sugar heparin was denoted as heparin-12 and abbreviated as Hep-12 (Table 1).

Name	Short Name	dp[a]	Disaccharide Structural Code[b]
Heparin-12	Hep-12	12	$[I2S6]_6$
Heparin-24	Hep-24	24	$[I2S6]_{12}$
Heparin-36	Hep-36	36	$[I2S6]_{18}$
Heparan sulfate-12	HS-12	12	$[I2S6]_2[G0A0]_2[I2S6]_2$
Heparan sulfate-24	HS-24	24	$[I2S6]_2[G0A0]_8[I2S6]_2$
Heparan sulfate-36	HS-36	36	$[I2S6]_2[G0A0]_{14}[I2S6]_2$
Heparan sulfate-36 with middle sulfation	HS-36mid	36	$[I2S6]_2[G0A0]_6[I2S6]_2[G0A0]_6[I2S6]_2$

Table 1. Summary of designations for heparin and HS molecules. The full name, abbreviated name, degree of polymerization, and four-character disaccharide structural code are listed for seven heparin/HS molecules built in this study.
[a] dp = degree of polymerization.
[b] See Lawrence et al., 2008.

The heparin and HS disaccharide chains contain a reducing end, indicated by the terminal glucosamine residue, and a non-reducing end, indicated by the terminal hexuronic acid residue (Figure 4A). These GAG chains were always depicted in our study with the non-reducing end on the left and the reducing end on the right. Each individual sugar was numbered from right to left within the molecule, with the reducing end having the lowest number (sugar number 1) and the non-reducing end having the highest number (e.g., sugar number 12 in Hep-12) (Figure 4B).

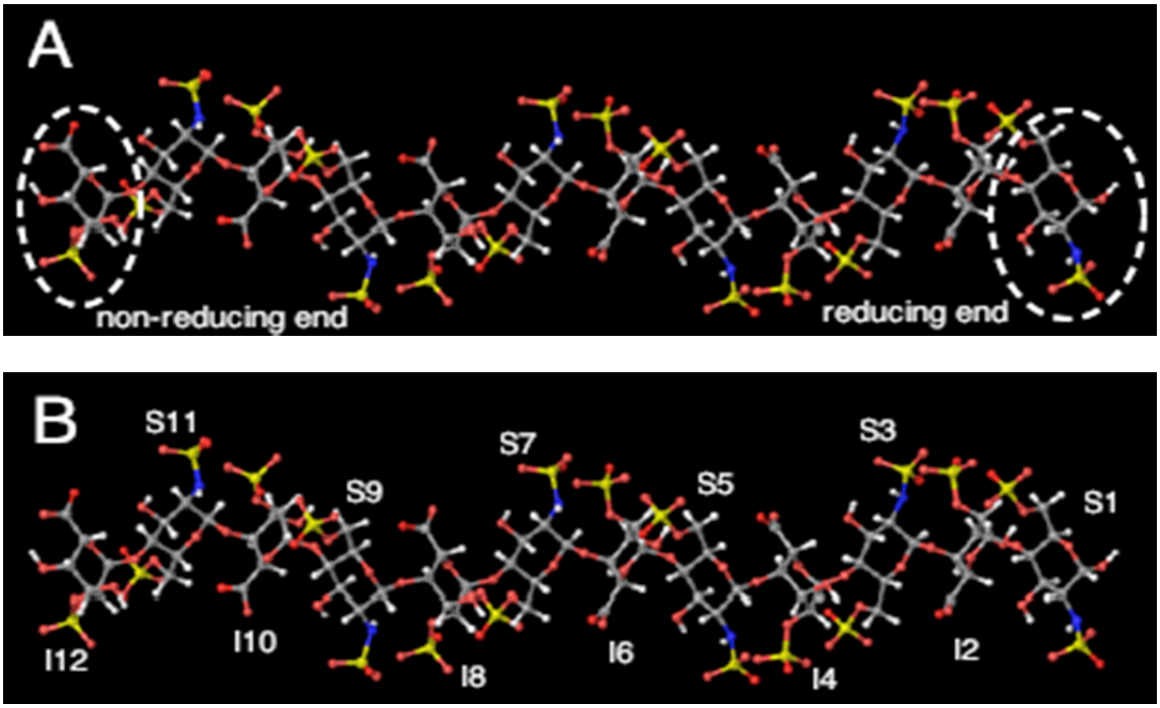


Figure 4. Nomenclature for representative GAG chain. **(A)** Non-reducing and reducing ends of heparin-12 (Hep-12) as encircled by the white dashed lines. **(B)** Sugar numbering system for Hep-12 starting at the reducing end with 1 and ending at the non-reducing end with 12. I = iduronic acid, S = *N*-sulfated glucosamine.

The backbone of the GAG chain is formed by glycosidic linkages between neighboring sugars, in which each linkage can be specified by two dihedral angles. The glycosidic dihedral angles that are pertinent for defining the structures of our linear polysaccharide are angle phi (Φ), defined by atoms $O5_i$-$C1_i$-$O4_{i-1}$-$C4_{i-1}$, and angle psi (Ψ), defined by atoms $C1_i$-$O4_{i-1}$-$C4_{i-1}$-$C3_{i-1}$, where *i* represents sugar *i*, and (*i*-1) represents sugar (*i*-1). Glycosidic dihedral angles were measured from the non-reducing end toward the reducing end of the molecule. Thus, dihedral angles measured for the glycosidic linkage between sugar *i* and sugar (*i*-1) were designated as belonging to sugar *i*.

Software and Equipment

All work for this project was conducted in laboratories in the Department of Biochemistry and the Department of Ophthalmology at Boston University School of Medicine (BUSM). Dr. Jean Spencer's laboratory at BUSM has a Thinkmate SuperServer 7046GTA-TF workstation (Thinkmate, Waltham, MA) that is dedicated to molecular modeling and molecular dynamics simulations. The system has an NVIDIA Tesla GPU computer processor with 448 cores. For software designed to take advantage of the simultaneously working multiple cores, extremely complicated data sets can execute in the fraction of time required by non-GPU stand-alone systems. The software on Thinkmate runs on a Linux platform (Red Hat Enterprise Linux, Red Hat, Raleigh, NC) and includes the commercial Schrödinger Suite 2014 (Schrödinger, New York, NY) with Maestro, MacroModel, Protein Preparation Wizard, LigPrep, and Glide, and freeware packages such as Desmond (D. E. Shaw Research), Visual Molecular Dynamics (VMD, Version 1.86) (University of Illinois at Urbana-Champaign, Urbana, IL), OpenOffice Calc (Apache Software Foundation, Forest Hill, MD) and R (R Foundation, Vienna, AU). In addition, Dr. Spencer's laboratory has several personal computers with Windows 7 (Microsoft, Redmond, WA) for data analysis and word processing (Microsoft Office 2010: Word, Excel, and PowerPoint).

Molecular Modeling

We implemented Maestro, MacroModel, and Protein Preparation Wizard to construct molecules and to prepare them for molecular dynamics simulations and molecular docking experiments. Maestro was used for building and visualizing structures,

MacroModel for energy minimization of the heparin/HS models, and Protein Preparation Wizard for refinement of the fibronectin model. The initial input files were molecular structures from the Protein Data Bank (PDB) of heparin (1HPN) (Mulloy et al., 1993) and fibronectin (1FNH) (Sharma et al., 1999). We modified these structures in various ways that included changing bonds, charges, and conformations, and adding/removing atoms and structural units. The resulting models were subjected to potential energy minimization using a molecular mechanics OPLS (optimized potential for liquid simulations)-2005 force field (Banks et al., 2005). This operation produced stable, low free-energy molecules as output.

The heparin and HS molecules were constructed from a 12-sugar heparin of fully sulfated disaccharides in PDB 1HPN. This heparin structure was frequently used in molecular dynamics and molecular docking studies in the literature (Becker et al., 2005; Carpentier et al., 2014; Spencer et al., 2006; Verli et al., 2004). We built larger molecules (Hep-24, Hep-36, HS-24, HS-36, and HS-36mid) than this base molecule because chains greater than 22 sugars were reported as the required length for modulating fibronectin for enhanced VEGF binding (Mitsi et al., 2006). The 24-sugar and 36-sugar molecules met this chain length criterion, and the 36-sugar molecule was also long enough to extend from one end of the fibronectin domains of interest to the other end.

In addition to chain extension, the construction of HS molecules required the placement of clusters (or domains) of nonsulfated disaccharides in the chain. There was minimal information in the literature on building HS chains, so the basic heparin chains were modified by replacing disaccharides in the central region with nonsulfated disaccharides. This generated HS molecules with a simple pattern of sulfated end

domains. Since the sulfation pattern of the HS molecule, rather than the sugar sequence, appeared to play an important role in HS-protein interactions (Spencer et al., 2010), we created a second 36-sugar HS molecule with a pattern of sulfated middle and end domains. In the final part of this project, we built an optimized HS molecule of 30 sugars using these same techniques and a sulfation pattern determined from molecular docking results.

Because these methods of constructing and modifying molecules were considered a major challenge in implementing this project, they are described in more detail in the Molecular Modeling section of Results.

Molecular Dynamics

We utilized Desmond software, a Schrödinger compatible program, to generate molecular dynamics (MD) simulations. The output structures from MacroModel and Protein Preparation Wizard were uploaded into Desmond for the simulations. Each simulation was set up in a periodic "box" with dimensions selected to prevent the molecule or molecular complex from "seeing" itself. The center of gravity of the model was positioned at the center of the box, and a minimum of 10 Å was maintained between the atoms in the model and the boundary of the box (de Souza et al., 1997). The box also contained explicit water molecules (TIP3P model) (Jorgensen et al., 1983), Na^+ and Cl^- ions to mimic physiological conditions (0.15 M), and Na^+ counter ions to maintain system electrical neutrality. The final solvated system contained 1.2-1.4 x 10^5 atoms. Molecular mechanical force field parameters were assigned to all components in the box using OPLS-2005 (Banks et al., 2005; Jorgensen et al., 1988), a force field model which

worked well for carbohydrates of repeating disaccharides (Stortz et al., 2009). The solvated system was equilibrated by undergoing a six-stage relaxation phase (default protocol in Desmond). The actual simulation (or production phase) was then executed for 10 ns at constant temperature and constant pressure (310 K, 1 atm). During the simulation the coordinates of all atoms in the box were saved every 4.8 ps (trajectory snapshot), and the system parameters (temperature, pressure, volume, and energy) were saved every 1.2 ps. Execution time varied from 2-8 weeks depending on the size of the solvated system. The settings of all parameters for MD simulations are shown in Appendix III (Figures A-III and B-III).

Three groups of MD simulations were completed for this project. The first group involved placing each heparin/HS model into an appropriately sized box and running the simulation. The simulation was stopped at 1.2 ns for preliminary analysis and then restarted and run to 10 ns. The results were analyzed for the effects of chain length and sulfation patterns on molecular flexibility.

The second group of MD simulations consisted of fibronectin alone or in combination with one of the 36-sugar heparin/HS models. At the start of these simulations, the presence of the heparin/HS molecule required manual alignment (non-reducing end to reducing end) above the fibronectin surface (domains 12-14). The heparin/HS molecule was roughly centered on the cationic cradle, a patch of positively charged residues in domain 13, and then positioned by adjusting three intermolecular distances to about 20-25 Å while minimizing atom-atom contacts. The three distances were measured from (1) the glycosidic oxygen after sugar 30 to the α-carbon of Pro 83 (domain 12), (2) the glycosidic oxygen after sugar 18 to the α-carbon of Phe 124 (domain

13), and (3) the glycosidic oxygen after sugar 6 to the α-carbon of Glu 226 (domain 14). After the required alignment was completed, the simulation was run to 10 ns. The results were analyzed for the effect of heparin/HS on "opening" fibronectin and the number of intermolecular bonds formed during this activity.

The final MD simulation group contained only one 10-ns run of an optimized HS molecule with fibronectin. The 30-sugar HS model was constructed from the results of the molecular docking studies and was aligned above fibronectin according to the observed binding patterns while avoiding atom-atom contacts. The distance between the central glycosidic oxygen (after sugar 15) and the α-carbon on Phe 124 (domain 13) was adjusted to approximately 20 Å, a positional constraint from the previous setups.

Molecular Docking

We used LigPrep and Glide to carry out molecular docking studies to identify optimal interaction/binding sites between hexasaccharide probes and fibronectin. Prior to actual docking, LigPrep was used to import hexasaccharide models from MacroModel and generate a library of ligands with different conformations for each model. These conformations were energy-minimized (OPLS-2005) and ranked, and 32 of the best conformations (most stable) were saved to an output file. This file, together with the fibronectin output file from Protein Preparation Wizard, was read into Glide for molecular docking within a specified "grid" space (25 x 25 x 25 Å) around a region of fibronectin. Five overlapping grids were generated in order to perform docking across the entire three domains of fibronectin. In each grid, the probe was aligned within 10 Å of the fibronectin surface before docking was initiated. Many of the setup parameters were

based on previous docking studies of GAGs with proteins (Gandhi et al., 2009; Sapay et al., 2011).

Glide docking was implemented as an exhaustive search of possible locations and orientations for each probe within the protein space defined by each grid. A series of filters of increasing refinement was applied to remove unlikely solutions from the search. The use of the extra-precision (XP) mode of Glide allowed the program to determine with highest accuracy the ligand-protein poses and the associated binding free energy, or Glide Score (GScore). The GScore is an empirical scoring function that approximates the binding free energy for the ligand binding to the protein. The complex equation has many terms that can be simplified as the summation of four types of energy within the system:

$$\text{GScore} = E_{coul} + E_{vdW} + E_{bind} + E_{penalty}$$

E_{coul} is the energy associated with the electrostatic force between atoms (Coulomb's Law), E_{vdW} is the energy associated with the van der Waals forces, E_{bind} accounts for all remaining forces that favor binding, such as those from hydrophobic interactions and hydrogen bonds, and $E_{penalty}$ is associated with the forces present from ligand strain that hinder binding (Friesner et al., 2006).

A table with all ligand-protein poses ranked from lowest to highest GScore was output from the program. Optimal interaction/binding sites were selected from the ligand-protein poses with the lowest (most negative) binding free energies.

Analysis

We applied the Visual Molecular Dynamics (VMD) program to analyze Maestro model files and Desmond trajectory files. VMD provided the capability to measure

descriptors of molecular shape (e.g., angles, distances, and ring conformations) and to characterize inter-molecular interactions (e.g., hydrogen bonds and van der Waals interactions). With trajectory files, VMD allowed measurements made on one frame to be automatically made on all frames, providing an easy means to track parameters with time. The output files from VMD were imported into Calc and Excel for data organization and plot/figure generation, and Excel (Analysis ToolPak) and R were used for statistical analysis of the results.

RESULTS

Molecular Modeling

Heparin

All molecules constructed for this project were built from the base molecule of 12-sugar heparin from the Protein Data Bank (PDB code: 1HPN). The 1HPN chain structure of $[I2S6]_6$ consisted of six repeating 1,4-linked disaccharides of α-L-iduronic acid with 2-*O*-sulfation and α-D-glucosamine with *N*-sulfation and 6-*O*-sulfation (Mulloy et al., 1993). The downloaded file contained two nuclear magnetic resonance (NMR) solution conformers of this molecule with glucosamine residues in the $^4C_1$ chair conformation, and iduronic acid residues in either the $^1C_4$ chair or the $^2S_0$ skew-boat conformation. Since equilibrium favors the chair conformation (60:40) for internal iduronic acid residues in heparin chains (Ferro et al., 1986), the structure of the $^1C_4$ conformer was selected for further modeling. The molecule was manually adjusted to have correct bond orders and formal charges (total charge -24), and the terminal oxygen atoms (representing incomplete glycosidic linkages) were capped with hydrogen atoms. The potential energy of the molecule was minimized using the molecular mechanics OPLS-2005 force field with water as an implicit solvent. The resulting minimized structure (Hep-12) was checked for problems with overlapping atoms or altered ring conformations, but no issues were found. Manual measurements of the 11 glycosidic bond angles, 22 glycosidic bond lengths, and 22 glycosidic dihedral angles in Hep-12 showed small changes from the values of the initial 1HPN structure. Table 2 compares the glycosidic dihedral angles for Hep-12 and 1HPN.

To build the 24- and 36-sugar heparin molecules, we split the Hep-12 structure into

three 4-sugar fragments ($[I2S6]_2$ building blocks), replicated the middle fragment as needed, and re-formed the appropriate bonds between the fragments. Figure 5 illustrates the process for constructing the 24-sugar heparin in which the middle fragment was replicated three times for a total of four middle fragments between the two end fragments (non-reducing and reducing), giving a chain structure of $[I2S6]_{12}$. Inter-fragment connections using an α(1-4) linkage were made between the reducing end of one fragment and the non-reducing end of the preceding fragment. Glycosidic dihedral angles (ϕ, ψ) between the S6-I2 linkages were adjusted to the average values of Hep-12 from Table 2.

Model	Dihedral Angles	I2-S6	S6-I2	S6-G0	G0-A0	A0-G0	A0-I2
1HPN	ϕ ψ	-77.05 (0.10) 133.0 (0.2)	78.7 (1.0) 88.4 (0.3)	NA	NA	NA	NA
Hep-12	ϕ Ψ	-72.47 (0.09) 138.2 (0.4)	78.9 (0.8) 87.2 (1.0)	NA	NA	NA	NA
HS-12	ϕ ψ	-77.14 (0.14) 128.7 (0.5)	67.69 (0.08) 84.3 (0.6)	66.4 (1.0) 104.7 (0.8)	-73.70 (0.10) 119.0 (0.4)	78.0 (0.9) 108.55 (0.14)	71.67 (0.10) 69.9 (0.4)

Table 2. Glycosidic dihedral angles for 1HPN, Hep-12, and HS-12. Values are the average (standard deviation) for each model. Glycosidic dihedral angles are defined as $\phi = O5_i\text{-}C1_i\text{-}O4_{i\text{-}1}\text{-}C4_{i\text{-}1}$ and $\psi = C1_i\text{-}O4_{i\text{-}1}\text{-}C4_{i\text{-}1}\text{-}C3_{i\text{-}1}$, where *i* is sugar *i* in the chain. For 1HPN and Hep-12: all linkages (N = 11). For HS-12: I2-S6 (N = 4); S6-I2 and G0-A0 (N = 2); S6-G0, A0-G0, and A0-I2 (N = 1).

The same procedure was used for the construction of 36-sugar heparin except that a total of seven middle fragments were required for the chain structure of $[I2S6]_{18}$. As a result of the building process, bond orders and formal charges (total charges -48 for Hep-24 and -72 for Hep-36) were correct for both molecules. The potential energy of each

model was then minimized with the OPLS-2005 force field with implicit water, and the resulting minimized structures were examined for problems. Manual measurements of the glycosidic bond angles (23 in Hep-24; 35 in Hep-36), glycosidic bond lengths (46 in Hep-24; 70 in Hep-36), and glycosidic dihedral angles (46 in Hep-24; 70 in Hep-36) completed the modeling process for the larger heparin molecules. The average values of these measurements were similar across all models. The final minimized structures of Hep-12, Hep-24 and Hep-36 are shown in Figure 6.

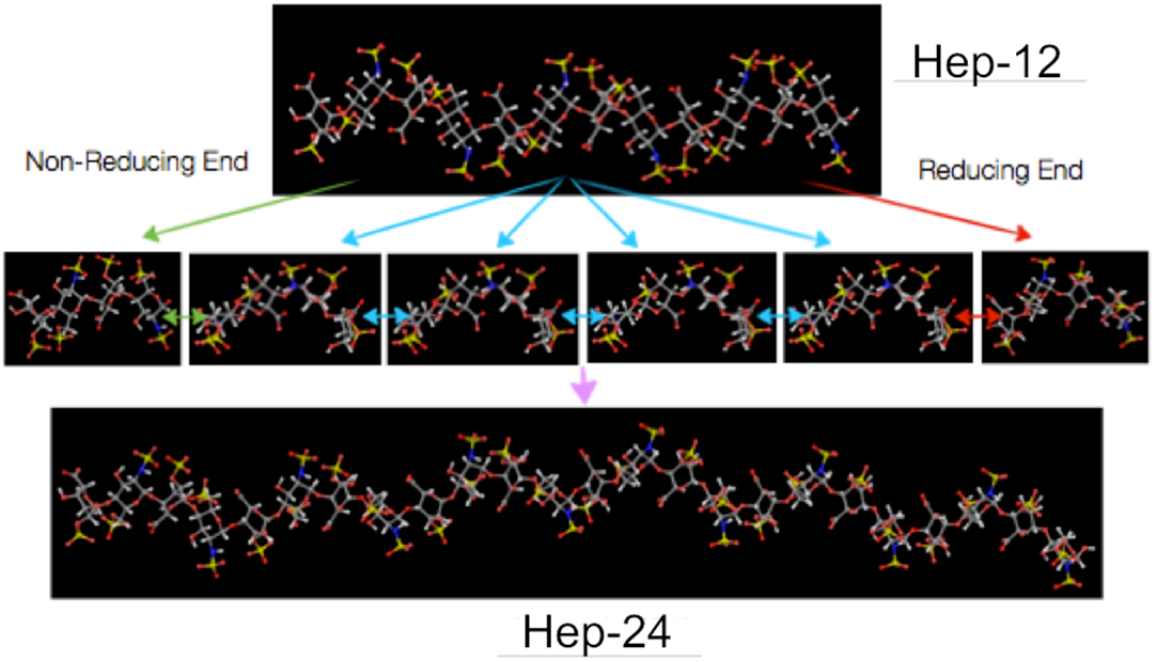


Figure 5. Molecular modeling process to construct heparin-24 (Hep-24) from heparin-12 (Hep-12). The starting Hep-12 model is an energy-minimized $^{1}C_{4}$ chair conformer of 12-sugar heparin from the Protein Data Bank (PDB code: 1HPN). The structure is initially split into three 4-sugar fragments. The two end fragments remain on the ends (green and red arrows), and the middle fragment is replicated for a total of four middle fragments (blue arrows). The fragments are joined by α(1-4) linkages with adjusted dihedral angles (double-sided arrows) and energy-minimized to form Hep-24. Atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur.

Heparan Sulfate

A 12-sugar heparan sulfate (HS) molecule was made from the Hep-12 model prior to energy minimization. For this structure the heparin end fragments ($[I2S6]_2$) were retained, while the middle fragment was extracted and replaced with a new middle fragment containing two nonsulfated disaccharides of β-D-glucuronic acid and α-D-glucosamine with *N*-acetylation ($[G0A0]_2$). Each nonsulfated disaccharide was made from preformed sugars joined internally by a β(1,4) linkage and connected by α(1,4) linkages to each other and to the end fragments. In this way we were able to establish a simple sulfation pattern (sulfated ends) in the HS model represented by the chain $[I2S6]_2[G0A0]_2[I2S6]_2$. The glycosidic dihedral angles (ϕ, ψ) of the internal G0-A0 linkage and the external A0-G0, S6-G0 and A0-I2 linkages were adjusted to values adapted from Blundell et al. (2009). The completed structure with correct bond orders and formal charges (total charge -18) was subjected to potential energy minimization according to the protocol used for the heparin models. The final model (HS-12) is shown in Figure 6, and the associated measurements of glycosidic dihedral angles are given in Table 2.

Larger HS molecules with 24 and 36 sugars and having sulfated ends were constructed from the HS-12 structure by the same process described for the larger heparin molecules. The resulting HS chains, $[I2S6]_2[G0A0]_8[I2S6]_2$ for 24 sugars and $[I2S6]_2[G0A0]_{14}[I2S6]_2$ for 36 sugars, were energy-minimized to the models HS-24 (total charge -24) and HS-36 (total charge -30) depicted in Figure 6. A final model was made from the HS-36 structure by extracting the central two nonsulfated disaccharides ($[G0A0]_2$) and replacing them with the heparin middle fragment of two highly sulfated

disaccharides ($[I2S6]_2$). The resulting chain $[I2S6]_2[G0A0]_6[I2S6]_2[G0A0]_6[I2S6]_2$ represented a 36-sugar HS model with a different sulfation pattern (sulfated middle and ends). The energy-minimized model (HS-36mid) is included in Figure 6 with the other HS models (HS-12, HS-24, and HS-36). Measurements of the glycosidic dihedral angles for all HS models were similar to the values for HS-12 in Table 2.

These same techniques were used for the modeling of an optimized HS molecule of unknown length and sulfation pattern based upon later molecular docking results (see Results: Final Optimized Structure and Simulation).

Fibronectin

The fibronectin model was imported from PDB 1FNH containing an X-ray diffraction structure of the heparin-binding domains 12, 13 and 14 (type III repeats) (Sharma et al., 1999). After water molecules were deleted, the fibronectin structure was adjusted by adding all hydrogen atoms, correcting bond orders and charge states, and capping the terminal ends (*N*-acetyl and *N*-methyl amide groups to N and C termini respectively). The hydrogen-bonding network was optimized by reorienting hydroxyl groups, flipping amide groups of asparagine (Asn) and glutamine (Gln), and selecting protonation and tautomeric states of histidine (His). The overall charge of the adjusted molecule was +10, resulting from 22 negatively charged aspartic acid (Asp) and glutamic acid (Glu) residues and 32 positively charged arginine (Arg) and lysine (Lys) residues. A final step of restrained energy minimization (4217 atoms) was carried out by using the OPLS-2005 force field and a termination criterion of less than 0.3 Å for the root-mean-square deviation (RMSD) of the heavy atoms. The final processed fibronectin molecule is shown in Figure 7.

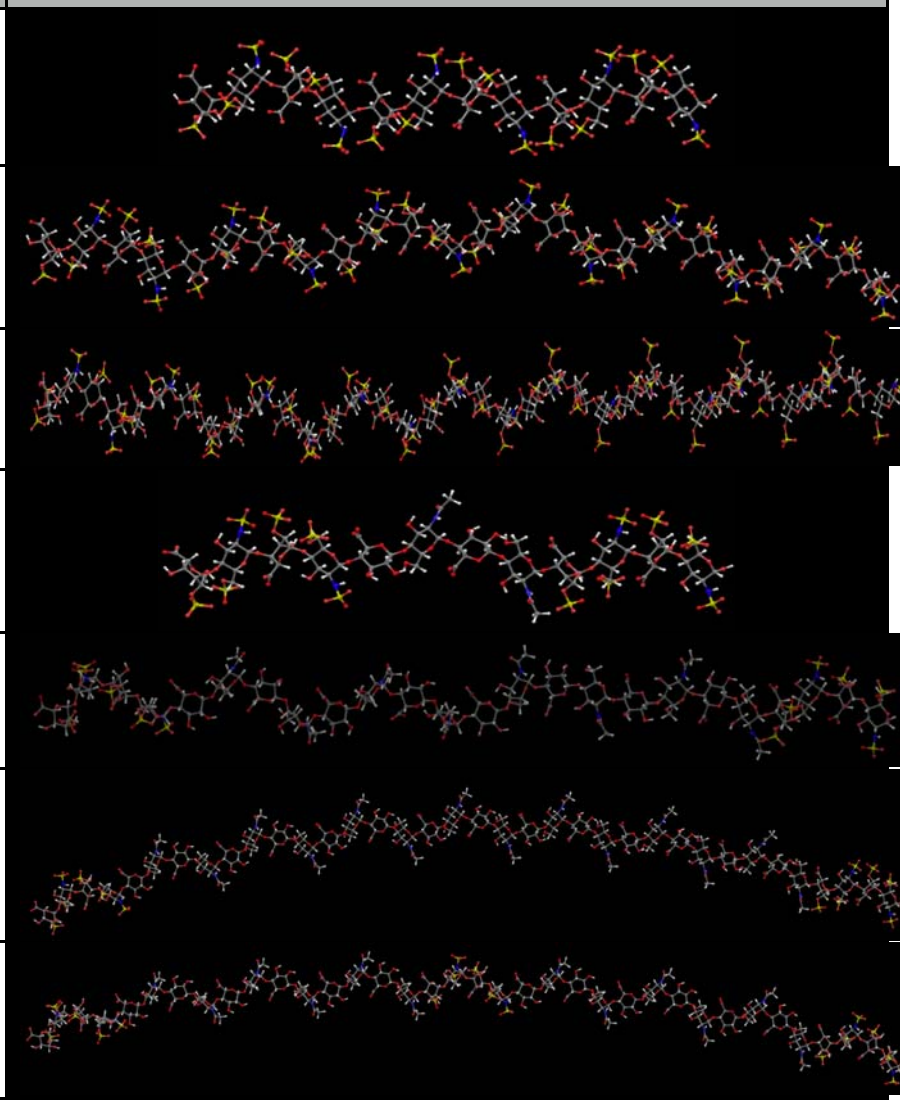
Structure	Image
Hep-12	
Hep-24	
Hep-36	
HS-12	
HS-24	
HS-36	
HS-36mid	

Figure 6. Models of heparin and HS with 12-36 sugars. The heparin disaccharide chains are $[I2S6]_6$ for Hep-12, $[I2S6]_{12}$ for Hep-24, and $[I2S6]_{18}$ for Hep-36. The HS disaccharide chains are $[I2S6]_2[G0A0]_2[I2S6]_2$ for HS-12, $[I2S6]_2[G0A0]_8[I2S6]_2$ for HS-24, $[I2S6]_2[G0A0]_{14}[I2S6]_2$ for HS-36, and $[I2S6]_2[G0A0]_6[I2S6]_2[G0A0]_6[I2S6]_2$ for HS-36mid. Atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur. For scale, Hep-12 is about 50 Å.

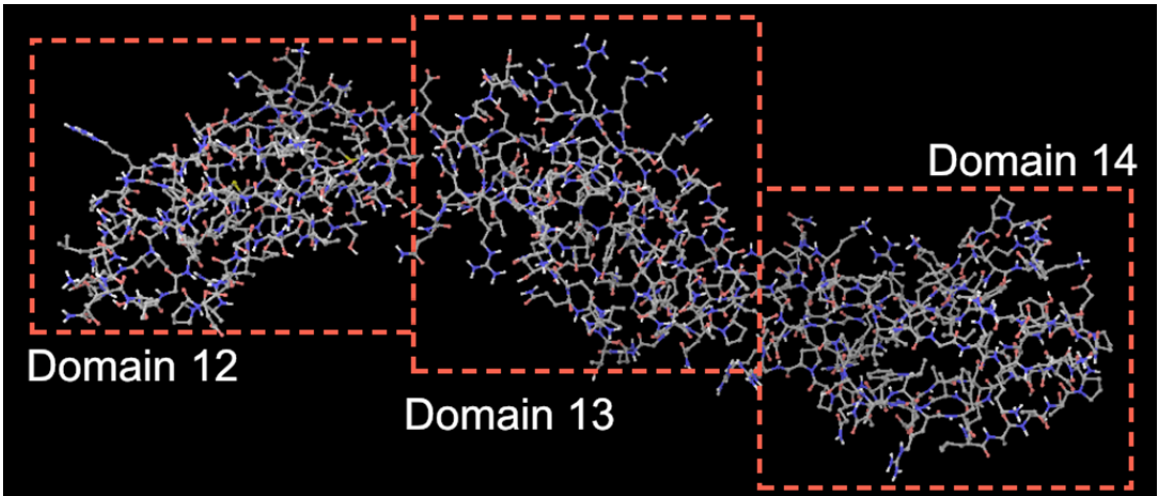


Figure 7. Fibronectin after processing and energy minimization by Protein Preparation Wizard. Domains 12, 13, and 14 are shown from left to right in red dashed boxes.

Molecular Dynamics

Heparin and Heparan Sulfate Systems

The seven heparin and heparan sulfate (HS) molecules, which were built with molecular modeling software (see Results: Molecular Modeling), were tested in isolation within a molecular dynamics (MD) simulation using Desmond software. Simulations required a solvation "box'" for each molecule (water with 0.15 M NaCl). The 12-sugar molecules were placed in the center of small boxes (70 x 70 x 70 Å), while larger boxes were used for the 24-sugar molecules (120 x 80 x 80 Å) and 36-sugar molecules (180 x 90 x 90 Å). Each molecule was run for 1.2 ns in simulation (310 K, 1 atm) and then stopped for data analysis. After measurements were completed, the simulations were continued to 10 ns. A summary of the physical parameters for the MD simulations for all seven molecules is included in Appendix I (Table A-I).

The simulations were first evaluated for stability by examining the fluctuations in temperature, total energy, and potential energy. A representative example of the simulations is shown in Figure 8 for HS-36mid. The data indicated no apparent drift in

any of the parameters with time. The percentage error (standard deviation/mean x 100) in temperature was 0.3% across all systems, and the percentage error in total energy varied from 0.06% to 0.10% (Table A-I).

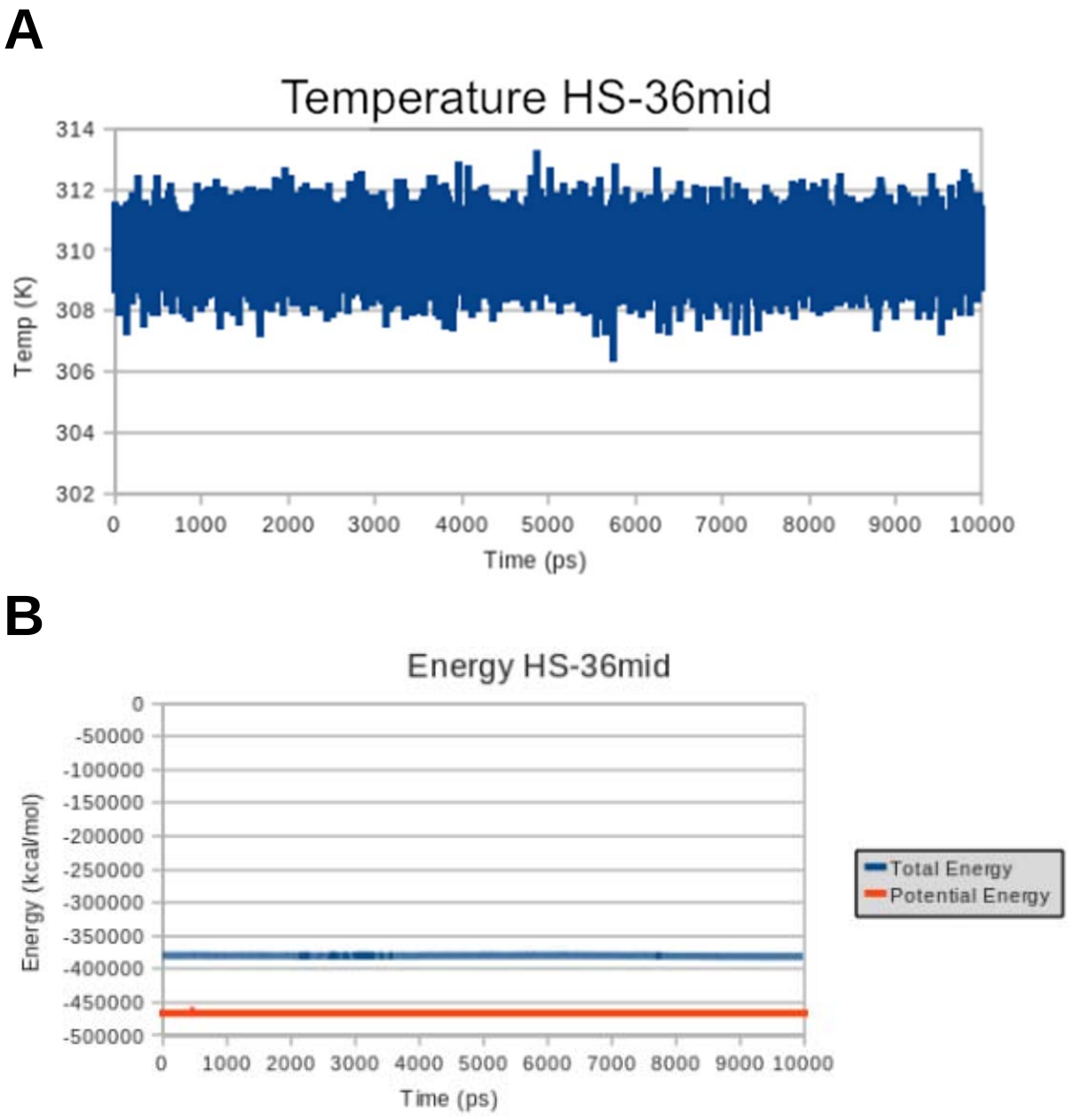


Figure 8. Thermodynamic stability of MD simulation for HS-36mid at physiological conditions. The variations in **(A)** temperature and **(B)** energy over 10 ns demonstrate a stable system for molecular dynamics. Average (±standard deviation) values: temp = 310.0 (±0.8) K; total energy = -3.800 (±0.004) x 10^5 kcal/mol; potential energy = -4.660 (±0.003) x 10^5 kcal/mol. N = 2084.

Another indicator of the quality of the simulation was the minimum approach distance for the atoms in the periodic box. This distance was greater than 50 Å for the 12-sugar molecules, 30 Å for the 24-sugar molecules, and 20 Å for the 36-sugar molecules (Table A-I). Examples of trajectory snapshots from the molecular dynamics simulations are displayed in Figure 9 for Hep-24 and HS-24 at 0 ns (frame 0), 0.6 ns (frame 125), and 1.2 ns (frame 250). As illustrated by the snapshots, these molecules fluctuated between extended and bent configurations during the 1.2 ns period.

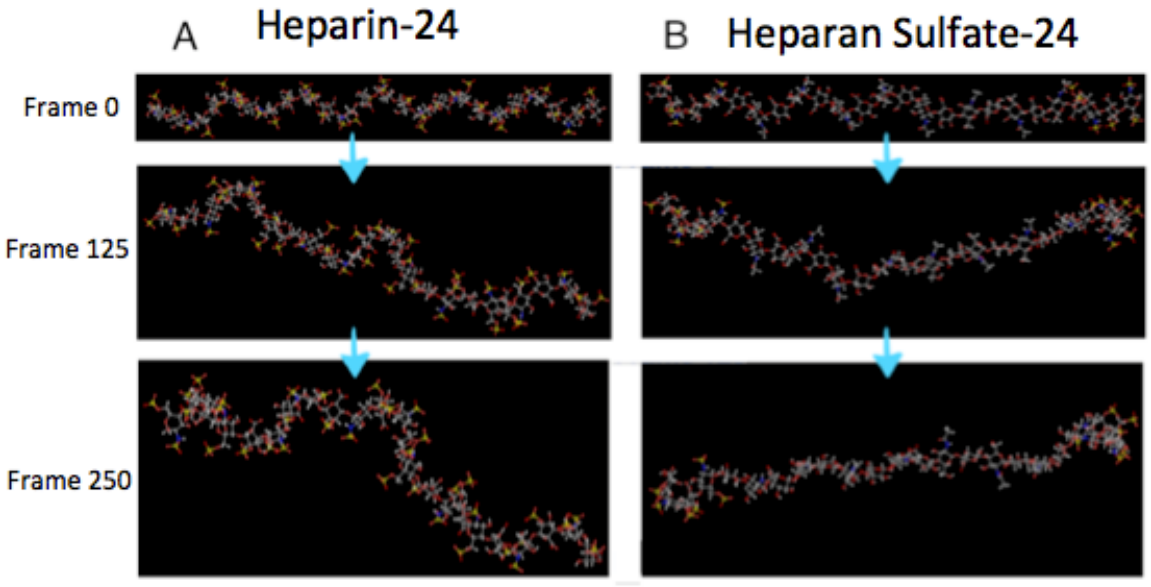


Figure 9. Trajectory snapshots from MD simulations of Hep-24 and HS-24 at physiological conditions. The images show the molecular fluctuations of **(A)** heparin-24 (Hep-24) and **(B)** heparan sulfate (HS-24) at 0 ns (frame 0), 0.6 ns (frame 125), and 1.2 ns (frame 250) of the simulations. Atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur.

To describe the dynamic aspect of these molecules with time, two shape descriptors were defined. The first was the modified distance (D_m) from the glycosidic oxygen atom between the first two disaccharides in the chain to the glycosidic oxygen atom between the last two disaccharides in the chain (Figure 10B). This parameter ignored the disaccharide at each end of the chain to compensate for possible end effects. The second

descriptor was the modified angle (θ_m) in which the vertex of the angle was the central glycosidic oxygen atom of the chain and the endpoints were the glycosidic oxygen atoms between the first two disaccharides and the last two disaccharides in the chain (Figure 10A).

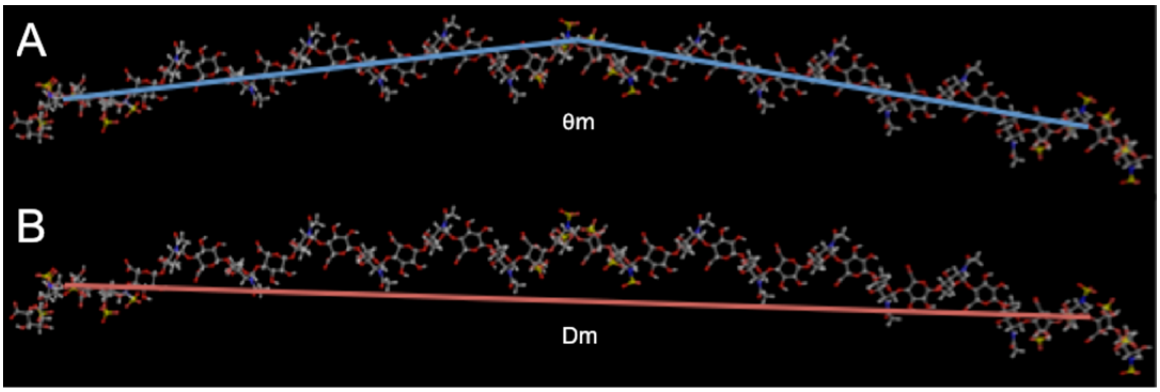


Figure 10. Shape descriptors for heparin/HS molecules during MD simulation. Two shape descriptors are defined to describe molecular flexibility: **(A)** a modified angle (θ_m), represented by the blue line extending from the vertex at the central glycosidic oxygen to the endpoints at the glycosidic oxygens between the first two disaccharides and the last two disaccharides, and **(B)** a modified distance (D_m), represented by the red line extending from the glycosidic oxygen between the first two disaccharides to the glycosidic oxygen between the last two disaccharides. Atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur.

The values of D_m and θ_m were measured with VMD software from the trajectory frames collected every 4.8 ps during the 1.2 ns simulation (251 frames total) for each of the seven heparin/HS molecules. The results are plotted against each other (D_m versus θ_m) in Figure 11 to characterize the flexibility of the heparin/HS chains at physiological conditions. The average values of D_m and θ_m from the plots are tabulated in Table 3.

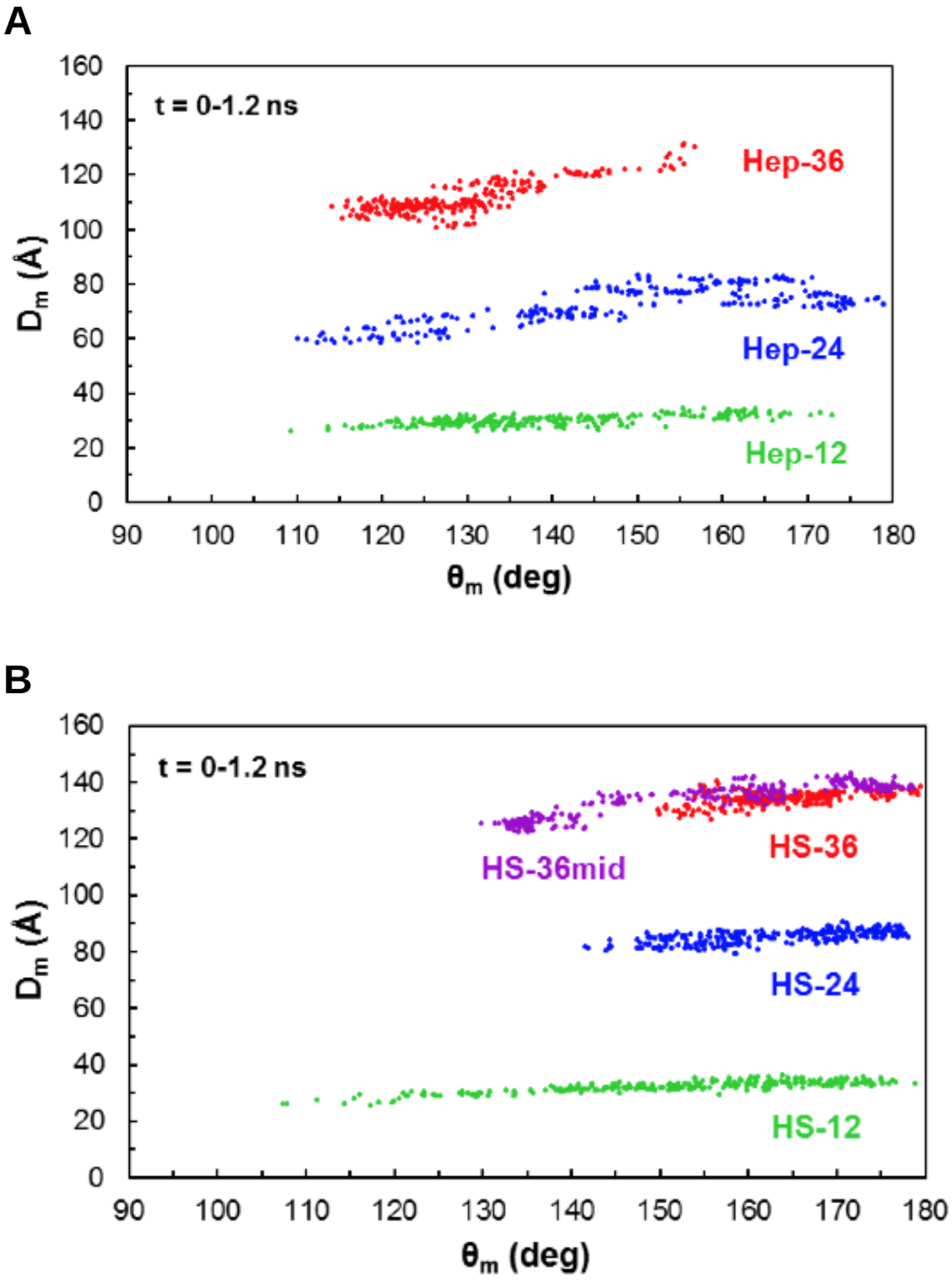


Figure 11. Molecular flexibility of heparin/HS molecules during MD simulations at physiological conditions. Results are plotted as modified distance (D_m) versus modified angle (θ_m) for **(A)** three heparin molecules (Hep-12, Hep-24, and Hep-36) and **(B)** four HS molecules (HS-12, HS-24, HS-36, and HS-36mid). Time = 1.2 ns; N = 251/system.

Molecule	D_m (Å)	θ_m (deg)	Molecule	D_m (Å)	θ_m (deg)
Hep-12	30 (±2)	140 (±14)	HS-12	32 (±2)	152 (±2)
Hep-24	72 (±7)	149 (±18)	HS-24	85 (±2)	163 (±9)
Hep-36	111 (±6)	130 (±9)	HS-36	134 (±2)	164 (±7)
			HS-36mid	134 (±6)	155 (±15)

Table 3. Summary of heparin/HS flexibility during MD simulations. Values are the average (±standard deviation) of the descriptors (D_m, θ_m) over time for each molecule in Figure 11. D_m = modified distance; θ_m = modified angle. Time = 1.2 ns; N = 251/system.

Fibronectin Systems

More complex MD simulations were performed with fibronectin alone or in combination with one of the 36-sugar molecules (Hep-36, HS-36, or HS-36mid). The fibronectin structure from Protein Preparation Wizard (see Results: Molecular Modeling) was placed in the center of a 180 x 90 x 90 Å solvation box (water with 0.15 M NaCl) and subjected to MD simulation with Desmond for 10 ns (310 K, 1 atm). Similar runs were repeated with fibronectin and the addition of a heparin/HS molecule to the system. Each sugar chain was manually aligned (non-reducing end to reducing end) above the fibronectin surface (domains 12-14) by adjusting three intermolecular distances to 20-25 Å. An example of the initial alignment is depicted in Figure 12 for Hep-36. The relation of the heparin-fibronectin complex to the box boundaries is indicated in Figure 13A, and the complete simulation box with water molecules and Na^+ and Cl^- ions is shown in Figure 13B. A summary of all MD simulations with fibronectin systems is presented in Appendix I (Table A-I).

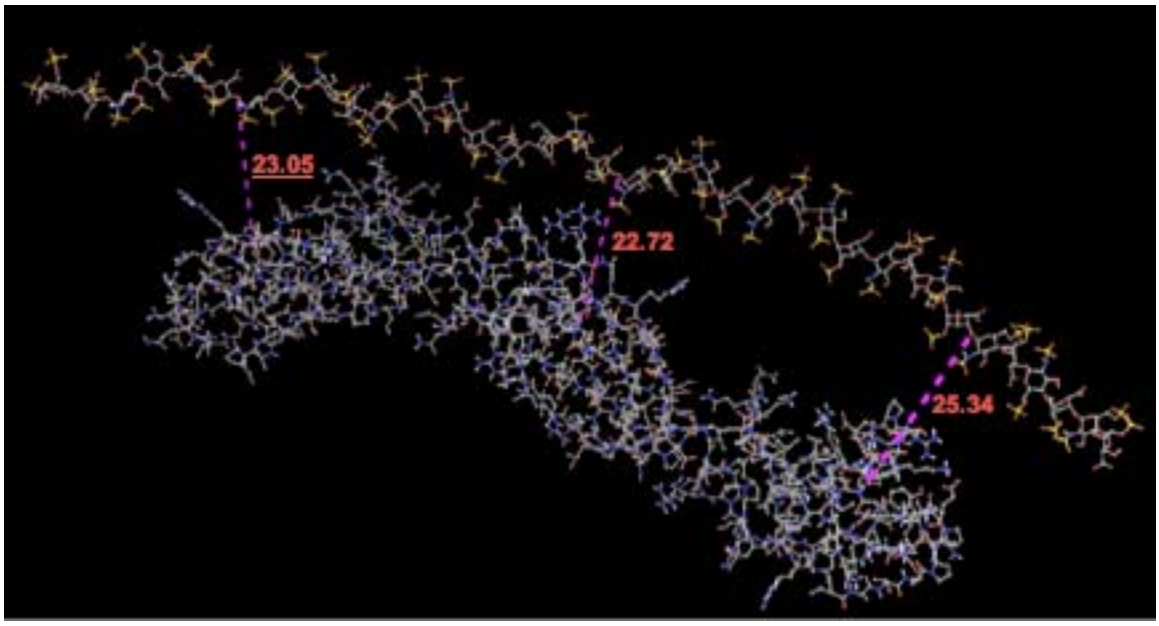


Figure 12. Alignment of Hep-36 above fibronectin before MD simulation. The Hep-36 molecule is manually aligned (non-reducing end to reducing end) above fibronectin domains 12, 13 and 14 (left to right). Three intermolecular distances (20-25 Å) are adjusted: (1) glycosidic oxygen after sugar 30 to α-carbon of Pro 83 (domain 12), (2) glycosidic oxygen after sugar 18 to α-carbon of Phe 124 (domain 13), and (3) glycosidic oxygen after sugar 6 to α-carbon of Glu 226 (domain 14). Atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur.

The stability of the fibronectin systems during the simulation was comparable to the results found for the heparin/HS systems. The percentage error in temperature was 0.3%, and the percentage error in total energy was 0.04-0.16%. A minimum approach distance greater than 20 Å was maintained throughout the simulations (Table A-I).

Various shape descriptors of fibronectin were screened in order to detect their sensitivity to possible "opening" of the protein. These descriptors consisted of the length between terminal residues of domains 12 and 14 (length 12-14), the inter-domain angle between domains 12 and 13 (angle 12-13) and domains 13-14 (angle 13-14), and the radius of gyration (R_g) for domains 12 and 13 (R_g 12-13) and domains 13 and 14 (R_g 13-14). The length and angle descriptors were defined using alpha-carbon atoms in the

indicated residues: length 12-14 = Pro 3 to Thr 271; angle 12-13 = Pro 3 to Asn 93 to Ala 183; and angle 13-14 = Asn 93 to Ala 182 to Thr 271.

Preliminary measurements of the selected shape descriptors with VMD software indicated that the inter-domain angle 12-13 had the greatest variation during the simulation period. This suggested that angle 12-13 may have the best chance of responding to a conformational change in the protein structure. Based on this finding, the shape descriptor for fibronectin "opening" was chosen as the inter-domain angle 12-13, or joint angle. This angle is illustrated in Figure 14A. An increase in the joint angle implies an "opening" of fibronectin to reveal additional VEGF binding sites.

Complete measurements of the joint angle were made on the trajectory files (2084 frames total) from the 10-ns MD simulations of the four fibronectin systems. The results are shown in Figure 14, where for each system, the last trajectory shapshot is displayed on the left, and the change in joint angle with time is plotted on the right. The average values of the joint angle from each plot are summarized in Table 4.

Although both heparin and HS exhibited a dynamic association with fibronectin throughout the simulation period, the HS molecules, particularly HS-36mid, appeared to have a closer association with the fibronectin surface (Figure 14C, D). Because non-ionic interactions (hydrogen bonds and van der Waals interactions) play an important role in this type of association, these interactions were analyzed for the last trajectory snapshot in each of the systems (Figure 14B-D). The number of non-ionic interactions for Hep-36/fibronectin, HS-36/fibronectin, and HS-36mid/fibronectin are compared in Figure 15.

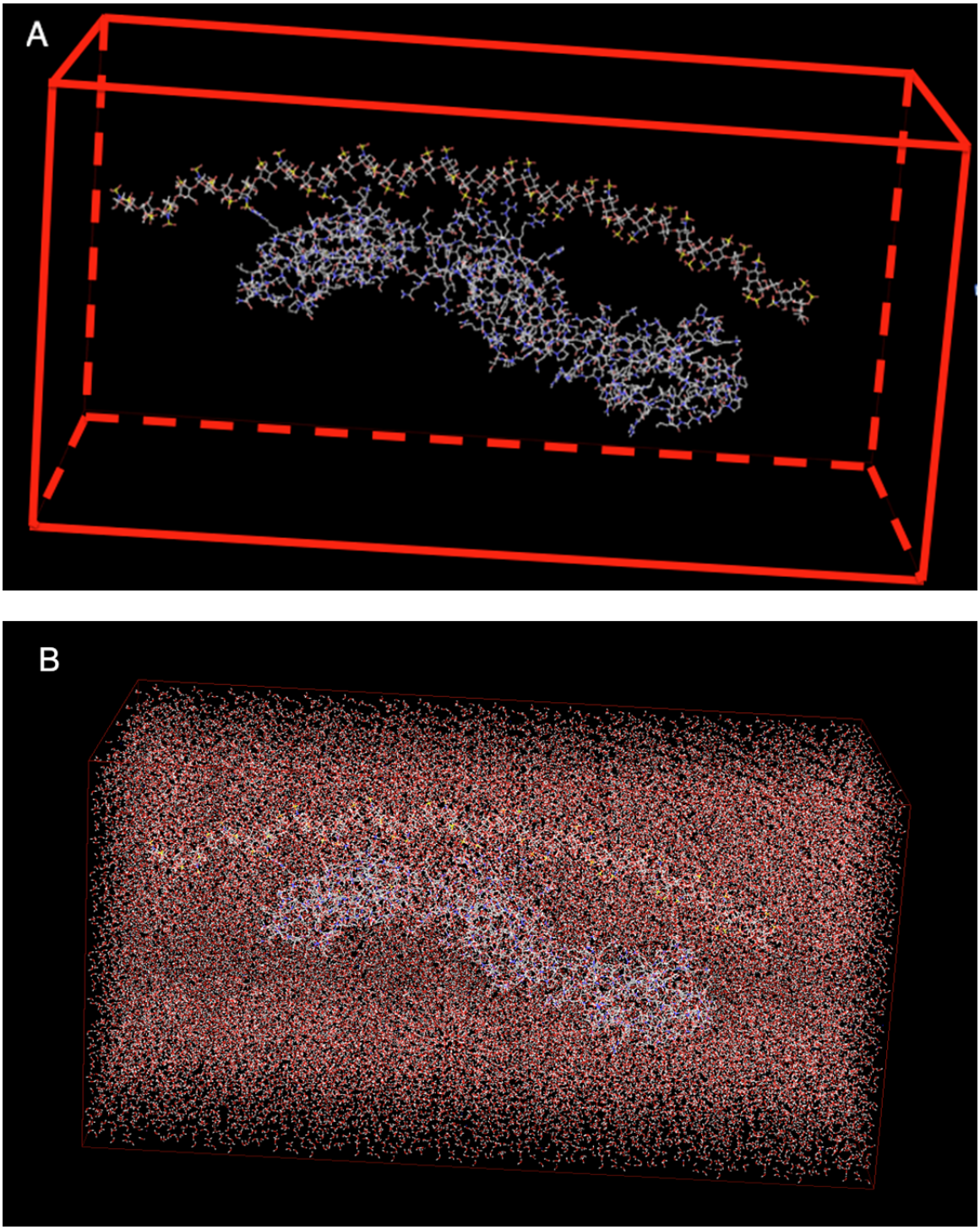


Figure 13. MD simulation box for Hep-36 and fibronectin. **(A)** The aligned complex of Hep-36 and fibronectin is centered in a box with dimensions of 180 x 90 x 90 Å. **(B)** The box containing Hep-36 and fibronectin is filled with water molecules and Na^+ and Cl^- ions to create a neutral environment at physiological conditions of 0.15 M. Total atoms = 1.39 x 10^5. Atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur.

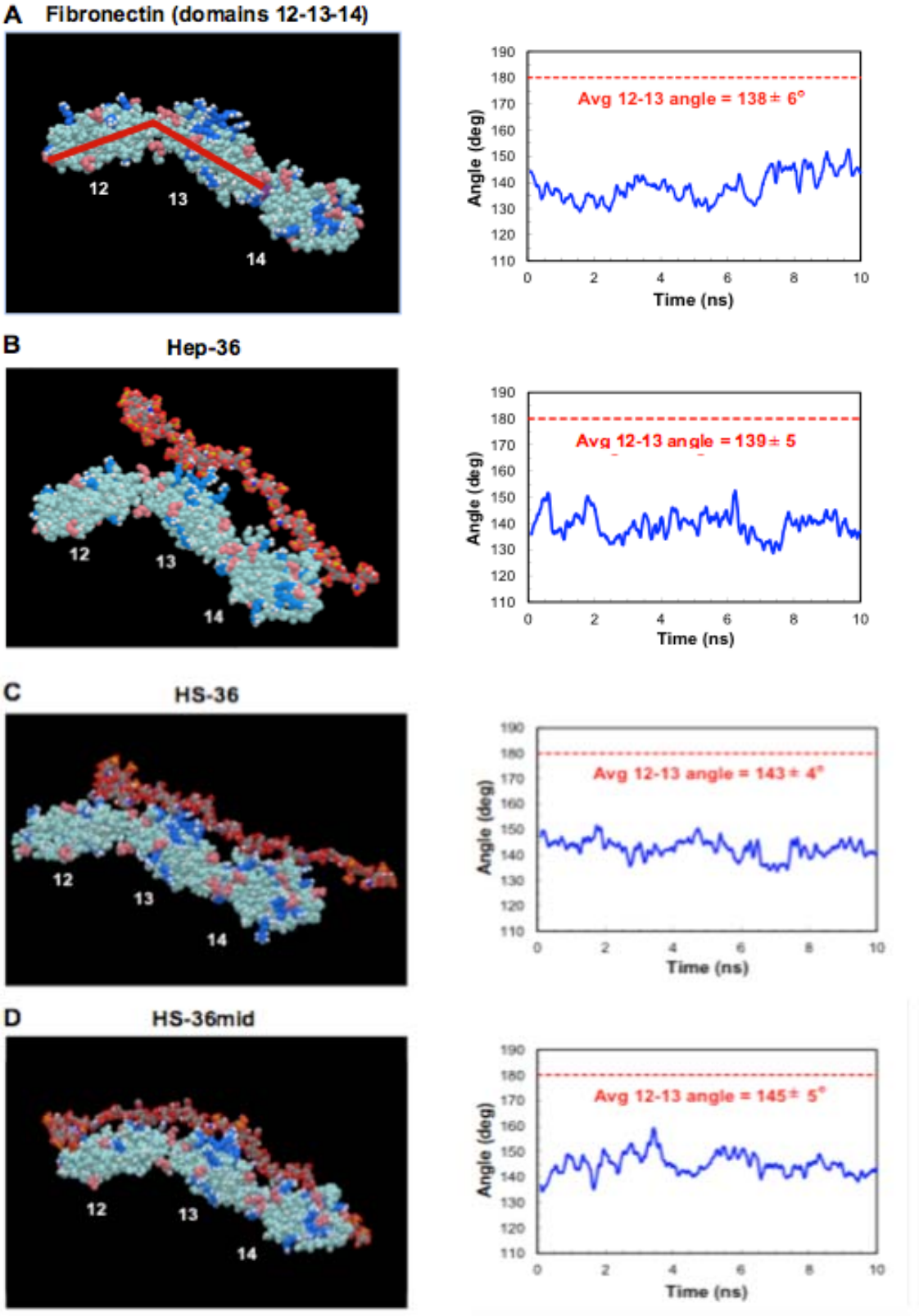


Figure 14. Changes in fibronectin joint angle mediated by heparin/HS molecules during MD simulations at physiological conditions. The panels show the results of 10-ns

simulations for **(A)** fibronectin (domains 12, 13, and 14) in isolation with the joint angle (inter-domain angle 12-13) in red, **(B)** Hep-36 and fibronectin, **(C)** HS-36 and fibronectin, and **(D)** HS-36mid and fibronectin. The last trajectory snapshot from each run is on the left, and a plot of joint angle versus time is on the right. Fibronectin atom colors: dark blue = basic, pink = acidic; heparin/HS atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur. N = 2084/system.

System	Inter-Domain Angle 12-13 (degrees)
Fibronectin	138 (±6)
Hep-36/Fibronectin	139 (±5)
HS-36/Fibronectin	143 (±4)
HS-36mid/Fibronectin	145 (±5)

Table 4. Summary of changes in fibronectin joint angle during MD simulations. Values are the average (±standard deviation) over time for each system in Figure 14. Joint angle = inter-domain angle 12-13. Time = 10 ns; N = 2084/system.

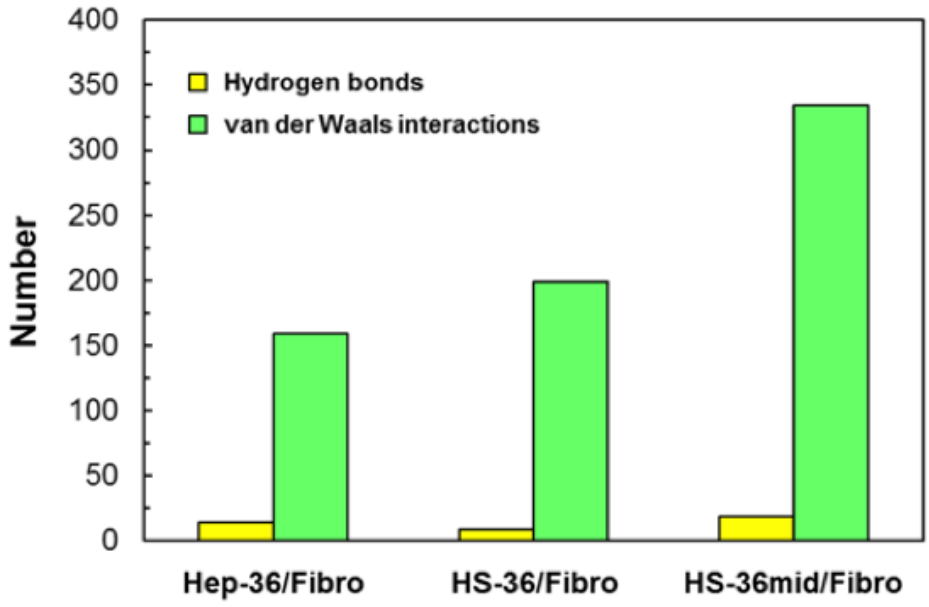


Figure 15. Non-ionic interactions between heparin/HS molecules and fibronectin after MD simulations at physiological conditions. The bar graph shows the number of non-ionic interactions (hydrogen bonds and van der Waals interactions) detected after 10 ns of MD simulations for Hep-36 and fibronectin, HS-36 and fibronectin, and HS-36mid and fibronectin.

Molecular Docking

Molecular docking studies were completed with the Schrödinger software Glide and supporting modules LigPrep, Protein Preparation Wizard, and MacroModel. Two hexasaccharide probes, one fully sulfated $[I2S6]_3$ and one nonsulfated $[G0A0]_3$, were created with molecular modeling software (MacroModel). These two probes were extracted from the central portions of previously modeled structures: Hep-24 for the fully sulfated probe and HS-24 for the nonsulfated probe (see Results: Molecular Modeling). Once the central six sugars were extracted from each of the larger molecules, hydrogen atoms were added to the ends of the hexasaccharides, and the structures were energy minimized. The resulting two probes were uploaded to LigPrep for pose production. The two hexasaccharide probes are shown in Figure 16.

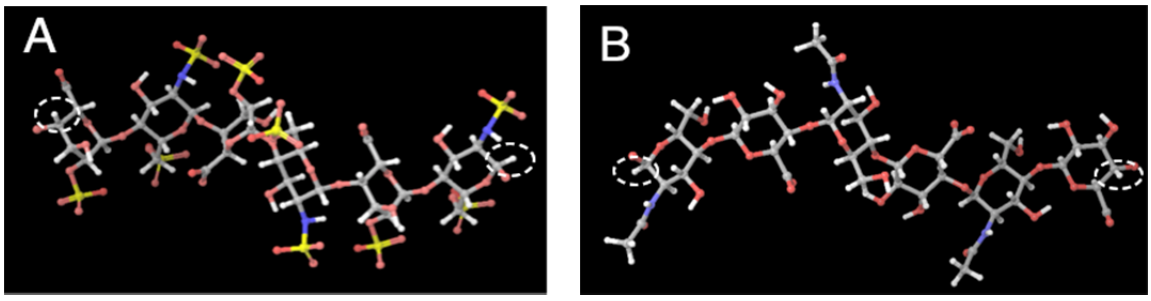


Figure 16. Hexasaccharide probes used for docking studies. Models represent **(A)** the fully sulfated probe $[I2S6]_3$ extracted from Hep-24 and **(B)** the nonsulfated probe $[G0A0]_3$ extracted from HS-24. White dashed circles depict added hydrogen atoms. Atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur.

The LigPrep function generated 32 poses for each hexasaccharide probe by examining diverse structural variations (e.g., tautomeric states, ring conformations, and stereoisomers). The final poses had optimal orientation and free energy states. All poses were saved to an output file of ligand molecules for use in docking. Examples of some of the output structures for the fully sulfated hexasaccharide probe are depicted in Figure 17.

After the LigPrep file was completed, a file containing the fibronectin molecule as prepared by Protein Preparation Wizard (see Results: Molecular Modeling) was uploaded. Because fibronectin (represented by domains, 12, 13 and 14) was too large to perform docking on the whole structure at once, the molecule was split into five smaller regions: (1) domain 12, (2) inter-domain space 12-13, (3) domain 13, (4) inter-domain space 13-14, and (5) domain 14. These regions are illustrated in Figure 18. A docking grid (25 x 25 x 25 Å) was generated around each region by centering the grid at the centroid of the set of residues in the domain or inter-domain space of interest. The grid represented the properties of the enclosed fibronectin structure (e.g., electrostatic potential) at each grid point so that docking calculations would be more efficient. Each of the five grids was saved as an output file.

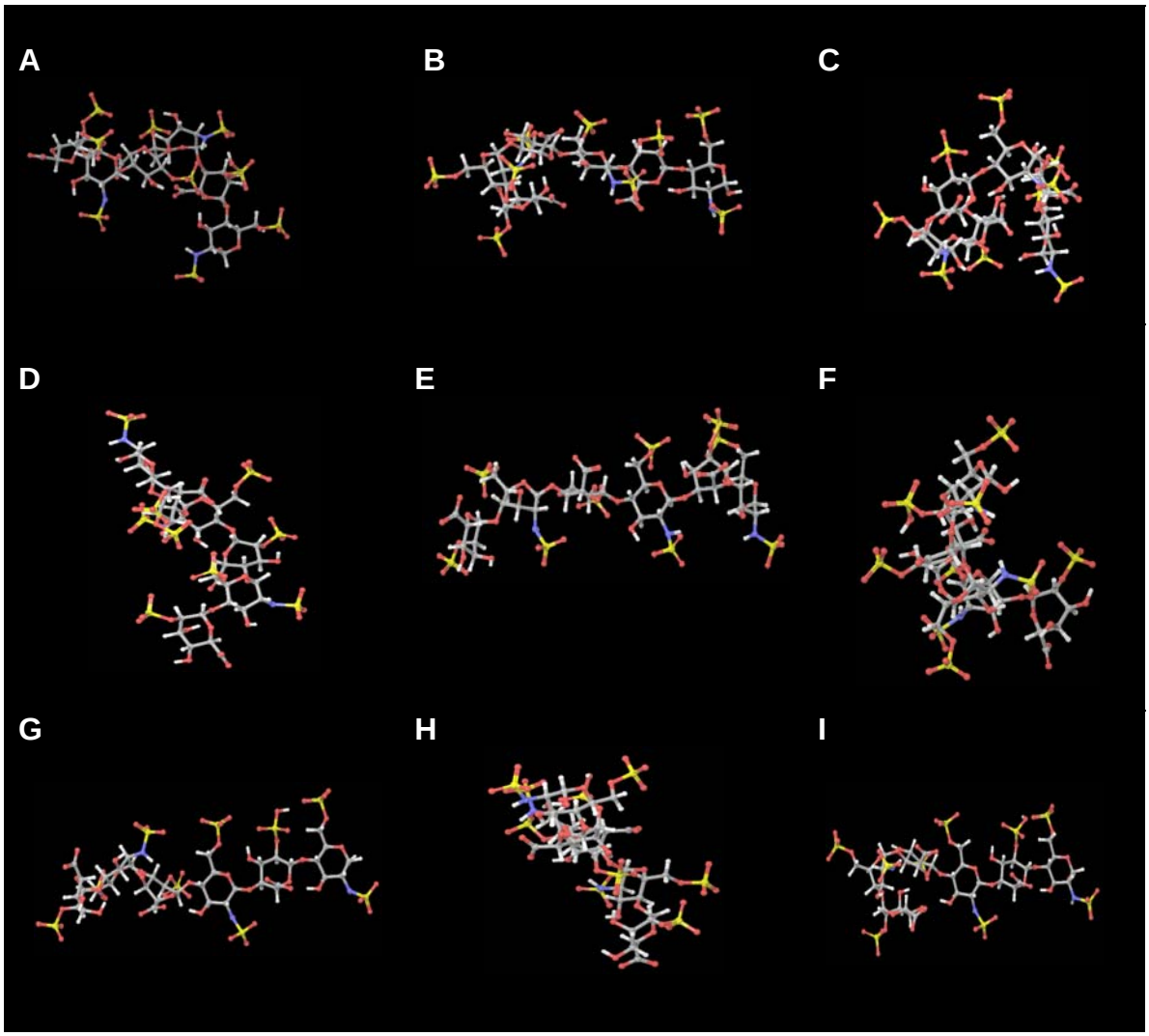


Figure 17. Nine poses of the fully sulfated hexasaccharide probe used for docking studies. A total of 32 poses with optimal orientation and free energy states were generated by LigPrep software from the original fully sulfated hexasaccharide probe (Figure 16A). Atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur.

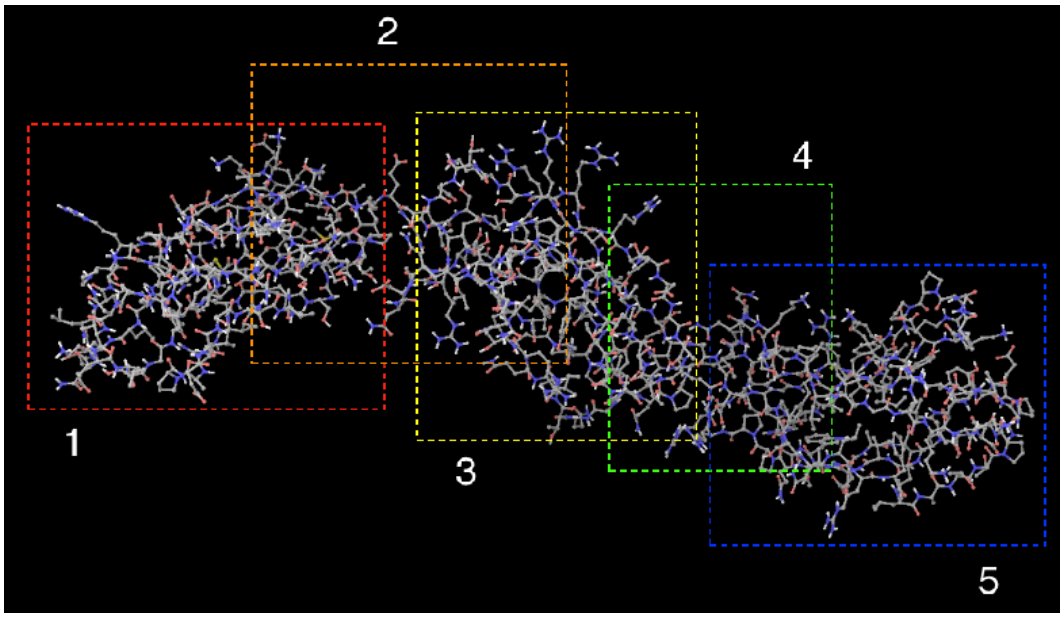


Figure 18. Regions of fibronectin used for docking studies. The colored boxes delineate five smaller regions of fibronectin: (1) domain 12, (2) inter-domain 12-13, (3) domain 13, (4) inter-domain 13-14, and (5) domain 14. Each region (25 x 25 x 25 Å) was separately evaluated in docking experiments. Atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur.

The fibronectin domains and inter-domain spaces were each tested separately in the docking studies. Each ligand was aligned within 10 Å of the protein in order to ensure optimal docking. This cutoff distance was drawn from previous docking studies performed by Gandhi et al. (2009). Once the probe was aligned, the docking simulation was run by uploading the grid output file and the ligand output file to the Glide docking software. Docking was performed in the Extra Precision (XP) mode for more detailed analysis. The output from XP docking was called the XP-Viewer file and contained a table of GlideScores (GScores, or optimal/stable free energies) from best (most negative) to worst (most positive) for the ligand poses input from the LigPrep file.

A summary of all docking runs is listed in Appendix I, Table B-I. Glide allowed the probes to be viewed in their most stable shape and orientation with respect to fibronectin in the pose viewer mode. Optimal fully sulfated and nonsulfated probes are depicted in

Figure 19 for each of the five fibronectin regions, and Table 5 shows the GScores for these poses.

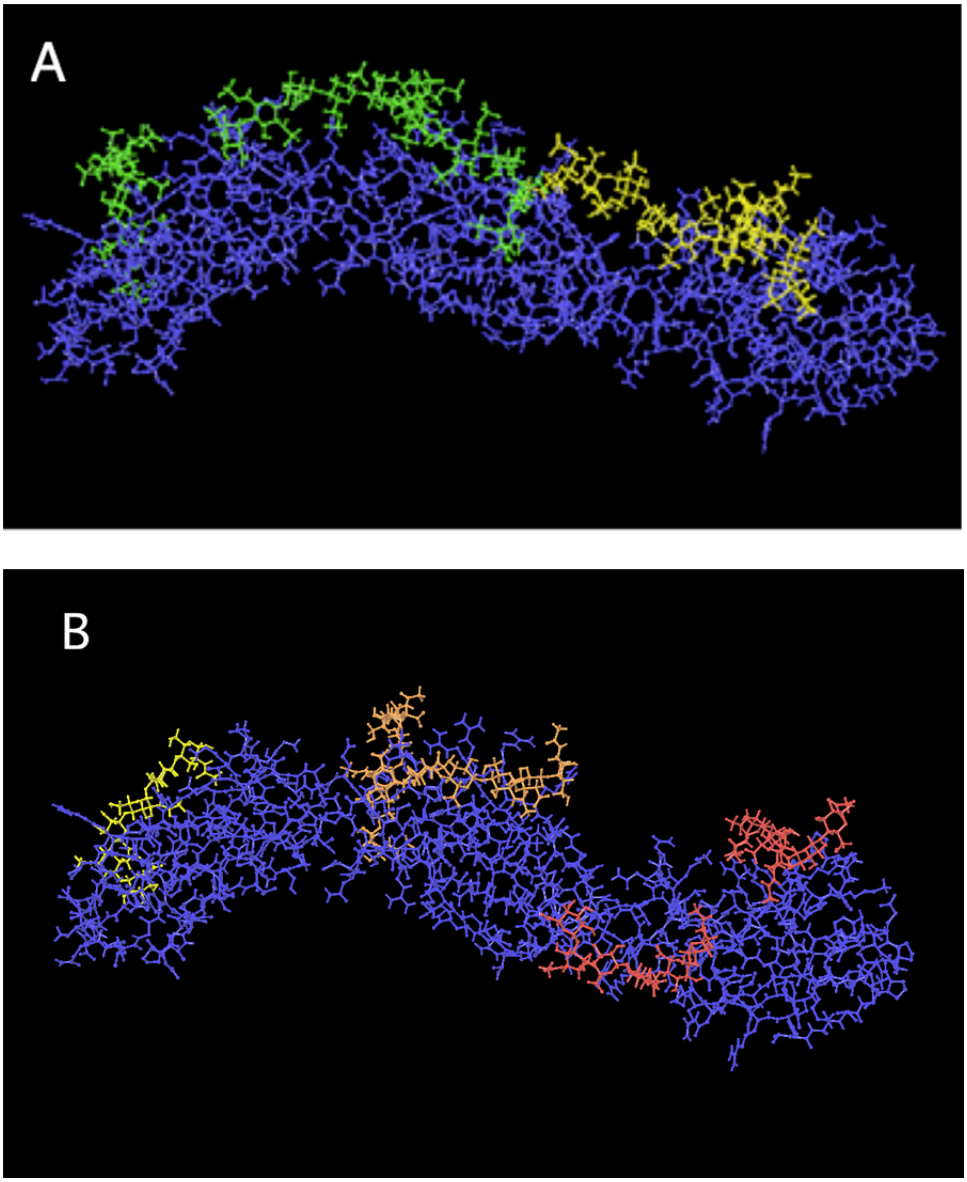


Figure 19. Optimal hexasaccharide-fibronectin interactions from docking studies. The five docking regions of fibronectin are depicted with optimal ligand poses: **(A)** fully sulfated hexasaccharides designated as sulf12, sulf1213, sulf13, sulf1314, and sulf14 (left to right) and **(B)** nonsulfated hexasaccharides designated as nonsulf12, nonsulf1213, nonsulf13, nonsulf1314, and nonsulf14 (left to right). The hexasaccharide probes are color-coded to show GScore or binding free energy (kcal/mol): green = -7.00 to -8.99, yellow = -9.00 to -10.99, orange = -11.00 to -12.99, and red = < -13.00. The more negative free energy score corresponds to a more optimal ligand location and configuration. The fibronectin molecule is oriented with domains 12, 13 and 14 from left to right.

Region	Fibronectin	Ligand	GScore (kcal/mol)
1	Domain 12	Sulf12	-8.7
		Nonsulf12	-10.6
2	Inter-domain 12-13	Sulf1213	-7.7
		Nonsulf1213	-12.7
3	Domain 13	Sulf13	-7.6
		Nonsulf13	-12.7
4	Inter-domain 13-14	Sulf1314	-9.9
		Nonsulf1314	-13.0
5	Domain 14	Sulf14	-10.1
		Nonsulf14	-13.7

Table 5. GScores of top ligand poses in each fibronectin region from Glide docking studies. The best GScores (most negative free energy of binding) are listed for a fully sulfated hexasaccharide (sulf) and a nonsulfated hexasaccharide (nonsulf) from each region of fibronectin. The ligand poses are shown in Figure 19. GScore = GlideScore from Glide software.

Final Optimized Structure and Simulation

The idea behind the construction of an optimized HS molecule for molecular dynamics simulation of the HS-fibronectin interaction was to select and join together the ligand poses with the best GScores across the five different regions of the fibronectin molecule. These ligands (fully sulfated and nonsulfated) with the best GScores for each region were identified in Table 5 from the Glide docking experiments. While the nonsulfated ligands had the best scores in all five regions, the optimal poses in the inter-domain spaces did not span the gaps and were positioned either towards the bottom of the

protein or behind and crosswise to its main axis (Figure 19B; poses nonsulf1213 and nonsulf1314). These positions were not conducive to forming a contiguous molecule with the other three nonsulfated ligands which were posed more towards the top of the protein (Figure 19B; poses nonsulf12, nonsulf13, and nonsulf14). By contrast, the optimal poses for two of the fully sulfated ligands were positioned directly above the inter-domain spaces and were spatially close to the optimal poses of the three nonsulfated ligands. (Figure 19A; poses sulf1213 and sulf1314). These two fully sulfated ligands were selected with the three nonsulfated ligands to form the hexasaccharide components of an HS molecule with nonsulfated hexasaccharides over the fibronectin domains and fully sulfated hexasaccharides over the inter-domain spaces. Conjoining the five hexasaccharides produced the chain $[G0A0]_3[I2S6]_3[G0A0]_3[I2S6]_3[G0A0]_3$ relative to fibronectin domains 12 to 14. The resulting 30-sugar HS molecule (HS-30) is shown in Figure 20.

The alignment of HS-30 with fibronectin for MD simulation is illustrated in Figure 21. In general, the molecule was positioned in the vicinity of the docking sites associated with each hexasaccharide component. The central oxygen atom of the chain (a glycosidic oxygen atom after sugar 15), denoted by the white dashed circle in Figure 21, was aligned 20 Å above the Phe 124 residue in domain 13 of fibronectin. This procedure was similar to the previous alignments used for the central oxygen in the 36-sugar heparin/HS structures (Hep-36, HS-36, and HS-36mid). The main difference with the HS-30 alignment was that the chain ran across the top of fibronectin from back to front, whereas the other heparin/HS chains ran parallel to the top of fibronectin.

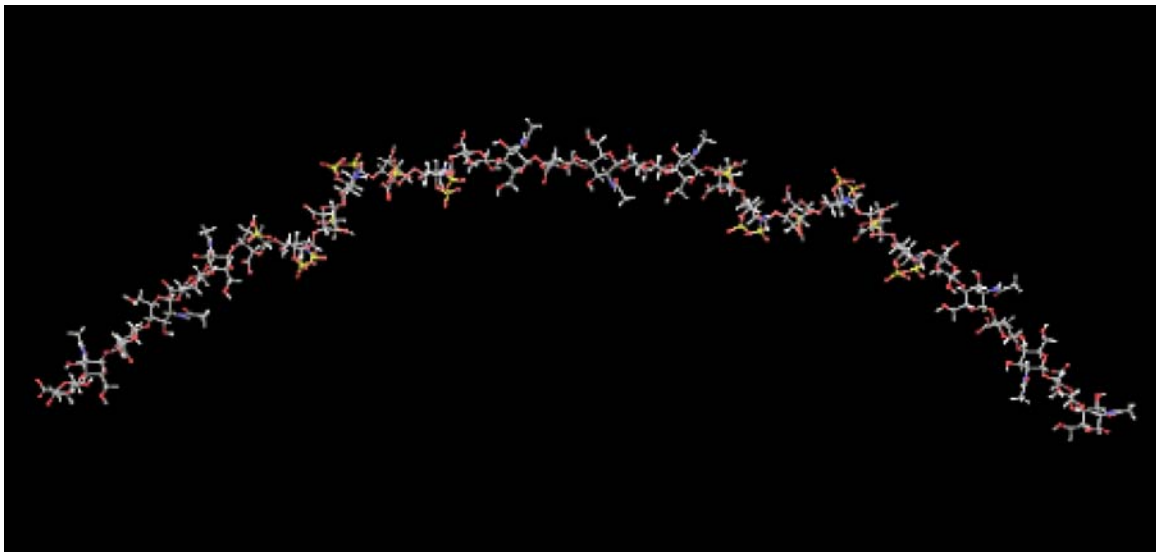


Figure 20. Final optimized HS-30 structure for fibronectin binding. The 30-sugar HS molecule is composed of five optimal hexasaccharide poses from the five regions of fibronectin. The order of the hexasaccharides in the chain (left to right) is $[G0A0]_3[I2S6]_3[G0A0]_3[I2S6]_3[G0A0]_3$, where $[G0A0]_3$ is a nonsulfated hexasaccharide and $[I2S6]_3$ is a fully sulfated hexasaccharide. Atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur.

The 10-ns MD simulation of HS-30 and fibronectin ran for 2.5 months. Unfortunately, the data from the simulation were not completely accessible as a result of technical issues with the software. Although the software indicated that the simulation completed successfully, the trajectory files were corrupted and produced no snapshots of the HS-fibronectin complex. It was possible, however, to extract energy, temperature, and pressure data from the simulation output files, suggesting that the simulation did run to completion even though the trajectory files were not generated properly.

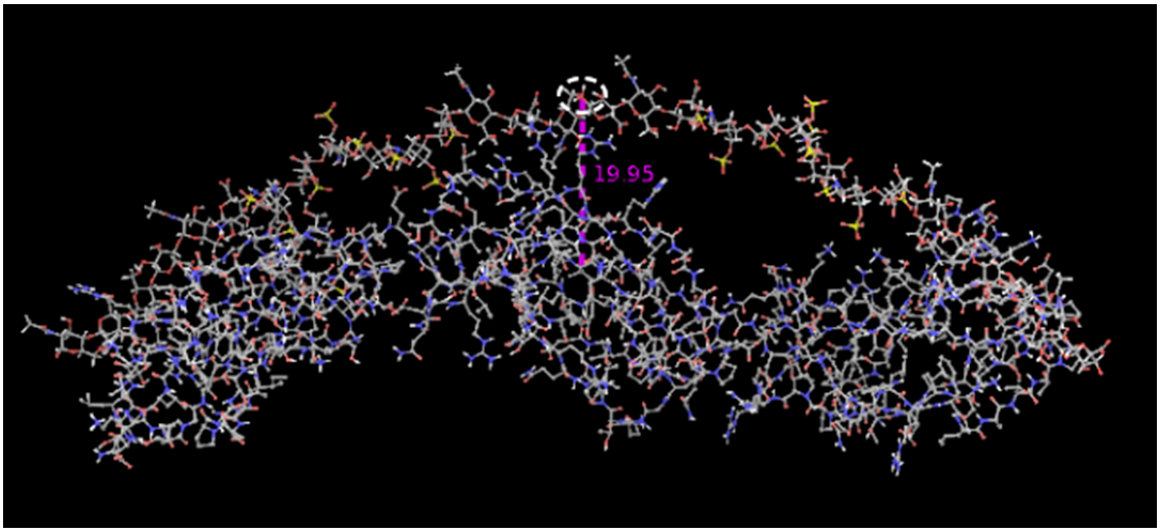


Figure 21. Alignment of HS-30 above fibronectin before MD simulation. The HS-30 molecule $[G0A0]_3[I2S6]_3[G0A0]_3[I2S6]_3[G0A0]_3$ is manually aligned above fibronectin domains 12 to 14 (left to right). The alignment distance (19.95 Å) is shown between the central glycosidic oxygen (white dashed circle) and the α-carbon of Phe 124 (domain 13). Atom colors: blue = nitrogen, gray = carbon, red = oxygen, white = hydrogen, yellow = sulfur

DISCUSSION AND CONCLUSIONS

Molecular Modeling

Results demonstrated that we were able to successfully build heparin and HS molecules of 12 and 24 sugars in length as well as three novel 36-sugar structures with varying sulfation patterns. The seven constructed heparin/HS molecules are presented in Figure 6. The heparin models (Hep-12, Hep-24, and Hep-36) were helical chains with a cluster of three sulfate groups that alternate from side-to-side at regular intervals (about 17 Å) on either side of the axis. Each turn in the helix was composed of approximately four sugars. These structural characteristics mimicked the original 12-sugar conformation that was used as the basis of the construction (Mulloy et al., 1993). The HS models (HS-12, HS-24, HS-36, and HS-36mid) also displayed a helical pattern, but these chains were slightly longer than the heparin structures with the same number of disaccharides. In addition, the HS molecules appeared to curve with increasing chain length. This effect was especially apparent in the 36-sugar HS models, although the middle sulfation of HS-36mid offset the curvature to some extent. The presence of this curvature in the larger HS molecules may reflect the greater degree of rotation that occurs at glucuronic acid-containing disaccharides compared with iduronic acid-containing disaccharides (Conrad, 1998).

Construction of these heparin/HS molecules was not a trivial task, and proper modification of the sugars and glycosidic linkages was a time-consuming process. Nevertheless, we developed a logical and robust methodology that can be followed by other researchers and is expected to be useful in building heparin/HS molecules of any size and pattern. Long-chain heparin and HS structures provide more information about

physiological mechanisms involved in protein interactions, and few studies have used a computational approach to address this issue because of the difficulty in building molecular models. Additionally, the sulfation pattern, as opposed to the sugar sequence of the chain, appears to be of more importance in controlling HS-protein interactions (Kreuger et al., 2006; Spencer et al., 2010). Therefore, by building longer HS chains, we were able to successfully incorporate various sulfation patterns in order to examine the impact of domain organization on fibronectin modulation.

Molecular Dynamics

All Systems

The MD simulations were thermodynamically stable for both the heparin/HS systems and the fibronectin systems. These results are shown in Figure 8 for HS-36mid and in Table A-I for all systems. The temperature and total energy of the system remained stable and did not drift with time during any of the 10-ns simulations. The observed fluctuations were reasonable for large systems (10^5 atoms) and were in accordance with previous studies (Gandhi et al., 2009; Kitchen et al., 1990). The system pressure was not as stable as temperature or energy, but the relatively large fluctuations (±60 atm) were within expectations for NPT (constant number of atoms, constant pressure, and constant temperature)-type simulations (Mikhailov et al., 1997).

The MD simulations also maintained a high level of quality by preventing artifacts from occurring as a result of the periodic boundary conditions invoked by the computational method. A molecule within a box with periodic boundary conditions has the ability to interact with itself if the box is not large enough. The molecule essentially leaves the box on one side and re-enters the box on the opposite side. In order to prevent

this self-interaction, a minimum 10-Å layer of water is recommended between the molecule and the boundary box (de Souza and Ornstein, 1997). A box must therefore be large enough so that it guarantees a minimum approach distance of 20 Å between any part of the molecule and its images in neighboring boxes. Measurements of the minimum approach distance for all MD simulations (Table A-I) indicated that this value was greater than 20 Å, thus ensuring that the solvation boxes were large enough to prevent self-interaction during the simulations.

Heparin and Heparan Sulfate Systems

The results of the MD simulations showed the flexibility in solution of each of the seven heparin/HS molecules (Figure 11). In plots of modified distance (D_m) versus modified angle (θ_m), greater molecular flexibility was inferred from the lower values of θ_m as the ends of the molecule come closer together. The extended shapes of the plots along the *x*-axis (θ_m direction) implied that both heparin and HS were flexible molecules at physiological conditions. However, heparin appeared to be more flexible than HS in solution, particularly at the longer lengths of 24 and 36 sugars. This was confirmed in Table 3, where comparison of heparin and HS showed smaller values of θ_m for heparin at every chain size. For example, the average (±standard deviation) value of θ_m for Hep-36 was 130 (±9) degrees, while the average for HS-36 was 164 (±7) degrees. Trajectory snapshots, such as those in Figure 9 for Hep-24 and HS-24, further supported the greater flexibility of heparin.

Molecular flexibility was also influenced by the sulfation pattern of the HS molecule. This was demonstrated by the greater spread of points for HS-36mid compared with HS-36 (Figure 11B). If Figure 11A and Figure 11B were overlaid, the points of HS-

36mid tracked from the cluster of HS-36 points to the cluster of Hep-36 points. This same trend was seen in Table 3, where the average θ_m of 155 (±15) degrees for HS-36mid was less than the value for HS-36 but greater than the value for Hep-36.

Although the MD simulations clearly depicted a heparin molecule with more flexibility than an HS molecule of similar size, the literature presents a more conflicting view. Some experimental studies show fully sulfated sequences of heparin to be more flexible than nonsulfated sequences of HS (Hricovini et al., 1997; Skidmore et al., 2008), while others show the opposite result (Khan et al., 2013; Mobli et al., 2008).

Fibronectin Systems

The most interesting feature of the MD simulations in the fibronectin systems was the effect of the 36-sugar heparin/HS molecules on the "opening" of the protein. This "opening" was represented by an increase in the size of the joint angle between fibronectin domains 12 and 13 (inter-domain angle 12-13). As shown in Figure 14, the HS molecules had a greater effect on increasing this joint angle than the heparin molecule. While fibronectin alone had a joint angle of 138 (±6) degrees, the presence of Hep-36 had a small effect on the size of the angle, increasing it to 139 (±5) (Table 4). The presence of HS-36 and HS-36mid, however, increased the joint angle to 143 (±4) and 145 (±5) degrees, respectively. The magnitude of this "opening" appeared to be dependent on the sulfation pattern of the HS molecules. This trend suggested that an HS molecule with an optimized sulfation pattern may be able to "open" the fibronectin molecule even further. Although these joint angle increments were small, the corresponding conformational change in the protein may be significant enough to provide new binding sites (Goh et al., 2004; Yu and Koshland, 2001). Longer simulations (100-1000 ns) could

also be useful in amplifying these trends and providing more dramatic changes in the fibronectin conformation (Meller, 2001).

During the 10-ns simulation, all three 36-sugar heparin/HS molecules displayed a dynamic relationship with fibronectin. A comparison of the final snapshots (Figure 14B-D) showed some differences in their interactions with the protein. The heparin molecule was bound mainly at the cationic cradle (cluster of positively charged residues) of domain 13. Even though heparin was the most flexible of the molecules in solution (see Results: Molecular Dynamics), it was held some distance from other regions of the fibronectin surface because of electrostatic interference from its fully sulfated chain. In contrast, the HS molecules were bound much closer to the surface. The HS-36mid molecule appeared to have the tightest association with the surface, being bound at the cationic cradle and following the topology on either side into the adjacent domain. These observations were reflected in the increasing number of non-ionic interactions (hydrogen bonds and van der Waals interactions) from Hep-36 through HS-36 to HS-36mid (Figure 15). Although ionic interactions between the negative charges on the sugar chains and the positive charges on fibronectin were most likely responsible for the global positioning of the heparin/HS molecules, the sulfation pattern of the HS molecules allowed the nonsulfated domains to approach closer to the surface without electrostatic interference and to establish a tighter binding through non-ionic interactions. This enhanced association may be responsible for the effective "opening" of fibronectin by HS molecules. These results again supported the idea that an HS sulfation pattern which mimicked the charged and uncharged topology of fibronectin may be the optimal candidate for "opening" fibronectin and revealing new sites for VEGF binding.

Molecular Docking

Docking simulations with fully sulfated and nonsulfated hexasaccharide probes ran successfully and produced tables of ligands that had the lowest free energy of binding with fibronectin. It was evident through our docking trials that the nonsulfated probes consistently had a better binding affinity with fibronectin than the fully sulfated probes. For each of the five docking regions, the GScores were more negative for the nonsulfated ligands (Table 5). This may reflect the refinements incorporated into the GScore, which in response to underestimation by other scoring functions, included improved scoring of hydrophobic interactions and hydrogen bonds (Friesner et al., 2006).

As such, these results implied that an HS molecule with a specific pattern of nonsulfated regions might have a higher affinity for fibronectin and therefore a greater chance of increasing the fibronectin joint angle. This supports our findings from the molecular dynamics studies in which HS molecules were more successful in "opening" fibronectin. By locating the optimal poses of the fully sulfated and nonsulfated probes on the fibronectin surface (Figure 19), we were able to visualize possible combinations of the probes that would determine a specific sulfation pattern and a specific alignment for an HS molecule on fibronectin. This optimized HS molecule would have high expectations for "opening" fibronectin in molecular dynamics simulations.

Final Optimized Structure and Simulation

Based on the docking results, we created a novel 30-sugar HS molecule (HS-30) with an optimized sulfation pattern $[G0A0]_6[I2S6]_6[G0A0]_6[I2S6]_6[G0A0]_6$ and an optimized placement above the fibronectin surface (Figures 20 and 21). A 10-ns MD simulation of HS-30 and fibronectin was completed after several months of execution;

however, the resulting trajectory files were corrupted even though energy, temperature, and pressure were recorded for the entire run.

The likely source of this problem was traced to a software update for Glide that was installed prior to the docking runs and the final MD simulation. As a result, the setup parameters for the MD simulation in Desmond reverted to default settings which were not adjusted to the normally used values. The run was performed under a different pathway that was unable to properly save the data generated from the simulation. Additionally, the simulation was executed on only 1 CPU when a multiple-CPU configuration was necessary for running the large system. By resetting the pathway and reconfiguring the CPUs, we should be able to successfully rerun this simulation. Also, as a safety measure for these long runs, the simulation should be carried out as a series of short intervals (checkpoint files) to ensure that the system is producing the correct files at each step.

Another relevant factor in the MD simulations was the large size (1.1-1.4 x 10^5 atoms) of the systems in the study. This size represents the upper limit for the operation of Desmond software. Although this issue was not directly responsible for the failure of the simulation with HS-30 and fibronectin, it makes the re-evaluation of our software a high priority, particularly when larger heparin/HS-protein complexes may be in the future. A possible candidate for MD simulations is NAMD (NAnoscale Molecular Dynamics; University of Illinois at Urbana-Champaign, Urbana, IL). NAMD is freely available software designed for parallel computing and capable of simulating very large systems of 10^6 atoms or more (Phillips et al., 2005).

General Summary and Future Work

We have made significant progress in developing computational techniques to investigate the HS-fibronectin interaction at the molecular level. The results of this project demonstrated the following:

(1) Heparin and HS molecules with 12 to 36 sugars were successfully constructed using fully sulfated disaccharides for heparin and variable patterns of fully sulfated and nonsulfated disaccharides for HS.

(2) MD simulations of heparin/HS molecules demonstrated that heparin was more flexible than HS at physiological conditions and that changes in the HS sulfation pattern affected its flexibility.

(3) MD simulations of each heparin/HS molecule interacting with fibronectin (domains 12-14) under physiological conditions showed the ability to detect fibronectin "opening" by monitoring changes in the joint angle between domains 12 and 13 (predicted hidden binding site for VEGF).

(4) MD simulations indicated that HS was more effective than heparin in increasing the fibronectin joint angle and that changes in the sulfation pattern modulated this "opening."

(5) MD simulations suggested that the enhanced capability of HS to "open" fibronectin was the result of a closer association with the fibronectin surface and the generation of a larger number of non-ionic interactions (hydrogen and van der Waals).

(6) Molecular docking studies with fully sulfated and nonsulfated hexasaccharide probes revealed potential sulfation patterns and alignment for an optimized HS

molecule with maximum effect on fibronectin "opening."

(7) A novel 30-sugar HS molecule was designed from molecular docking results for evaluation as an "opener" in MD simulations with fibronectin.

Future work on this project will continue to build on the information from the MD simulations and molecular docking studies and will focus on five main steps:

(1) Re-evaluate the computer system and software and determine the best setup for working with large MD simulations of heparin/HS interactions with fibronectin and other proteins.

(2) Create several optimized HS molecules from molecular docking results (including the 30-sugar HS) and assess their behavior in MD simulations with fibronectin.

(3) Choose one or two HS molecules with the greatest capability to modulate fibronectin "opening" and re-analyze the associated MD simulation for types and number of intermolecular interactions.

(4) Determine the dominant intermolecular interactions during the "opening" of fibronectin and specify the atoms on residues and sugars that play critical roles in this mechanism.

(5) Devise several approaches for disrupting this mechanism and evaluate the effectiveness in laboratory binding experiments between VEGF and fibronectin.

These approaches will provide the innovative basis for the development of new therapeutic options with greater specificity for the treatment of neovascular AMD and the prevention of blindness. They will also validate the use of computational techniques for exploring the often complex interactions between HS and proteins.

BIBLIOGRAPHY

Ambesi A, McKeown-Longo PJ. (2014). Conformational remodeling of the fibronectin matrix selectively regulates VEGF signaling. *J Cell Sci.*, 127(17):3805-3816.

Andersen HC. (1980). Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.*, 72:2384

Banks JL, Beard HS, Cao Y, Cho AE, Damm W, Farid R, Felts AK, Halgren TA, Mainz DT, Maple JR, Murphy R, Philipp DM, Repasky MP, Zhang LY, Berne BJ, Friesner RA, Gallicchio E, Levy RM. (2005). Integrated modeling program, applied chemical theory (IMPACT). *Journal of Computational Chemistry*, 26(16):1752-1780.

Barkalow FJ, Schwarzbauer JE. (1991). Localization of the major heparin-binding site in fibronectin. *J. Biol. Chem.*, 266:7812-7818.

Becker CF, Guimaraes JA, Verli H. (2005) Molecular dynamics and atomic charge calculations in the study of heparin conformation in aqueous solution. *Carbohydrate Research*, 340(8):1499-1507.

Bernfield M, Gotte M, Park PW, Reizes O, Fitzgerald ML, Lincecum J, Zako M. (1999). Functions of cell surface heparan sulfate proteoglycans. *Annu Rev Biochem*, 68:729-777.

Birch DG, Liang FQ. (2007). Age-related macular degeneration: a target for nanotechnology derived medicines. *Int J Nanomedicine*, 2(1):66-77.

Bishop JR, Schuksz M, Esko JD. (2007). Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature*, 446(7139):1030-1037.

Blundell CD, Robert IS, Sheehan JK, Almond . (2009) Investigating the molecular basis for the virulence of *Escherichia coli* K5 by nuclear magnetic resonance analysis of the capsule polysaccharide. *J Mol Microbiol Biotechnol*, 17:71-82.

Brown DM, Regillo CD. (2007). Anti-VEGF agents in the treatment of neovascular age-related macular degeneration: applying clinical trial results to the treatment of everyday patients. *Am J Ophthalmology*,144(4):627-637.

Carpentier M, Denys A, Allain F, Vergoten G. (2014). Molecular docking of heparin oligosaccharides with Hep-II heparin-binding domain of fibronectin reveals an interplay between the different positions of sulfate groups. *Glycoconj J.*, 31(2):161-169.

Chakravarthy U, Evans J, Rosenfield PJ. (2010). Age-related macular degeneration. *BMJ*, 344:c981.

Conrad HE. (1998) *Heparin-Binding Proteins*, Chap 2. Academic Press: San Diego, CA.

Curtis LH, Hammill BG, Qualls LG, Dimartino LD, Wang F, Schulman KA, Cousins

SW. (2012). Treatment patterns for neovascular age-related macular degeneration: analysis of 284 380 Medicare beneficiaries. *Am J Ophthalmology*, 153(6):1116-1124.

de Souza ON, Ornstein RL. (1997). Effect of periodic box size on aqueous molecular dyanmics simulation of a DNA dodecamer with particle-mesh Ewald method. *Biophys J*, 72(6): 2395-2397.

Dulaney SB, Huang X. (2012). Strategies in synthesis of heparin/heparan sulfate oligosaccharides: 2000-present. *Adv. Carbohydr. Chem. Biochem.,* 67:95-136.

Esko JD, Selleck SB. (2002). Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu Rev Biochem*, 71:435-471.

Ferrara N, Gerber HP, LeCouter J. (2003). The biology of VEGF and its receptors. *Nat Med*, 9(6):669-676.

Ferro DR, Provasoli A, Ragazzi M, Torri G, Casu B, Gatti G, Jacquinet J-C, Sinäy P, Petitou M, Choay J. (1986). Evidence of conformational equilibrium of the sulfated L-iduronate residue in heparin and in synthetic heparin mono- and oligosaccharides: NMR and force-field studies. *J Am Chem Soc*, 108:6773-6778.

Friedman D, O'Colmain B, Munoz B, Tomany S, McCarty C, de Jong P, Nemesure B, Mitchell P, Kempen J. (2004). Prevalence of age-related macular degeneration in the United States. *Arch Ophthal*, 122(4):564-572.

Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC and Mainz DT. (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.,* 49:6177-6196.

Gallagher JT. (2001). Heparan sulfate: growth control with a restricted sequence menu. *J Clin Invest*, 108(3):357-361.

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.

Gandhi NS, Mancera RL. (2009) Free energy calculations of glycosaminoglycan-protein interactions. *Glycobiology,* 19(10):1103-1115.

Goerges AL, Nugent MA. (2004). pH regulates vascular endothelial growth factor binding to fibronectin: a mechanism for control of extracellular matrix storage and release. *J Biol Chem*, 279(3):2307-2315.

Goh CS, Milburn D, Gerstein M. (2004). Conformational changes associated with protein-protein interactions. *Curr Opin Struct Biol*, 14(1):104-109.

Hansson T, Oostenbrink C, van Gunsteren WK. (2002). Molecular dynamics simulations. *Current Opinion in Structural Biology,* 12(2): 190-196.

Hricovini M, Guerrini M, Bisio A, Torri G, Petitou M, Casu B. (1997) Conformation of heparin pentasaccharide bound to antithrombin III. *Biochem J.*, 359(2): 265-272.

Jorgensen WL, Chandrasekhar J, Madura JD. (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys,* 79:926-935.

Jorgensen WL. (2004) The many roles of computation in drug discovery. *Science*, 303(5665):1813-1818.

Khan S, Fung KW, Rodriguez E, Patel R, Gor J, Mulloy B, Perkins SJ. (2013) The solution structure of heparan sulfate differs from that of heparin: implications for function. *J Biol chem.*, 288(39)27737-51.

Kitchen DB, Hirata F, Westbrook JD, Levy R, Kofke D, Yarmush M. (1990). Conserving energy during molecular dynamics simulations of water, proteins, and proteins in water. *J Comput Chem* 11(10):1169-1180.

Kreuger J, Spillmann D, Li JP, Lindahl U. (2006). Interactions between heparan sulfate and proteins: the concept of specificity. *J Cell Biol* 174(3):323-327.

Krieger E, Geretti E, 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.

Lawrence R, Lu H, Rosenberg RD, Esko JD, Zhang L. (2008) Disaccharide structure code for easy representation of constituent oligosaccharides from glycosaminoglycans. *Nat Methods*, 5(4):291-291.

Leach AR. (2001). *Molecular Modelling: Principles and Applications*, 2nd ed, Chap 7. Pearson Education Limited: Harlow, England.

Lengauer T, Rarey M. (1996) Computational methods for biomolecular docking. *Curr Opin Struct Biol*, 6(3):402-406.

Liao C, Sitzmann M, Pugliese A, Nicklaus MC. (2011). Software and resources for computational medicinal chemistry. *Future Med Chem*, 3(8):1057-1085.

Linhardt RJ, Liu J. (2012) Synthetic heparin. *Curr Opin Pharmacol.* 12(2):217-219.

Lyon M, Rushton G, Askari JA, Humphries MJ, Gallagher JT. (2000) Elucidation of the structural features of heparan sulfate important for interaction with the hep-2 domain of fibronectin. *Journal of Biological Chemistry*. 275:4599-4606.

Martyna GJ, Tobias DJ, Klein ML. (1994). Constant pressure molecular dynamics algorithms. *J Chem Phys,* 101: 4177.

McNaught AD. (1996) Nomenclature of carbohydrates. *Pure & Appl. Chem.,* 68(10): 1919-2008.

Meller J. (2001). Molecular dynamics. In: *Encyclopedia of Life Sciences*, John Wiley & Sons Ltd: Chichester, UK.

Mikhailov D, Linhardt RJ, Mayo KH. (1997) NMR solution conformation of heparin-derived hexasaccharide, *PNAS*, 328:51-61.

Mintz-Hittner HA, Kennedy KA, Chuang AZ. (2011). Efficacy of intravitreal bevacizumab for stage 3+ retinopathy of prematurity. *N Engl J Med*, 364:603-615.

Mitsi M, Forsten-Williams K, Gopalakrishnan M, Nugent MA. (2008). A catalytic role of heparin within the extracellular matrix. *J Biol Chem*, 283(50):34796-34807.

Mitsi M, Hong Z, Costello CE, Nugent MA. (2006). Heparin-mediated conformational changes in fibronectin expose vascular endothelial growth factor binding sites. *Biochemistry*, 45(34):10319-10328.

Mulloy B, Forster MJ. (2000) Conformation and dynamics of heparin and heparan sulfate. *Glycobioloy*, 10(11): 1147-1156.

Nicholson BP, Schachat AP. (2010). A review of clinical trials of anti-VEGF agents for diabetic retinopathy. *Graefes Arch Clin Exp Ophthalmol*, 248(7):915-930.

Pankov R, Yamada K. (2002) Fibronectin at a glance. *J Cell Sci*, 115: 3861-3863.

Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K. (2005). Scalable molecular dynamics with NAMD. *J Comput Chem* 26(16):1781-1802.

Rahimi N. (2006). Vascular endothelial growth factor receptors: molecular mechanisms of activation and therapeutic potentials. *Exp Eye Res*, 83(5):1005-1016.

Ricard-Blum S, Féraud O, Lortat-Jacob H, Rencurosi A, Fukai N, Dkhissi F, Vittet D, Imberty A, Olsen BR, van der Rest M. (2004). Characterization of endostatin binding to heparin and heparan sulfate by surface plasmon resonance and molecular modeling. *J Biol Chem*, 279(4):2927-2936.

Robinson CJ, Mulloy B, Gallagher JT, Stringer SE. (2006). $VEGF_{165}$-binding sites within heparan sulfate encompass two highly sulfated domains and can be liberated by K5 lyase. *J Biol Chem*, 281(3):1731-1740.

Sandwall E, Bodevin S, Nasser NJ, Nevo E, Avivi A, Vlodavsky I, Li JP. (2009).

Molecular structure of heparan sulfate from Spalax. Implications of heparanase and hypoxia. *J Biol Chem*, 284(6):3814-3822.

Sapay N, Cabannes E, Petitou M, Imberty A. (2011) Molecular modeling of the interaction between heparan sulfate and cellular growth factors: bringing pieces together. *Glycobiology,* 21(9):1181-1193.

Schlingemann RO. (2004). Role of growth factors and the wound healing response in age-related macular degeneration. *Graefes Arch Clin Exp Ophthalmol*, 242(1):91-101.

Schmidt-Erfurth U, Eldem B, Guymer R, Korobelnik J, Schlingemann RO, Axer-Siegel R, Wiedemann P, Simader C, Gekkieva M, Weichselberger A. (2011). Efficacy and safety of monthly versus quarterly ranibizumab treatment in neovascular age-related macular degeneration: the EXCITE study. *Ophthalmology*, 118(5):831-839.

Sharma A, Askari JA, Humphries MJ, Jones EY, Stuart, DI. (1999) Crystal structure of heparin- and integrin-binding segment of human fibronectin. *EMBO J,* 18(6):1468-1479.

Skidmore MA, Guimond SE, Rudd TR, Fernig DG, Turnbull JE, Yates EA. (2008). The activities of heparan sulfate and its analogue heparin are dictated by biosynthesis, sequence, and conformation, *Connect Tissue Res* 49:140-144.

Spencer, JL, Bernanke JA, Buczek-Thomas JA, Nugent MA. (2010) A computational approach for deciphering the organization of glycosaminoglycans. *PLoS One*., 5(2):e9389

Spencer JL, Stone PJ, Hugent MA. (2006) New insights into the inhibition of human neutrophil elastase by heparin. *Biochemistry*, 45: 9104-9120.

Stortz CA, Johnson GP, French AD, Csonka GI. (2009) Comparison of different force fields for the study of disaccharides. *Carbohydrate Research*, 344: 2217-2228.

Stringer SE, Forster MJ, Mulloy B, Bishop CR, Graham GJ, Gallagher JT. (2002). Characterization of the binding site on heparan sulfate for macrophage inflammatory protein 1α. *Blood*, 100(5):1543-1550.

Verli H, Guimarães JA. (2005). Insights into the induced fit mechanism in antithrombin-heparin interaction using molecular dynamics simulations. *J Mol Graphics Modell*, 24:203-212.

Verli H, Guimaraes JA. (2004) Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution. *Carbohydrate Research*, 339(2):281-290.

Wang Y, Zheng Q-C, Kong C-P, Tian Y, Zhan J, Zhang J-L, Zhang H-X. (2015). Heparin makes differences: a molecular dynamics simulation study on the human βII-tryptase monomer. *Mol BioSyst*, 11:252-261.

Wijelath ES, Rahman S, Namekata M, Murray J, Nishimura T, Mostafavi-Pour Z, Patel

Y, Suda Y, Humphries MJ, Sobel M. (2006). Heparin-II domain of fibronectin is a vascular endothelial growth factor-binding domain: enhancement of VEGF biological activity by a singular growth factor/matrix protein synergism. *Circ Res.*, 99(8):853-860.

Yu EW, Koshland DE, Jr. (2001). Propagating conformational changes over long (and short) distances in proteins. PNAS, 98(17):9517-9520.

Zarbin MA. (2004). Current concepts in the pathogenesis of age-related macular degeneration. *Arch Ophthalmol*, 122(4):598-614.

Zulueta MML, Lin S-Y, Hu Y-P, Hung S-C. (2013). Synthetic heparin and heparan sulfate oligosaccharides and their protein interactions. *Curr Opin Chem Biol.* 17:1023-1029.

Appendix I

Structures in Simulation	Time (ns)	Box Size (Å)	Number of Atoms (10^4)	Minimum Approach Distance (Å)	Total Energy (10^5 kcal/mol)	Temp (K)	Pressure (atm)
Heparin-12	10.0	70x70x70	11.59	51	-3.139 (0.002)	310.0 (0.9)	4.34 (63)
Heparin-24	10.0	120x80x80	11.52	31	-3.165 (0.002)	310.0 (0.9)	3.70 (63)
Heparin-36	10.0	180x90x90	13.93	33	-3.860 (0.003)	310.0 (0.9)	5.11 (59)
Heparan Sulfate-12	10.0	70x70x70	11.59	60	-3.130 (0.003)	310.0 (0.9)	5.28 (63)
Heparan Sulfate-24	10.0	120x80x80	11.56	30	-3.120 (0.003)	310.0 (0.9)	5.34 (63)
Heparan Sulfate-36	10.0	180x90x90	14.03	32	-3.770 (0.003)	310.0 (0.8)	6.25 (58)
Heparan Sulfate-36 with middle sulfation	10.0	180x90x90	14.01	21	-3.800 (0.004)	310.0 (0.8)	5.37 (59)
Heparin-36 + Fibronectin	10.0	180x100x100	13.62	26	-4.960 (0.002)	310.0 (0.9)	4.94 (54)
Heparian Sulfate-36 + Fibronectin	10.0	180x100x100	13.59	28	-3.700 (0.004)	310.0 (0.8)	5.18 (63)
Heparan Sulfate-36 with middle sulfation + Fibronectin	10.0	180x100x100	13.77	22	-3.71 (0.006)	310.0 (0.8)	4.17 (63)
Fibronectin	10.0	150x80x100	11.49	NA	-3.04 (0.002)	310.0 (0.9)	4.00 (70)

Table A-I. Summary of molecular dynamics simulations. System parameters are given for each of the heparin/HS molecules in isolation, and the 36-sugar molecules interacting with fibronectin. Values for total energy, temperature and pressure are listed as average (standard deviation).

Ligand	GScore (kcal/mol)	Run Time	Grid Size (Å)
Sulf12	-8.7	3hr 30min	25x25x25
Nonsulf12	-10.6	10hr 30min	25x25x25
Sulf13	-7.6	1hr 30min	25x25x25
Nonsulf13	-12.7	11hr 10min	25x25x25
Sulf14	-10.1	5hr 21min	25x25x25
Nonsulf14	-13.7	10hr 50min	25x25x25
Sulf1213	-7.7	4hr 48min	25x25x25
Nonsulf1213	-12.7	11hr 56min	25x25x25
Sulf1314	-9.9	4hr 30min	25x25x25
Nonsulf1314	-13.0	13hr 9min	25x25x25

Table B-I. Summary of docking simulations. The prefix "sulf" represents the fully sulfated hexasaccharide probe, while the prefix "nonsulf" represents the nonsulfated probe. The numerical values following the prefix imply the fibronectin domain over which the probe was aligned for that simulation. The GScore presented is the best (most negative free energy) for each docking simulation.

Appendix II

Abbreviations

1FNH: Protein Data Bank code for Fibronectin

1HPN: Protein Data Bank code for Heparin

AMD: Age-related Macular Degeneration

BUSM: Boston University School of Medicine

CPU: Central Processing Unit

ECM: Extracellular Matrix

GAG: Glycosaminoglycan

GPU: Graphics Processing Unit

GUI: Graphic User Interface

Hep: Heparin

Hep-X: Heparin molecule comprised of X-number of sugars

HS: Heparan Sulfate

HS-X: Heparan Sulfate molecule comprised of X-number of sugars

NECO: New England College of Optometry

PDB: Protein Data Bank

RPE: Retinal Pigment Epithelium

VEGF: Vascular Endothelial Growth Factor

VMD: Visual Molecular Dynamics

Appendix III

Additional Figures

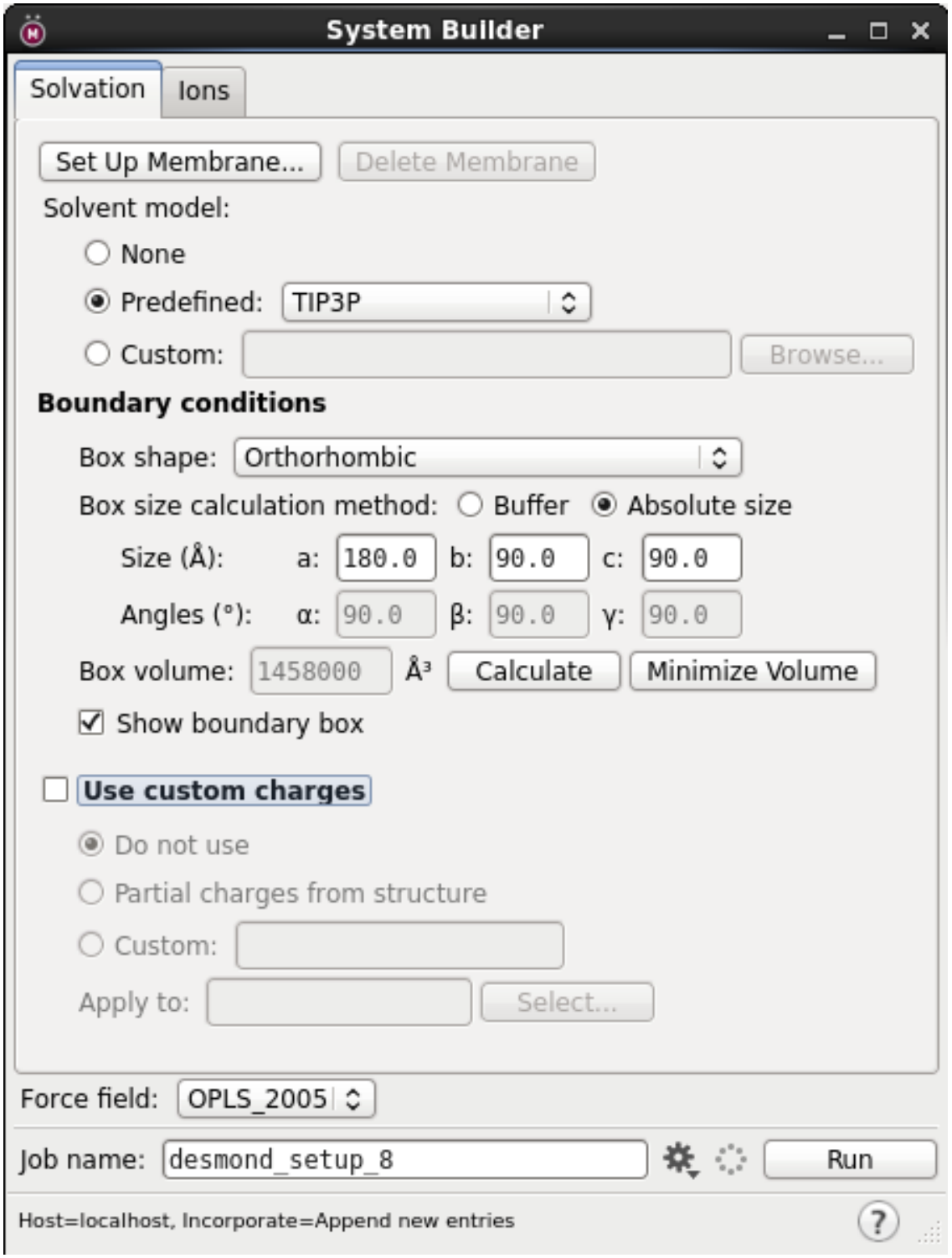


Figure A-III. Molecular dynamics setup screen from Schrodinger's Maestro software. Items selected are those that we used for each of the 36-sugar structures.

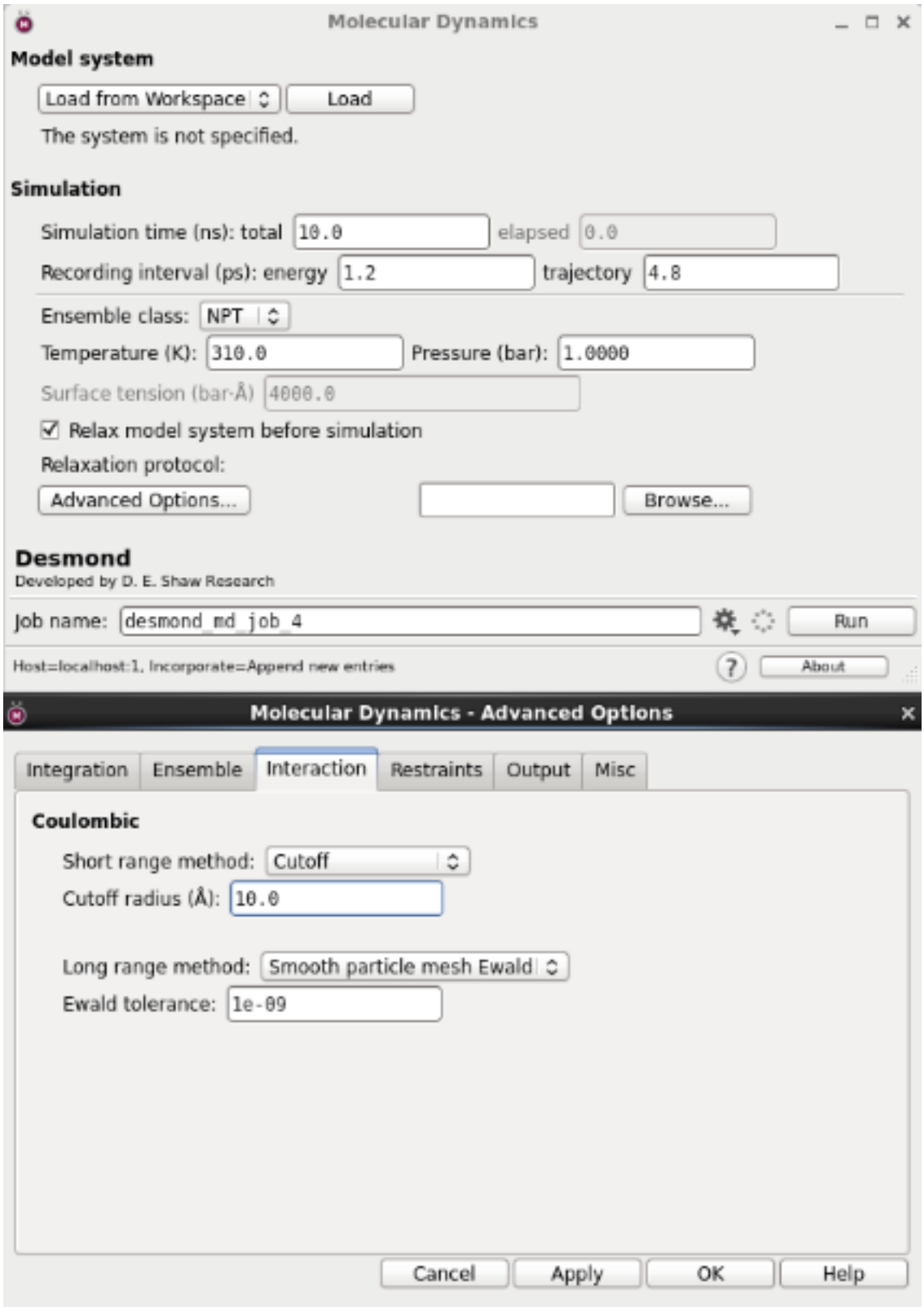


Figure B-III. The molecular dynamics (MD) input screen above depicts the input parameters used for all of our MD simulations.

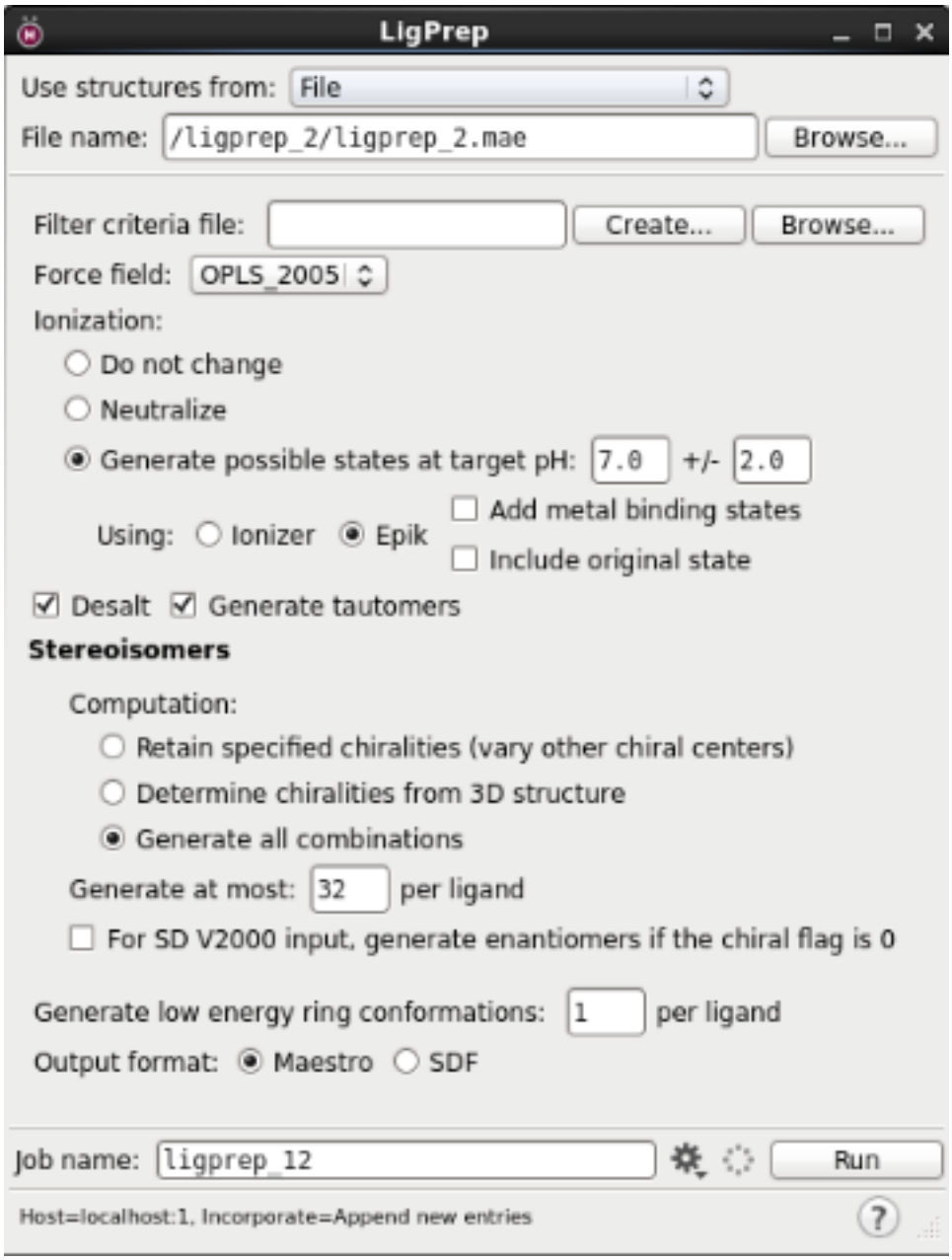


Figure C-III. The LigPrep input screen displays the input parameters for each of the hexasaccharide probes used in the docking studies.

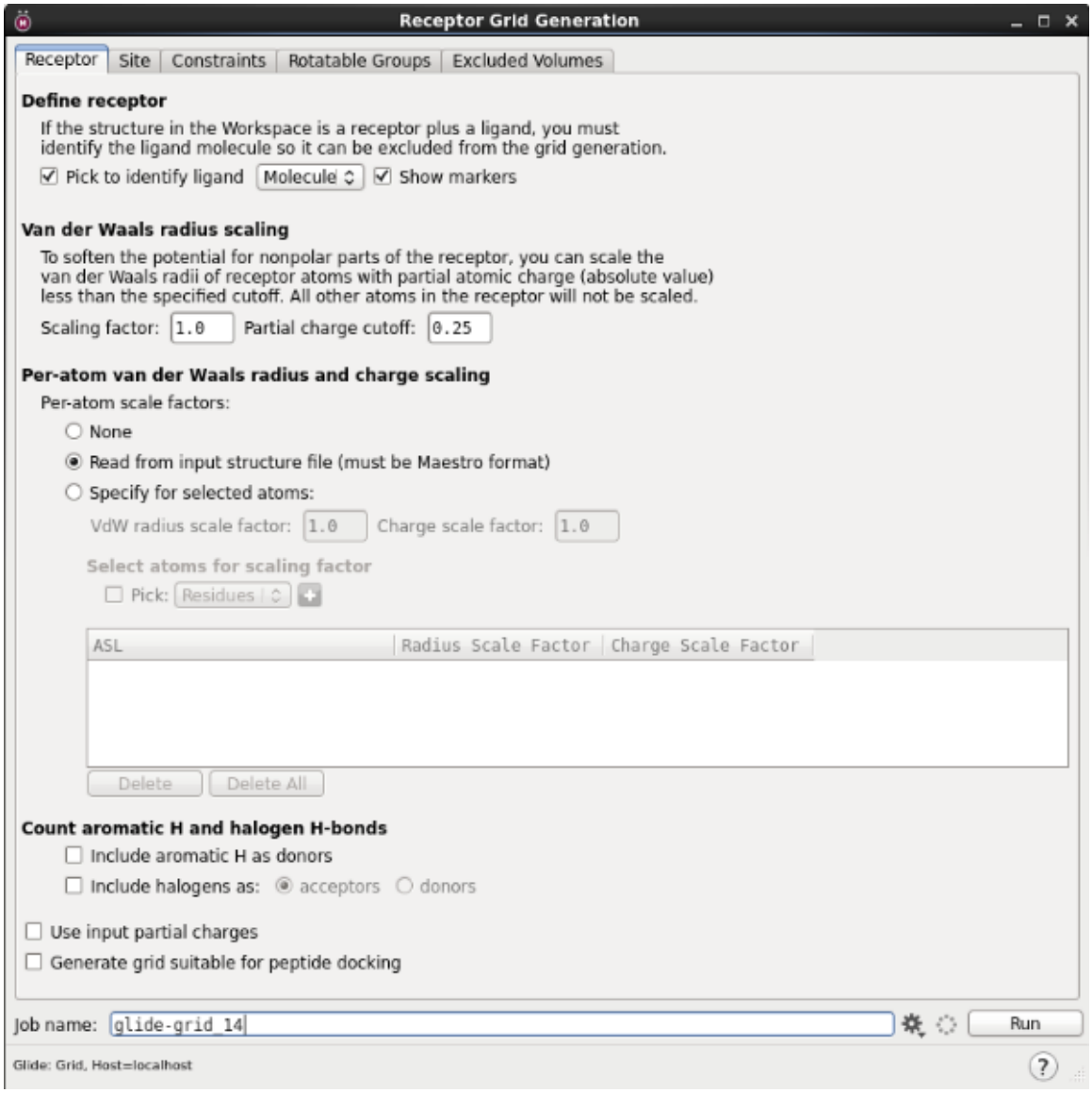


Figure D-III. A grid must be created before docking studies can be performed. The figure above shows the input screen for the grid generated for each of our docking studies in Glide.

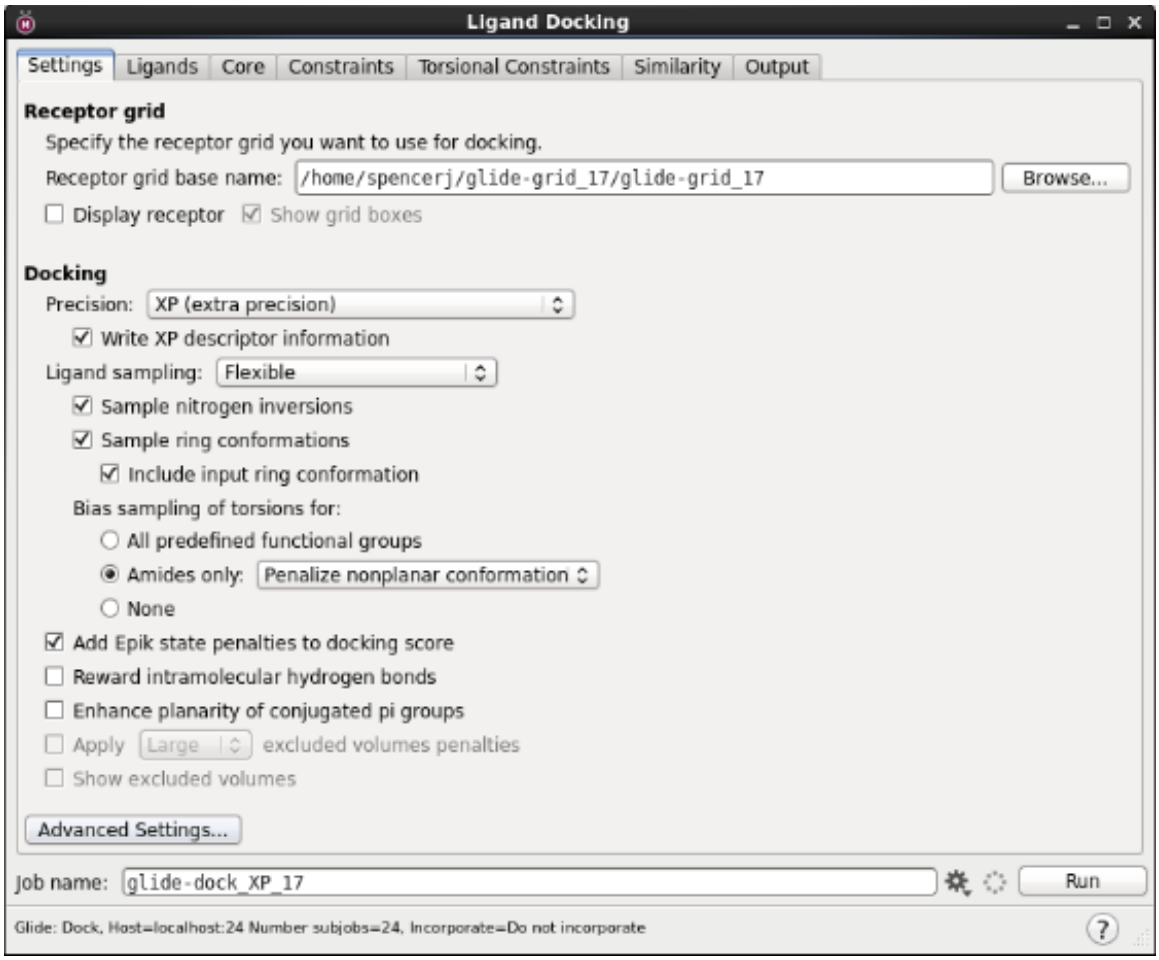


Figure E-III. The Glide input screen is shown above with appropriate parameters used for each of the docking simulations.