A Molecular Dynamics Study of the Vascular Endothelial Glycocalyx Layer



Maria Pikoula

Pembroke College

University of Oxford

Submitted to the Department of Engineering Science, University of Oxford, in fulfilment of the requirements for the degree of

Doctor of Philosophy

Hilary 2016

Dedication

To Eleftheria, Apostolos and Kyriakos: three mathematicians that taught me a lot more than maths.

Abstract

The luminal surface of endothelial cells which line the vasculature is coated with a layer of membrane-bound macromolecules of a mixed carbohydrate and protein nature, collectively described as a glycocalyx, from the greek meaning "sweet husk/covering". Experiments have consistently revealed the pivotal role of the endothelial glycocalyx layer in vasoregulation and the layer's contribution to mechanotransduction pathways. However, the exact mechanism by which the glycocalyx mediates and interprets fluid shear stress remains elusive. This study employs atomic-scale molecular simulation with the aim of investigating the conformational and orientation properties of the highly flexible components of the glycocalyx and their suitability as transduction molecules under hydrodynamic loading. To this aim, two molecular dynamics systems were constructed. The first system focused on the impact of flow on a tethered, branched, oligosaccharide. Fluid flow was shown to only moderately affect the conformation populations explored by the oligosaccharide, in comparison to static conditions. On the other hand, the glycan exhibited a significant orientation change, when compared to simple diffusion, aligning itself with the flow direction. The tethered end of the glycan, an asparagine amino-acid, experienced conformational changes as a result of this flow-induced bias. Results of the "glycan in flow" model suggest that shear flow through the layer can have an impact on the conformational properties of saccharide-decorated transmembrane proteins, thus probably acting as a mechano-transducer. The second system consisted of charged and non-charged heparan sulfate, a component found in large quantities in the endothelial glycocalyx layer. Systems of paired heparan sulfate strands were investigated under conditions of increasing proximity, which is the expected effect of compression of the layer under the effect of flow; a process explored formally using the adaptive biasing force method. This approach warranted the implementation of enhanced sampling for the high energy states of heparan sulfate in close proximity. Results of the heparan sulfate model suggest that areas of locally high charge density within the glycocalyx, generally areas of high sulfation, are characteristically more resistant to compression than non-sulfated areas. The sulfation mix therefore emerges as an important determinant of glycocalyx mechanical properties.

Acknowledgements

I would like to express my gratitude to my supervisor, Prof Yiannis Ventikos for his direction and unyielding support throughout the duration of my work on this thesis. This work could never have been accomplished without his motivation and guidance. I would also like to thank Dr Matthew B. Tessier as well as Prof Robert J. Woods, who welcomed me in their research group at the University of Georgia in Athens GA. I would never have been able to grasp the intricacies of complex carbohydrates without their help and I am very grateful for having collaborated with this group and for having worked closely with Dr Tessier in particular. I would also like to thank Prof Forbes Dewey Jr., without whom many aspects of this thesis would not have been possible, for his invaluable insights during my visit at the Massachusetts Institute of Technology. I also very much appreciate the time Dr Benedict Sattelle took from his work to retrieve a few of his old analysis scripts and share them with me. Thank you also to Prof Steven Payne for taking me on at such a late stage of my research and helping me go the final stretch. Last but not least, many thanks to all members of the old Fluidics and Bio-Complexity research group, and in particular Himanshu Kaul, who put up with me the most. Finally, this thesis would not be possible without the funding from the following, for which I am deeply grateful:

- RCUK Digital Economy Programme
- The Alexander S. Onassis Public Benefit Foundation
- The State Scholarships Foundation of Greece (IKY)

Contents

1 Introduction **1**
- 1.1 Motivation . . . 1
- 1.2 Overview of the clinical problem . . . 2
- 1.3 The endothelial glycocalyx . . . 3
- 1.4 Clinical applications . . . 4
- 1.5 The role of atomistic models . . . 7

2 Vascular Physiology and the EGL **10**
- 2.1 Vascular Physiology and Atherogenesis . . . 10
  - 2.1.1 Normal and Pathogenic Vascular Morphology . . . 11
  - 2.1.2 Mechanotransduction . . . 14
  - 2.1.3 Systemic Risk Factors . . . 21
- 2.2 Glycocalyx Overview . . . 24
  - 2.2.1 Early Findings . . . 24
  - 2.2.2 The EGL Structure . . . 25
  - 2.2.3 EGL properties . . . 26
  - 2.2.4 Physiological Function . . . 30
- 2.3 Proteoglycans and glycosaminoglycans . . . 33
  - 2.3.1 Structure . . . 34
  - 2.3.2 Ring Flexing . . . 35
  - 2.3.3 Heparan Sulfate Turnover in ECs . . . 38
  - 2.3.4 Properties and Vasoregulatory Function . . . 39
- 2.4 The multi-scale relevance of the EGL . . . 40

3 EGL Modelling **43**
- 3.1 Continuum and Coarse-grained Models . . . 43
  - 3.1.1 EGL Effects on Permeability and Adhesion . . . 43
  - 3.1.2 Mechanosensing models . . . 45
  - 3.1.3 EGL modelled as a polymer brush . . . 46
- 3.2 Atomic-scale Models . . . 48
- 3.3 Flow in Molecular Simulation . . . 52
  - 3.3.1 Flow Implementation methods . . . 52
  - 3.3.2 Applications to Biological Systems . . . 53
- 3.4 Feasibility of complete atomic scale models . . . 55
- 3.5 Overarching aims of the thesis . . . 55

4 Molecular Dynamics Modelling **58**
4.1 Classical Molecular Dynamics . . . 58
4.1.1 Force Fields . . . 59
4.1.2 Statistical Mechanics . . . 61
4.1.3 Computational Considerations . . . 63
4.2 Implementation Specifics . . . 68
4.2.1 NAMD . . . 69
4.2.2 The Glycam Force Field . . . 70
4.2.3 Integrator and time step . . . 70
4.2.4 Supporting Tools . . . 71
4.2.5 Supercomputing resources . . . 72
4.3 Work Flow . . . 72
4.4 Validation . . . 74
4.4.1 RDS and coordination number . . . 75
4.4.2 Torsion angles and Ramachandran plots . . . 77
4.4.3 Nuclear Overhauser effect distances . . . 78
4.5 The Adaptive Biasing Force . . . 79
4.5.1 The reaction coordinate . . . 79
4.5.2 Thermodynamic Integration for free energy calculations . . . 80
4.5.3 Sampling using the ABF method . . . 80
4.6 MD modelling of the EGL . . . 82

5 Case study: Tethered Glycan Under Physiological Flow **83**
5.1 Flow in the EGL . . . 83
5.2 System Set-up . . . 84
5.3 Results and Discussion . . . 87
5.3.1 Ion Hydration . . . 88
5.3.2 Calculation of imposed flow . . . 91
5.3.3 Effect of the tethered glycan on the water velocity profile . . . 93
5.3.4 Flow-induced orientation and conformational changes on the glycan . 93
5.3.5 Effect of flow on the amino acid . . . 100
5.4 Summary of findings . . . 104

6 Case study: GAG interactions in close proximity **106**
6.1 Charge density in the EGL . . . 106
6.2 System Set-up . . . 109
6.2.1 Monosaccharide simulations . . . 109
6.2.2 Tetrasaccharide simulations . . . 113
6.3 Prediction of tetramer PMF through monomer interaction . . . 117
6.3.1 PMF model . . . 117
6.3.2 Electrostatic model . . . 118
6.4 Results and Discussion . . . 119
6.4.1 Conformation analysis . . . 119
6.4.2 Simulation derived PMF curves . . . 133
6.5 Summary of findings . . . 141

7 Discussion on the two MD models **143**
7.1 Glycan in flow and glycan as a charged polymer brush . . . 143
7.2 Aims of this thesis . . . 145
7.2.1 Tethered Glycan Under Physiological Flow . . . 145
7.2.2 GAG interactions in close proximity . . . 148
7.3 Limitations . . . 151
7.3.1 Tethered Glycan Under Physiological Flow . . . 151
7.3.2 GAG interactions in close proximity . . . 151

8 Conclusions and future direction of MD EGL simulations **153**
8.1 Glycan in flow simulations . . . 153
8.1.1 Conclusions of glycan in flow model . . . 154
8.1.2 Future direction for flow simulations . . . 155
8.2 GAG simulations . . . 157
8.2.1 Conclusions on the GAG simulations . . . 157
8.2.2 Direction of GAG simulations . . . 158
8.3 Future work on EGL models . . . 159
Bibliography . . . 161

List of Acronyms

ABF: Adaptive biasing force
CoM: Centre of mass
CS: Chondroitin sulfate
DPD: Dissipative particle dynamics
DS: Dermatan sulfate
EC: Endothelial cell
ECM: Extra-cellular matrix
EGL: Endothelial glycocalyx layer
eNOS: endothelial nitric oxyde synthase
GAG: Glycosaminoglycan
HA: Hyaluronic acid
HDL: High density lipoprotein
HS: Heparan sulfate
HUVEC: Human umbilical vein endothelial cell
JNK: c-Jun N-terminal kinase
LDL: Low-density lipoprotein
MD: Molecular dynamics
NMR: Nuclear magnetic resonance
NOE: Nuclear Overhauser effect
NO: Nitric oxide
O2O: Oxygen-to-oxygen
PDB: Protein crystallographic database
PME: Particle mesh Ewald
PMF: Potential of mean force
RBC: Red blood cell
TI: Thermodynamic integration
VLDL: Very low density lipoprotein
VSMC: Vascular smooth muscle cell
WBC: White Blood Cell
WSS: Wall shear stress
μ-*PIV*: Micro-scale particle image velocimetry

Chapter 1

Introduction

1.1 Motivation

In the last decade, evidence-based practice has started to gain ground in the medical field. Apart from traditional scientific methods applied to clinical problems, such as *in vivo* and *in vitro* experiments, computational modelling is emerging as a promising technique in support of the evidence-led approach.

Several key factors come together to allow the application of personalised medicine, which allowed for interventions and treatments to be tailored to the physiology of the patient. In cardiovascular and cerebrovascular disease, the combination of:

- "state of the art" computational models of cells and tissue behaviour that emerge from multi-scale studies, coupled with
- advanced imaging techniques that allow for the implementation of patient-specific biochemical parameters and vasculature geometry

can lead to better etiological understanding, prognosis and treatment [1].

The focus of this project is in the field of computational modelling of the vascular lumen, starting right at the molecular level, where biochemical pathways of the cell are activated by mechanical stimuli in the atomistic scale and vice versa. Ultimately,

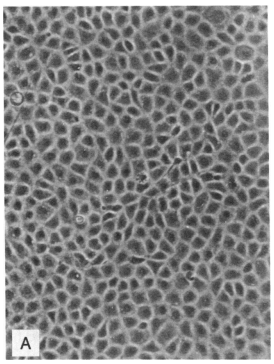
the main ambition driving this work is to understand and quantify these processes and enable the design of innovational interventional technology, translatable to clinical applications.

1.2 Overview of the clinical problem

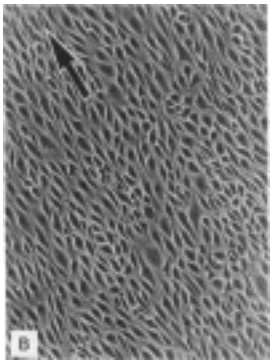
Atherosclerosis, the principal cause of heart attack and stroke, is the world's number one cause of death and disability [2]. Complications of atherosclerosis include peripheral vascular disease and peripheral arterial occlusive disease which manifests widely in the lower extremities ("smoker's leg"). Although it is well known to be lipid-driven, inflammation and blood flow have generated substantial interest as alternative or complementary explanations for the formation of atherosclerotic lesions (plaques) [3].

Other conditions related to the vasculature and blood flow regulation and abnormalities include aneurysms, diabetes, systemic inflammatory response syndrome and hypertension [4] [5] [6]. Hypertension in particular can be viewed as both cause and outcome of several of the above diseases, since blood pressure is affected by them but also plays a pivotal role in their evolution.

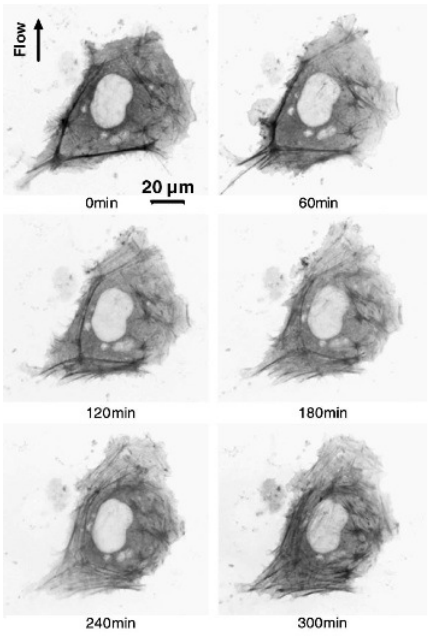
It is believed that central to the development of these conditions, is the function of a monolayer of cells that covers the vascular lumen, the endothelium. Endothelial cells (ECs) are influenced by frictional stress, the wall shear stress (WSS), generated by the blood. In vitro, ECs change from a raised polygonal to a flattened elongated morphology (Figure 1.1) under the influence of flow, with the degree of elongation correlated to the magnitude of the shear stress [7] [8]. It is widely accepted that, through this mechanism, ECs are responsible for, among other functions, the sensing of haemodynamic WSS on the luminal side and as such they are central mediators of vasoregulation [9].



(a) *Endothelial cells in static culture*



(b) *Endothelial cells alignment with flow*



(c) *Endothelial Cell elongation due to shear stress*

Figure 1.1: ***Endothelial Cell elongation due to shear stress:*** *(a) and (b) Morphological changes induced in confluent bovine aortic endothelial cells by exposure to shear stress in (b) laminar flow (arrow points out the flow direction) in comparison with static culture (a). [8] (c) Dynamic change in actin filament structure of a single endothelial cell exposed to fluid shear stress up to 300 min. [10]*

1.3 The endothelial glycocalyx

It is now accepted that the luminal surface of ECs which line the vasculature is coated with a layer of membrane-bound macromolecules [11] of a mixed carbohydrate and protein nature collectively described as a glycocalyx, from the greek

meaning "sweet husk/covering", given its polysaccharide structure. This endothelial glycocalyx layer (EGL) has come a long way since its discovery by means of electron microscopy in the 1960s. It is now believed to have unique properties that allow it to play a central role at the interface between ECs and the blood flow. Strategically placed, the EGL is a negatively charged molecular bush-like structure, bound on the endothelial cell membrane, connected and interacting with the cell's cytoskeleton. Furthermore, it exists in full interaction with the blood plasma, in combination with which it forms a hydrated active structure, also referred to as an endothelial surface layer (ESL) [12].

Three key functions are attributed to the EGL. It is considered a permeability modulator in the transcapillary exchange of water, creating local osmotic potential that is very different to the bulk of plasma in the capillary. Evidence also suggests that it plays a part in the regulation of red and white blood cell interactions with the vascular lumen, most notably during inflammatory response. Finally, it is thought to plat a vital role in the mechanism of WSS sensing by the ECs. Initial mechanosensing by the primary sensory structure is followed by signal transduction to the cytoskeleton of the cell, inducing biochemical responses as part of its self-regulation process [13]. The present work mainly focuses on the understanding of the latter role of the EGL as a mechanosensor.

1.4 Clinical applications

Several pathological conditions correlate with damage of the EGL, with effects that range between partial deterioration to complete destruction of the glycocalyx structure. Reperfusion injury[1] occurs in cases of complete or partial ischemia treatment which can cause near-complete glycocalyx degradation [12]. Inflammation

[1]Reperfusion injury is the tissue damage caused when blood supply returns to the tissue after a period of ischemia or lack of oxygen (anoxia, hypoxia).

and trauma have also been shown to correlate with glycocalyx shedding (a partial or full degradation of glycocalyx components) a process which triggers further inflammation, in a destructive feedback process. Atherosclerotic disease in arterial vessels correlates with diminished EGL thickness for reasons not well understood so far [14]. Diabetes has also been shown to correlate with glycocalyx thickness reduction and increased vascular permeability [15].

In most types of clinical intervention for the treatment of cardiovascular and cerebrovascular conditions, either drugs or implants (or in most cases both) come into contact with the vascular endothelium. It is imperative to study primarily the function of a healthy EGL and subsequently the way in which such therapeutics interact with ECs, in order to better understand their function and ultimately, to design them in a way that encourages the viability and proliferation of ECs and their delicate glycocalyx layer.

Early diagnosis and chemical protection

Signs of early EGL damage can be detected via blood sample analysis. Abnormal levels of concentration of vital EGL components are an indication of EGL shedding and therefore can be indicative of early atherosclerosis stages. Once damage is detected, preservation or, in cases of severe damage, restoration strategies, could play a key part in treatment of the conditions that negatively effect the integrity of the EGL.

A simple way to achieve protection is with high concentration of plasma proteins, most notably by use of albumin. Hydrocortisone and antithrombin have also exhibited protective effects against reperfusion injury [16]. Avoidance strategies suggest the use of antioxidants and avoidance of hypervolaemia (increase in the volume of circulating blood). Protection by means of supplying "pre-fabricated" glycocalyx components to compensate for glygocalyx deterioration is an interesting concept to investigate but has not been conclusively proven to have any restoring and/or

protective effect.

Coronary stents

At present, the deployment of intravascular stents has become a common and widely used minimally invasive treatment for coronary heart disease. Following the deployment of a coronary stent and disruption of an atheromatous plaque, the deformation of the arterial wall and the presence of the stent struts create a new fluid dynamic field, which can cause an abnormal biological response. To improve procedures of coronary re-vascularisation/re-endothelisation (e.g. reduce in-stent restenosis rates) the optimal strategy lies in further development of stent design, material and coatings.

Cerebral open stents/flow diverters

Endovascular treatment of cerebral aneurysms using stents has advanced markedly in recent years. Originally the stent was designed to act like a scaffold across the aneurysm neck while using coiling as the endovascular treatment [17]. More recently however the stent is also being used as a stand-alone device. It is believed that placement of a stent across the neck of an aneurysm alters the haemodynamics of the blood flow in and around the aneurysm region. Ohta et al. suggest that intracranial stents reduce the shear stress on the aneurysmal wall and increase viscosity within the aneurysm sac, both of which are believed to reduce the risk of rupture and increase the potential of aneurysm thrombosis [18].

However, the main concerns, particularly for this technique, are endothelisation and intimal growth over the stented segment, particularly at the aneurysm - parent vessel interface. Stent devices are thrombogenic as they stimulate platelet aggregation as soon as they are exposed to blood and so neointimal proliferation occurring within the stented segment could result in stenosis of the parent artery [19].

Further understanding of EGL function and ways to successfully preserve its structure could lead to better protection strategies during stent design and deploy-

ment. Such strategies can be targeted to improve the relation between the stent and the affected area to which it is implanted, as well as the preservation of physiological function in adjacent areas which are also affected by the presence of the implant.

1.5 The role of atomistic models

The focus of this work is the endothelial glycocalyx in the atomic scale using molecular dynamics (MD). It is our aim to demonstrate that investigation of particular aspects of the glycocalyx, coupled with advances in glycomics and MD methods will lead to better and highly detailed atomic scale models. In turn, such models will enable the development of robust coarse grained models grounded firmly in realistic atomistic representations. The next stage is the inclusion of coarse grained model derived parameters within already existing as well as new continuum and computational fluid dynamics models of endothelial cells in flow, complete with a realistic glycocalyx layer. This rationale is illustrated in Figure 1.2.

The contribution of this thesis to detailed atomistic models of the glycocalyx is twofold. In the first MD model, the importance of the interaction of physiological blood flow with the glycocalyx is investigated through adequately converged, all atom simulations. This is the first time realistic, physiological flow is implemented in an all-atom model and for fully converged simulations.

In the second atomistic mode, the focus is on the electrostatic nature of the glycocalyx and the effect of its compression as a result of blood flow as well as other factors. The dynamics of highly and poorly negatively charged glycocalyx regions in close proximity is investigated by constructing detailed all-atom models and implementing a novel enhanced sampling method in order to fully explore possible relative orientations of closely interacting EGL components.

A detailed presentation of the aims and objectives of the present work is given following the literature review at the end of Chapter 3. Subsequent chapters are

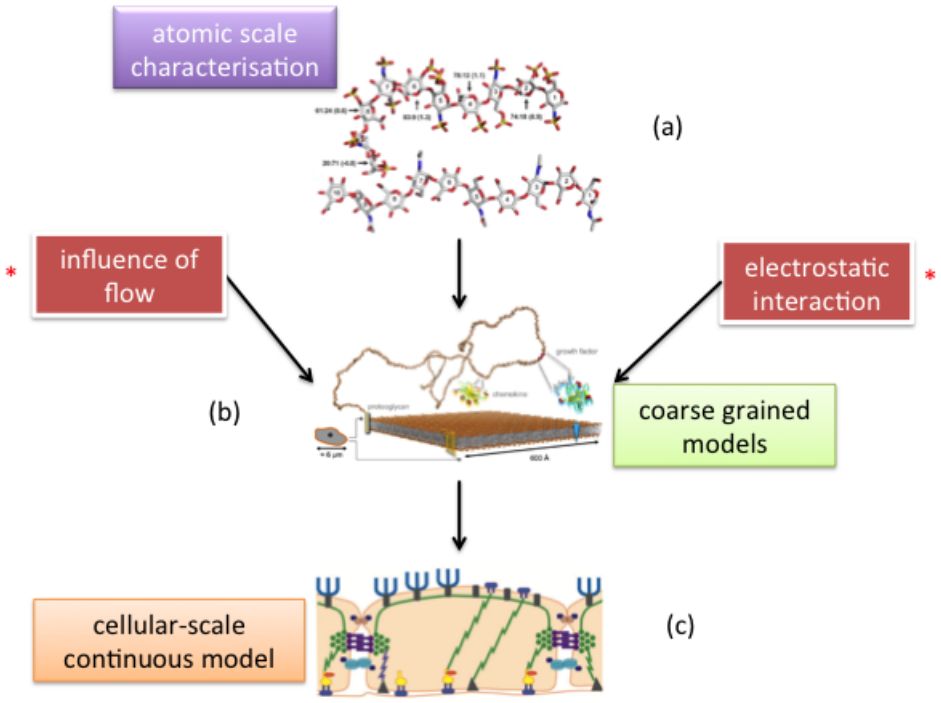


Figure 1.2: ***Situating of the work in this thesis in the wider context of endothelial cell glycocalyx modelling:*** *Asterisks denote modelling aspects covered in this work. Schematics by (a), (b) Sattelle et al. [20] and (c) Weinbaum et al. [13].*

organised as follows:

- **Chapter 2:** Description of vascular physiology with a particular focus on the endothelial glycocalyx layer and its constituent glycosaminoglycans as well as the importance and proposed hypotheses on mechanisms that govern mechanosensing in ECs.

- **Chapter 3:** Approaches to modelling of the endothelium and endothelial glycocalyx in particular; Atomic scale and continuous models; Flow implementation in atomistic models; Aims and objectives of this work.

- **Chapter 4:** MD theory and methodological implications for the present work; Overview of the computational algorithms and methods used in the following two case studies presented in chapters 5 and 6.

- **Chapter 5:** Description of a novel MD endothelial glycocalyx model inclusive of physiological flow implemented in the present work; Presentation of results and validation.

- **Chapter 6:** Description of an MD glycocalyx model for the study of closely interacting charged glycosaminoglycans, the building blocks of the endothelial glycocalyx; Presentation of results and validation.

- **Chapter 7:** Overview and discussion on both models presented in the results chapters and their contribution with respect to the original objectives as well as their limitations.

- **Chapter 8:** Conclusions derived from Chapters 5, 6 and 7; future directions of present models and their contribution to glycocalyx modelling.

Chapter 2

Vascular Physiology and the EGL

In this chapter we discuss the healthy and the diseased vasculature phenotype and the progression of atherosclerosis. The focus shifts from a broad overview to specific systemic and localised risk factors as well as protective structures, primarily in the endothelium. We examine the endothelial cells and in particular their glycoprotein-rich coating which forms the EGL and its role in transduction of mechanical signals to the endothelial cell. We conclude with a detailed discussion on the glygocalyx constituent components and their function in the layer and the endothelium.

2.1 Vascular Physiology and Atherogenesis

The mammalian circulatory system is well-tuned to provide each cell with the adequate amount of oxygen and nutrients required for the homeostasis of the organism. Disturbances to this complex self-regulation can take the form of either systemic factors or localised biomechanical events, that span a wide time-scale, ranging from milliseconds to years. It becomes therefore evident that the progression of vascular disease is a challenging problem to study. The exact effects of both systemic and localised factors, as well as their interplay, are not fully understood, although mechanisms of synergy have been proposed. In this section we provide an overview of the current state of the art in the field.

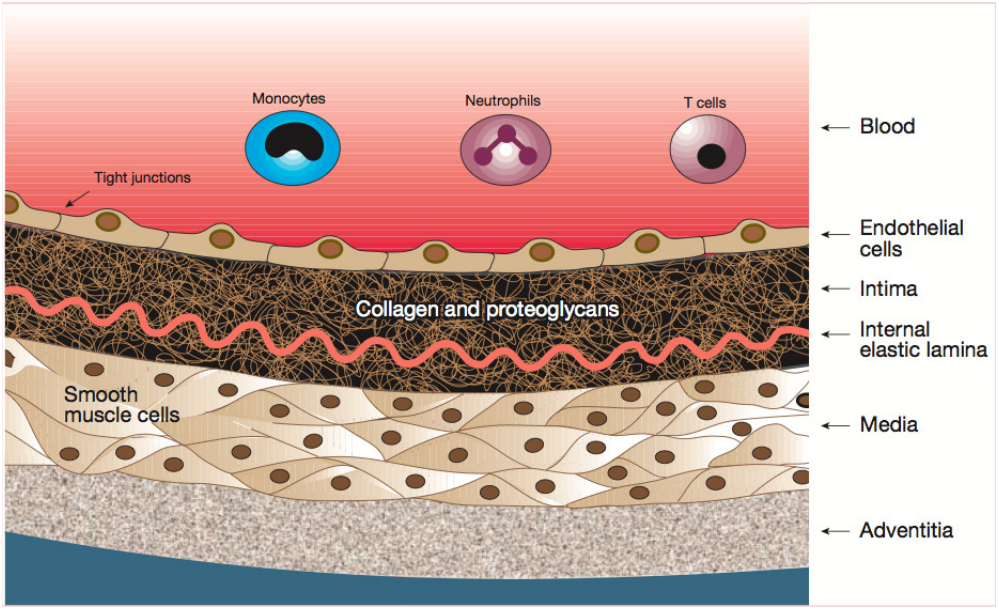


Figure 2.1: ***The layered structure of a healthy, large artery.*** *The intima layer size is exaggerated for clarity. [22]*

2.1.1 Normal and Pathogenic Vascular Morphology

The structure of arteries consists of three morphologically distinct layers: the *intima*, *media* and *adventitia* (see schematic in Figure 2.1). The intima is the thin, inner-most layer, made-up of extracellular matrix components such as collagen, elastin and proteoglycans. It is bound on the luminal side by the endothelial surface monolayer, an ultra-thin film composed of ECs. Vascular smooth muscle cells (VSMCs) and most of the elastin make up the media. The elastin is arranged in fenestrated sheets (lamellae) between which are longitudinally oriented collagen fibres, thin layers of proteoglycan-rich extracellular matrix (ECM), and VSMCs. Thin elastic fibres connect the lamellae into a three-dimensional continuous network and connect the lamellae with the VSMCs [21]. The adventitia is the outer-most layer and consists of connective tissue alongside VSMCs and fibroblasts.

The morphological and functional changes that occur in the pathology of atherosclerosis characterise the progressive nature of the disease. The stages of atherogenesis have been well-documented in animal models such as non-human primates, pigs, rabbits and rodents. In particular, mice deficient in apolipoprotein E or the low-density

lipoprotein (LDL) receptor are the models most used in genetic and physiological studies, as they develop advanced lesions [23].

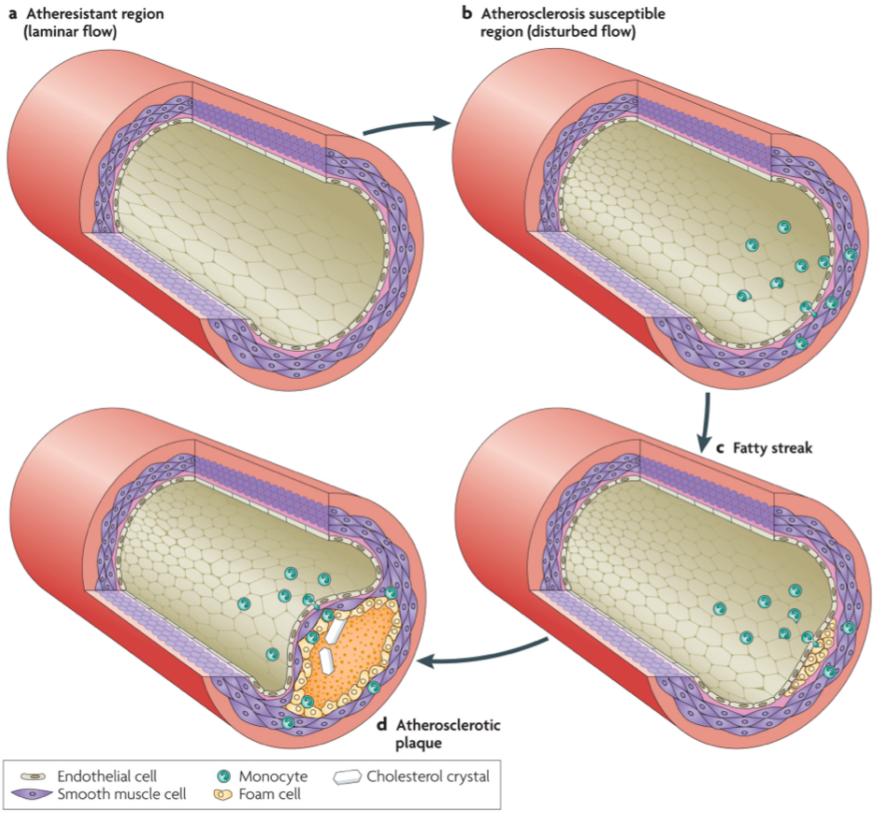


Figure 2.2: ***Progression of atherosclerosis in arteries.*** *(a) Healthy morphology in atheroresistant regions, (b) Region susceptible to atherosclerosis, (c) Formation of fatty streaks and (d) Atherosclerotic plaque [24].*

A schematic representation of the events leading to atherogenesis can be seen in Figure 2.2. In healthy arteries, ECs are aligned with the flow direction (as seen in Figure 2.3) and exhibit low death and proliferation rates [25]. Athero-promoting factors, such as adhesion receptors and leukocyte attractive cytokines maintain low expression levels (part (a) of Figure 2.2). As we will discuss further, such regions coincide with profiles of laminar, undisturbed flow. In contrast to this phenotype, in atherosclerosis-prone regions, ECs exhibit higher rates of turnover and higher leukocyte adhesion receptor expression rates, resulting in the binding of leukocytes and monocytes mainly, to the endothelium and also migrating to the vessel wall (intima layer), as seen in part (b) of Figure 2.2.

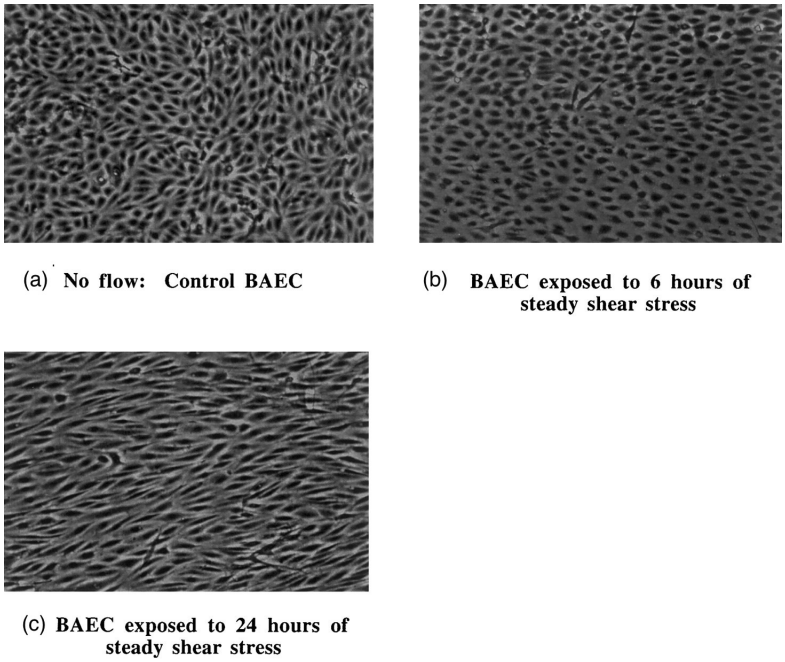


Figure 2.3: ***Flow-induced endothelial elongation.*** *Micrographs show control (no flow) bovine aortic endothelial cells (BAEC) or BAEC exposed in a parallel-plate flow chamber to a steady laminar shear stress of* 19 dyn/cm^2 *for* 6 *or* 24h. *BAEC monolayers were viewed through windows in the flow chamber. Cells were photographed using a cooled charge-coupled device camera (Photometrics SenSys) [26].*

Atherosclerosis does not arise uniformly on arterial surfaces throughout the vasculature but tends to be localised to specific sites, e.g. branches and curvature segments. In such regions, fluid mechanic factors deviate from their normal paterns. It is therefore accepted that flow patterns characterised as "disturbed", highly correlate with the location of vascular regions which are shown to be prone to atherogenesis [27]. In a subsequent stage of the disease, monocytes begin to proliferate, differentiate into macrophages and take up the lipoproteins, forming foam cells [28] (Figure 2.2, part (c)). Such factors can penetrate into and cause thickening of the VSMC-containing media layer. These so-called fatty streaks appear in athero-prone regions in individuals across all ages. Eventually, small fatty-streaks can grow to larger, more inflamed regions known as atherosclerotic plaques, as seen in part (d) of Figure 2.2. Their composition varies substantially, but they typically consist of cholesterol crystals, dead cells and foam cells. The rate of progression towards this more severe stage of the disease is governed by a multitude of factors that are

examined in following sections.

VSMCs can also migrate into the plaque from the underlying media. The death of VSMCs, foam cells and macrophages all contribute to a necrotic core, one of the defining characteristics of a vulnerable plaque [29]. Overtime, plaques can become increasingly complex, developing regions of calcification, ulceration at the luminal surface, and haemorrhage from small vessels that grow into the lesion from the media of the vessel. Advanced lesions can gradually grow to the point of blocking blood flow. However, the most important clinical complication is an acute occlusion due to the formation of a thrombus or blood clot, resulting in myocardial infarction or stroke. The thrombosis is often associated with rupture or erosion of the lesion [22].

2.1.2 Mechanotransduction

Mechanosensing is an essential physiological function in daily life as it is keeps the organism informed as to the environment and protecting it from danger. Mechanosensing is involved in two of the five senses: hearing (audition) and touch (tactition). The function of somatosensory mechanosensing originates at the cellular level.

All cells exist within a 3-dimensional environment which imposes on them myriads of mechanical stimuli. Cells sense both endogenous forces mainly generated from cytoskeletal contractility and exogenous ones originating from the extracellular environment. Early single-cell organisms developed the ability to sense, transmit and respond to mechanical forces, which eventually evolved to the sophisticated mechanisms, such as stretch activated ion-channels [30] [31] and integrin heterodimers [32], by which cells in multicellular organisms communicate and interact with the extracellular mechanical environment. Cells in multicellular tissues are subject to a variety of forces (Figure 2.5), including tensile, compressive, wall shear stress (WSS), the frictional force per unit area, and hydrostatic pressure. These physical cues are mainly transduced into biochemical signals, and cells respond to the cues in the form of morphological changes, migration, or replicative senescence through activat-

ing signalling pathways as well as changes in gene expression.

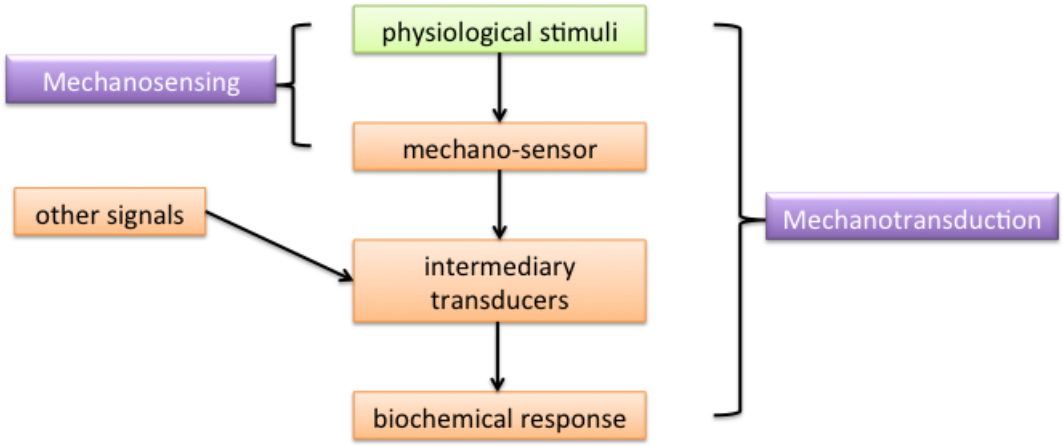


Figure 2.4: ***Mechanotransduction stages in cells.*** *Schematic diagram of the mechanotransduction process in cells and the role of mechanosensing as the primary event.*

This process, known as mechanotransduction, consists of several stages, as seen in Figure 2.4. In this work, the focus is on the mechanosensing stage that occurs upstream of the transduction into biochemical signals and downstream of the external mechanical stimuli. Response to these forces varies depending on cell physical properties, the adjacent cells and the structure and composition of the ECM in their immediate surroundings.

The ECM, being the principal extracellular component of tissues, provides a supporting scaffold to the cells and also plays an important role in the processes of adhesion, migration, apoptosis, proliferation and differentiation. Its molecular components include collagen, elastin, proteoglycans, fibronectin and laminin. Dependent on its composition in various tissues, it ranges from a hard solid structure (such as for bones and teeth) to more elastic structures, such as cartilage and tendons [33].

Cells associated with the vasculature respond to mechanical forces and modulate the artery diameter to accommodate the blood flow and provide cells across all tissues with vital blood perfusion [35]. Blood pressure causes circumferential stretching normal to the artery wall, which is sensed by both ECs, on the luminal surface and

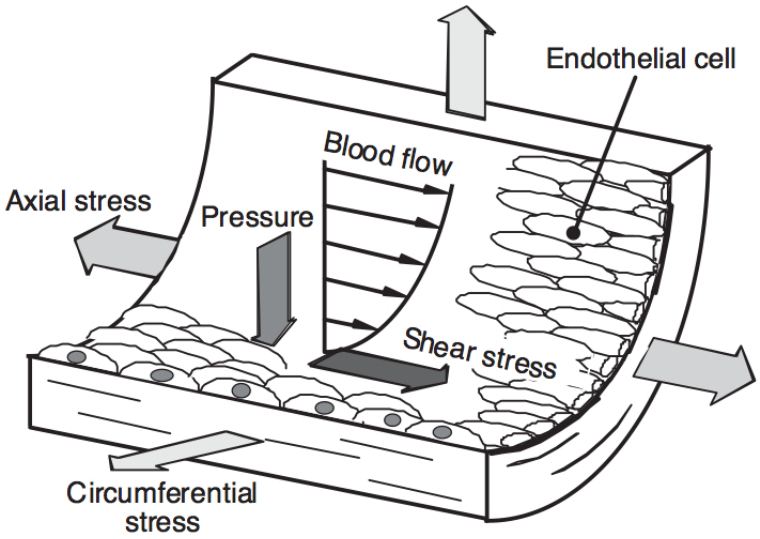


Figure 2.5: ***Forces acting on Endothelial Cells.*** *Schematic diagram showing the generation of shear stress parallel (tangential to the endothelial cell surface) by blood flow and the generation of normal stress (perpendicular to the endothelial cell surface) and circumferential stretch due to the action of pressure [34]*

VSMCs embedded in the vascular wall. VSMCs respond to blood pressure related stress by remodelling and increasing the vessel diameter during periods of elevated pressure. In pathological conditions of constantly elevated blood pressure, such as hypertension, this remodelling can be irreversible, resulting in the vessel wall losing its elasticity and therefore its flexibility to adapt in pressure changes [36].

ECs also respond to blood pressure changes. However, the most important factor affecting ECs is likely WSS. In healthy arteries, ECs respond to WSS magnitude and regulate the relaxation of the arterial smooth muscle by producing nitric oxide (NO), prostacyclin and other substances [38], as well as releasing K^+ cations through membrane channels [39]. An overview of the so far understood complex network of cellular responses to flow for both ECs and VSMSs is given in Figure 2.6.

Multiple pathways are involved in the shear stress signal transduction that leads to alterations in EC morphology and function and to activation of various transcription factors (Figure 2.7). Which pathways are primary and which are secondary remains unclear and the initial sensor that recognises shear stress has not been identified. It is likely that shear stress activates several pathways simultaneously [41].

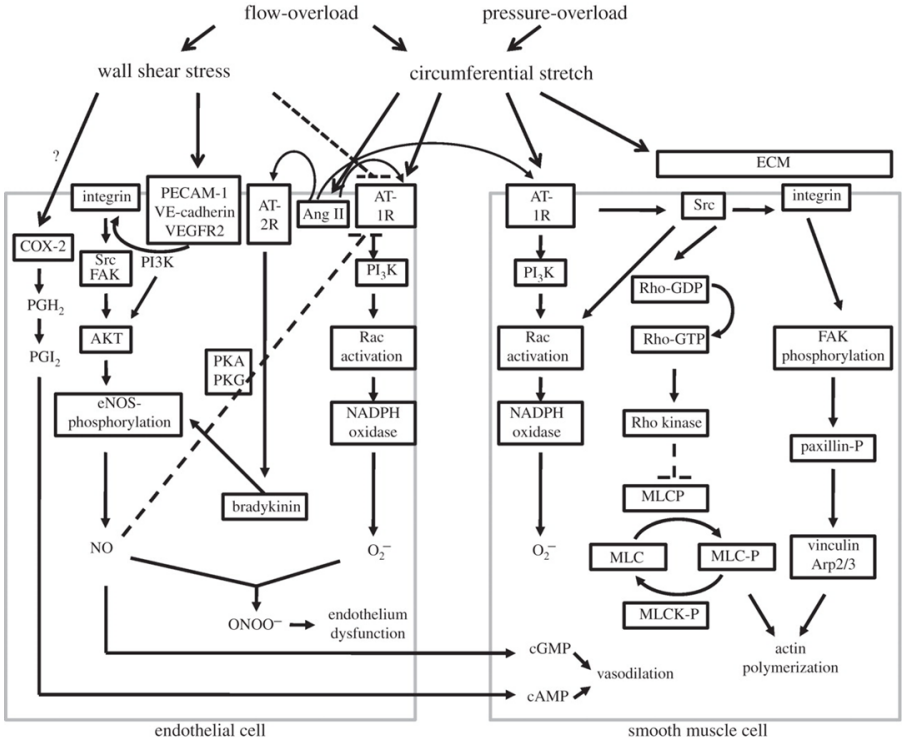


Figure 2.6: ***A schematic diagram of cellular responses from the flow and pressure overload.*** *The next level biochemical pathways induced by WSS and circumferential stretch are depicted in the graph. Two cell types, endothelial and smooth muscle cells are included in the pathway. In the graph, solid arrows indicate activation while dash blunts indicate negative regulations. Schematic by Lu et al. [37]*

Mechanosensing candidates

Although ECs have been shown to be responsible for most of the mechanical sensing of forces attributed to flow, the exact mechanisms and functional components that convert the physical stresses to biochemical responses remain unknown. Several mechanosensing candidates have been proposed (Figure 2.8), and although they have all been shown to contribute in the process, it is challenging to assess the hierarchy and interplay of events that ultimately lead to mechanotransduction.

Cytoskeleton Despite not being considered as a primary and direct mechanosensor, the cytoskeleton plays an important part in the propagation of mechanical cues through its microtubules, actin and intermediate filaments which physically connect different regions of the EC, transmitting signals from the apical domain where shear is applied [42]. Inhibition, by drugs or genetic methods, of actin, microtubules or

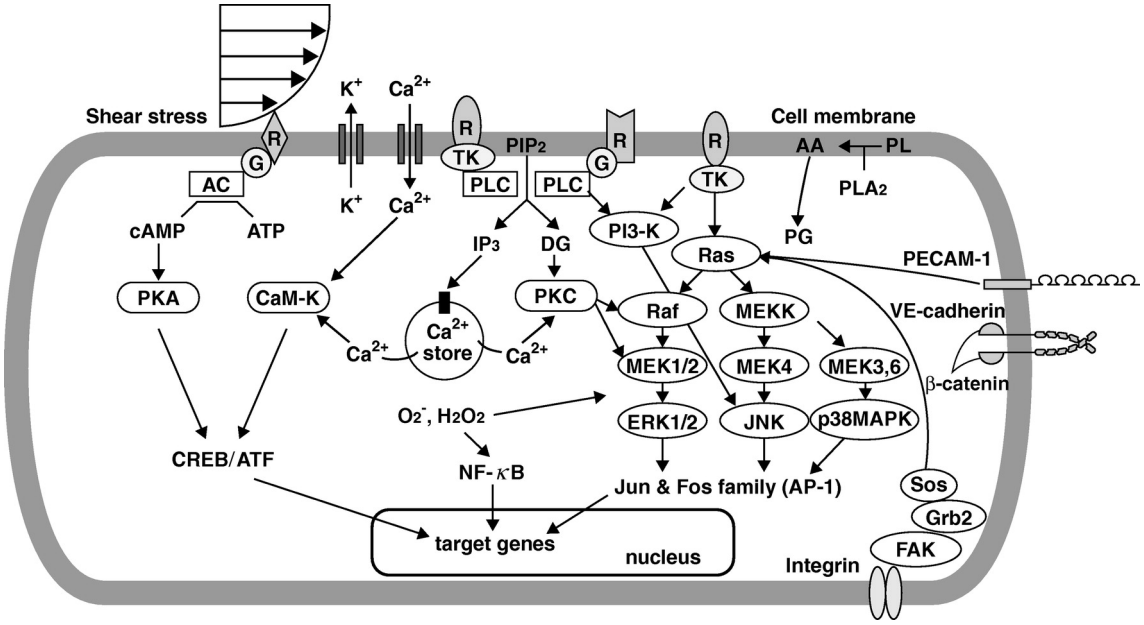


Figure 2.7: ***Signal transduction factors that have been shown to be activated by shear stress:*** *AA: arachidonic acid; AC: adenylate cyclase; ATP: adenosine 5'-triphosphate; CaM-K: calmodulin kinase; cAMP: cyclic adenosine 3',5'-monophosphate; CREB/ATP: cAMP-responsive element-binding protein/activating transcription factor; DG: 1,2-diacylglycerol; ERK: extracellular-signal-regulated kinase; FAK: focal adhesion kinase; G: G protein; IP3: inositol,4,5-triphosphate; JNK: c-jun N-terminal kinase; MEK: MAP kinase-ERK kinase; MEKK: MAP kinase-ERK kinase kinase; NFκB, nuclear factor kappa B. Jun, Fos family (AP-1): transcription factors; p38 MAPK: mitogen-activated kinase; O2$^-$ and H2O2: reactive oxygen species; PECAM-1: platelet endothelial cell adhesion molecule-1; PG: prostaglandin; PI3-K: phosphatidylinositol 3-phosphate; PKA: protein kinase A; PKC: protein kinase C; PL: phospholipids; PLA2: phospholipase A2; PLC: phospholipase C; Ras and Raf: small G proteins; Sos and Grb2, adaptor proteins. R: receptor; TK: tyrosine kinase. Schematic by Ando et al. [40]*

intermediate filaments blocks many ECs responses to flow [43], [44], [45].

Adhesion receptors Integrins are connective proteins between the cell cytoskeleton and the ECM. It has been shown that they are crucial in sensing of shear, as their elimination by integrin antibodies or other perturbations inhibit EC responses to flow [46], [47], [48], [49]. There is, however, little evidence to support a direct role for integrins as mechanosensors, in terms of primary sensing of shear. Although it is possible to activate integrins by altering their conformation under tension, the forces exerted by flow would not be sufficient to do so, as they are 2 to 3 orders of magnitude weaker than typical traction forces between a cell and its substrate [50].

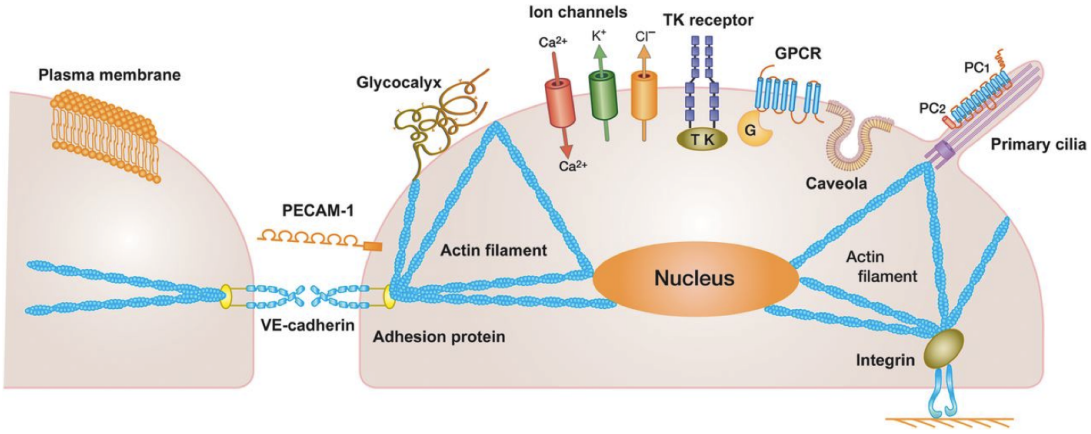


Figure 2.8: ***Schematic representation of shear sensor candidates.*** *TK: Tyrosine Kinase, GPCR: G-protein coupled receptor (by Ando et al. [57]) Not in scale.*

Upstream of integrin activation, force transduction mechanisms have been investigated. One example is a protein complex comprising of PECAM-1, VE-cadherin and the transmembrane Tyrosin Kinase vascular endothelial growth factor receptor 2 (VEGFR2) [51]. Evidence suggests that activation of this complex takes place in direct response to flow.

Luminal membrane proteins The luminal side of the cell is directly exposed to flow and thus shear, and naturally several membrane proteins have been investigated for their role as potential direct mechanosensors. One such example is the G-protein and G-protein coupled receptor. Flow and increased membrane fluidity have been shown to activate G-proteins [52]. It is, however, unclear whether G-protein is a direct mechanosensor (ligand dependent) or indirect (ligant independent), since findings from *in vitro* [53] and *in vivo* mice studies [54] are conflicting.

Changes in membrane fluidity or tension can also trigger signalling through mechanosensitive ion channels under flow [55]. Their activation is inhibited by high plasma membrane cholesterol levels, suggesting a mechanism by which plasma cholesterol can interfere with vasoregulation [56].

Primary cilia A cilium is an organelle found in eukaryotic cells. Cilia are thick protuberances that project from the much larger cell body. Primary cilia associated

proteins, which mediate flow sensing in the kidney, are expressed on ECs [58]. However, despite the evidence of vascular disfunction in mice and humans with genetic mutations that inhibit the expression of such proteins [59], the presence of primary cilia in arteries is rare, as they disassemble in response to flow [60]. It is therefore unlikely that primary cilia play a key role in sensing of normal levels of shear, although it is possible that they are implicated in low-shear sensing in sites susceptible to atherosclerosis [61].

Glycocalyx The last decade has seen major advances in our understanding of the critical role played by the layer of glycosaminoglycans (GAGs) that reside on the surface of vascular endothelium. The layer, called the glycocalyx, represents the key surface that interacts with the blood flow streaming through the arteries. The role of this layer is crucial; it is compromised in many disease states and in particular in regions of the artery wall where atherosclerosis is found [62]. The glycocalyx is a complex structure. The GAGs that populate the layer are very long flexible molecules that have a net negative charge and hence a complex interaction with one another as well as the transmembrane proteins to which they are attached. Enzymes that digest GAGs, such as heparinase and chondroitinase, inhibit response to WSS [13]. However, such interventions would also interfere with basal surface proteoglycans that form co-receptors with integrins [63] and, as integrins have been strongly associated with responses to flow [50], it is not clear whether direct mechanosensing via the glycocalyx is taking place. Because of its proteoglycan and GAG composition and structure, the glycocalyx may cover the entire surface of the EC and thus can interact with other EC sensors and transducers to play a role in sensing and transmitting haemodynamic forces [64]. A detailed analysis of the glycocalyx layer follows in subsequent sections of this chapter.

Blood flow and atherosclerosis

ECs are influenced by frictional stress, the WSS, generated by the blood. In vitro,

ECs change from a raised polygonal to a flattened elongated morphology under the influence of flow with the degree of elongation correlated to the magnitude of the shear stress [7] [8] [9]. Cells exposed to unidirectional and pulsatile flow *in vitro* and *in vivo* express multiple atheroprotective genes [65]. This type of flow which is prevalent to most of the vascular system, overall suppresses atherogenesis by decreasing EC turnover and promoting EC alignment with flow direction.

Often the term "disturbed flow" is used to describe patterns of low flow, flow separation, gradients, flow reversal and in some situations, turbulence. High rates of EC proliferation and apoptosis are associated with regions of disturbed flow *in vivo*. The ECs in those regions fail to align with the flow direction and exhibit an increased population of reactive oxygen species (ROS) and expression of inflammatory factors [27], [66]. The clinical implications of shear stress in vascular pathology can be systemic, for example influenced by exercise, and also site-specific given the varying vessel geometry, existence of lesions and post-intervention complications (for example after stent deployment) [41]. A comprehensive review of the role of disturbed flow on vascular endothelium pathophysiology has been produced by Chiu et al. [67]. The notion that low mean WSS and high oscillatory shear index dominate as mechanical localising factors in atherogenesis is currently the majority opinion in the field. However, the low versus high WSS as a leading cause of disease debate is still ongoing [68].

2.1.3 Systemic Risk Factors

Several systemic risk factors for cardiovascular disease have been identified in clinical practice. In these section we summarise a few of the most prominent ones. We refer to them as systemic in contrast to localised effects, such as damage to the vasculature due to local morphology or injury.

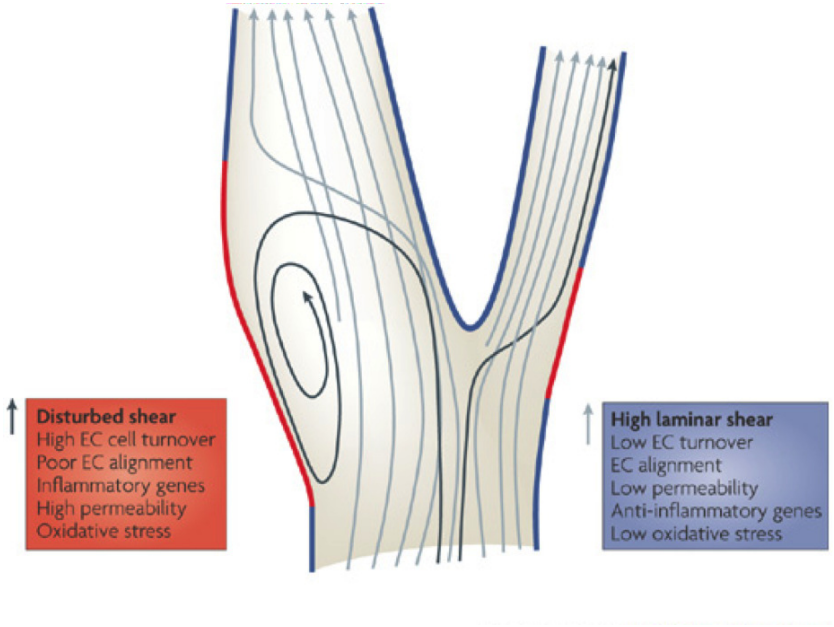


Figure 2.9: ***Flow pattern representation in the carotid bifurcation.*** *Regions of high laminar shear (blue) and disturbed shear (red) [24].*

Hypertension The pulsatile component of blood pressure has been identified as a major risk factor for atherosclerotic plaque formation and progression [6]. Blood pressure is lower in veins, which could explain why atherosclerosis is encountered in arteries. One of the explanations for the effects of hypertension lies in the pathways of protein kinases and associated transcription factors which are activated by mechanical stretch, regulating the expression of inflammatory mediators that contribute to leukocyte adhesion in atherosclerosis [69]. In the case of the c-Jun N-terminal kinases (JNKs), their activation is dependent on stress direction, and only stretch parallel to the actin stress fibres activates JKN, whereas perpendicular stretch does not [70]. It follows that ECs which are correctly aligned, would show little JKN activation. In regions of disturbed flow, ECs are poorly aligned, and therefore more sensitive to stretch-induced inflammatory pathway activation. Hypertensive arteries are also characterised by greater pulse wave velocity, an indicator of increased vascular stiffness [4], most likely due to the accumulation of collagen. The combined effects of disturbed flow and stretch in causing oxidative stress [36] could promote inflammatory pathways as described above, and inhibit

vasorelaxation by eliminating NO [71].

Hyperlipidaemia High levels of cholesterol is one of the most important and well-known atherosclerosis risk factors. LDLs and very low density lipoproteins (VLDLs) are the vehicles transporting cholesterol to the atheroprone regions of the arteries [72], where the endothelial permeability is increased [73]. ECM remodelling in the plaque increases the retention of cholesterol in the vessel wall [72]. Furthermore, oxidised LDLs are generated during oxidant stress and permeation of lipoproteins, causing further inflammation [74]. As already discussed, high plasma membrane cholesterol has been observed to participate in an ion-channel inhibiting mechanism [56]. In contrast to the action of LDL and VLDL lipoproteins, high-density lipoproteins (HDLs) directly oppose atherosclerosis, by acting at several key stages to interrupt the process of atherogenesis [75]. HDL is involved in functions of removing cholesterol from foam cells, inhibiting the oxidation of LDLs, and limiting the inflammatory processes that underlie atherosclerosis. HDLs also have antithrombotic properties [76].

Sedentary lifestyle The influence of exercise, which has been shown to be extremely beneficial against cardiovascular disease, cannot be explained solely on the basis of lowering the levels of known risk factors, such as blood pressure, glucose and lipoprotein levels [77]. The overall increase in fluid shear stress, due to elevated cardiovascular output, has been suggested as a possible explanation of the added benefit of exercise [78] [79] to the prevention of cardiovascular disease.

Diabetes Type 2 diabetes has been shown to increase the risk of death from cardiovascular disease by 2-6 fold [5]. Although the exact mechanisms by which diabetes contributes to atherosclerosis remain unknown, there is some evidence that it causes upregulation of several basement membrane proteins in endothelial cells, which have been linked to ECM remodelling, observed in atheroprone regions [80].

This effect was reproduced in *in vitro* cultures of ECs in high glucose conditions, which mimic the effects of diabetes [81]. Evidence also suggests a link between diabetes, cardiovascular risk and oxidative stress via the increased production of ROS [82].

2.2 Glycocalyx Overview

2.2.1 Early Findings

Early indications of an EGL *in vivo* were observed by Pries et al. [83] who performed flow resistance experiments in capillaries of up to 30 μm in diameter and compared the findings with glass tubes of same size. They recorded a dramatic difference in resistance, which decreased after enzymatic removal of the EGL [84]. In another study, Desjardins et al. [85] observed an increase in hematocrit after enzyme treatment to remove EGL components.

The EGL appeared to play a role in vascular disease since the early large vessel studies. Lewis et al. investigated EGL dimensions in coronary arteries with atherosclerosis in White Carleau pigeons [86]. EGL was thicker in healthy areas, in comparison to areas with observable atherosclerosis. Furthermore, high cholesterol diet correlated with decreased thickness in all arterial zones, indicating the protective role of the EGL.

Direct intravital microscopic evidence for the existence and function of the glycocalyx was obtained by Vink et al. who used a dye exclusion technique to estimate *in vivo* thickness of the EGL in capillaries [87] in the mid-nineties. By comparison of anatomical and functional diameter in hamster muscle capillaries they estimated an EGL thickness of 0.4-0.5 μm.

2.2.2 The EGL Structure

The schematic of Figure 2.10 gives a good representation on the glycocalyx and its position relative to the cell membrane along with the surrounding environment. It comprises glycoproteins bearing acidic oligosaccharides and terminal sialic acids and proteoglycans along with their associated GAG side chains. The proteoglycans found on the luminal surface of endothelial cells are syndecans-1, -2, and -4, and glypican-1 [11]. The polyanionic nature of its constituent GAGs provides it with a net negative charge [88].

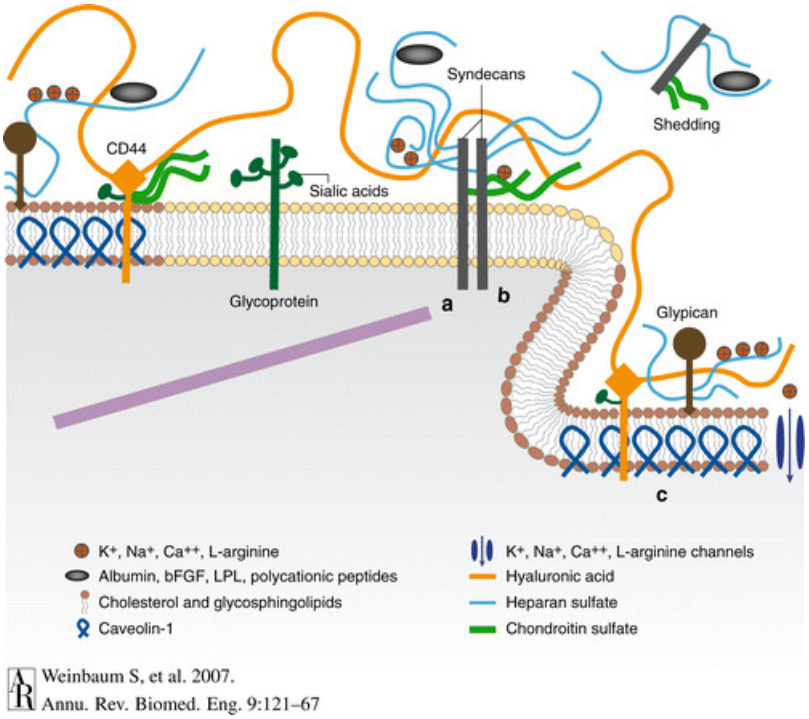


Figure 2.10: ***Glycocalyx layer structure.*** *A schematic representation of proteoglycans and glycoproteins on the surface of ECs. [13]*

GAGs are linear polydisperse heteropolysaccharides, characterized by distinct disaccharide unit repeats. Specific combinations of these give rise to different GAG families, such as the heparan sulfate (HS), chondroitin/dermatan sulfate (CS/DS), and hyaluronic acid or hyaluronan (HA) found on ECs. Proteoglycans are proteins that contain specific sites where sulfated GAGs are covalently attached. Local ionic strength and pH strongly influence the level of extensibility of a GAG chain, and

it appears that this is maximal in a NaCl solution having physiological concentration [13].

Zheng et al. investigated the hypothesis that the GAG components of the EGL are interdependent and that the EGL would therefore collapse after enzymatic removal of one or more of them. They tested this hypothesis using confocal microscopy and observed that the coverage and thickness of HS, CS, HA and absorbed albumin were similar pre- and post-treatment and that the thickness of individual GAGs were spatially nonuniform. Specifically, removal of HS or HA did not result in cleavage or collapse of any of the remaining components. Simultaneous removal of CS and HA by chondroitinase did not affect HS, but did reduce adsorbed albumin, although the effect was not large. They concluded therefore that, despite the inter-mixing of EGL components, they do not interact with one another [89] in terms of EGL structural integrity and distribution on the EC surface.

2.2.3 EGL properties

EGL Thickness and distribution Using a method referred to as microviscometry, which involves analysing micro-scale particle image velocimetry (μ-PIV[1]) data over the entire cross-section of the vessel, Damiano et al. [90] showed that in mouse cremaster-muscle venules (20-40 μm diameter), the mean hydrodynamically effective EGL thickness was ~ 0.5 μm in control vessels and ~ 0.2 μm after light-dye treatment to degrade the EGL.

Using the same technique, together with a detailed three-dimensional analysis of the fluid dynamics around the microsphere and throughout the glycocalyx, Savery et al. have studied the hydrodynamic relevance of the glycocalyx in arterioles [91]. They were able to show that the only possible velocity distribution within the glycocalyx that was consistent with the μ-PIV data had to be highly nonlinear and

[1]In μ-PIV, fluorescent microspheres are injected in the blood stream and subsequently visualised, thus providing information on the near-wall flow conditions

therefore indicative of a region near the vessel wall that significantly decelerates plasma flow.

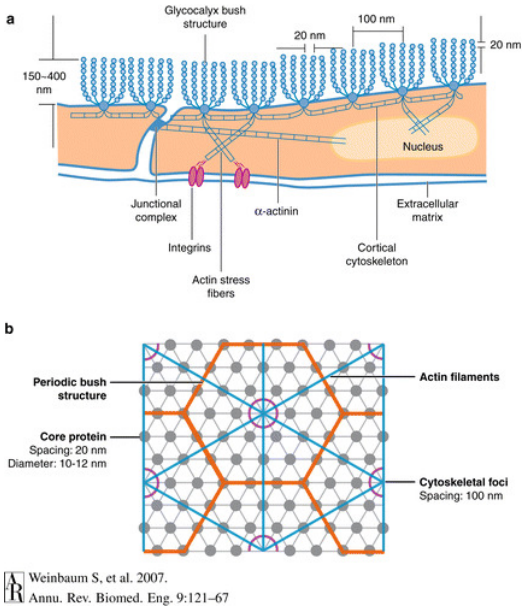


Figure 2.11: ***Structural model for the EGL.*** *(a) Sketch of the arrangement of core proteins in the EGL and its anchorage to the underlying actin cortical cytoskeleton. (b) En face view of the idealised mathematical model in Weinbaum et al. [92] showing the hexagonal arrangement of core proteins and cluster foci.*

As early as 2001, pioneering work by Squire et al. [93] revealed the ultrastructure of the glycocalyx in frog mesentery arteries using electron microscopy images. The results suggested that the glycocalyx consists of "fibres" - of non-specified molecular type but most likely proteoglycans - between 10 to 14 nm in diameter in a square lattice of 20-nm spacing both parallel and perpendicular to the EC surface. A much smaller periodicity of about 8 nm was also observed. Additionally, there was evidence of a quasi-hexagonal lattice of larger spacing, about 100 nm, across the luminal surface. This informed the Weinbaum EGL model, as seen in Figure 2.11. A decade later, these observations were found to be largely similar in mammalian (rat and rabbit) ECs in a study by Arkill et al. [94], using a similar methodology as Squire et al.

More recently, Bai et al. [95] discovered through *in-vitro* experiments, that HUVEC glycocalyx, fully developed under static conditions in culture (over a period

of ten days), experienced a re-distribution towards the cell edges, near the cell-cell junctions, rather than in the apical cell surface, as a result of shear-stress exposure. Similar findings were independently observed by Zeng et al. on rat fat pad endothelial cell studies [96]. After exposure to shear stress for 24 h, HS, chondroitin sulfate, glypican-1 and syndecan-1, were enhanced on the apical surface, with nearly uniform spatial distributions close to baseline levels that differed greatly from the distributions measured at 30 minutes after exposure to WSS. HS and glypican-1 still clustered near the cell boundaries after 24 h of shear, but caveolin-1/caveolae and actin were enhanced and concentrated across the apical aspects of the cell.

Mechanical Properties of the EGL In the elastohydrodynamic theory for the structural integrity of the EGL, the important parameter for determining how the proteoglycan core proteins transmit WSS from the free surface of the EGL to the actin cortical cytosceleton is the flexural rigidity, EI, of the core proteins of the bush-like structures shown in Figure 2.11 [92].

In absence of a method for direct measurement of elastic properties of the EGL, Weinbaum et al. [92] determine EI using an indirect approach in which they predict the measured characteristic relaxation time for the time-dependent restoration of the EGL after it has been crushed by the passage of a white blood cell (WBC) in a tightly fitting capillary. To quantify this effect, Vink et al. [97] measured the gap between the endothelial and red cell membranes of red cells that passed in the wake of the WBC. By measuring the time lag between the passage of the red cell and WBC, one could measure the time-dependent restoration of this gap. The recovery curve for the displacement of the red cell can be approximately described by a single exponential with a characteristic time of $0.38s$. EI was determined by constructing an approximate viscoelastic model for this recovery. This linearised model predicts that EI is 700 $pN\ nm^2$.

Subsequently, Han et al. [98] developed a more sophisticated large deformation

model for elastica that predicts the time-dependent changes in shape of the core proteins from their initial large deformation to their final restoration after the passage of the white cell. This more realistic model predicts that EI is 490 $pN\ nm^2$.

More recently, Bai et al. [99] measured the Young's modulus of HUVEC EGL using atomic force microscopy nano-indentation. Concretely, the EGL Young's modulus was deduced from Young's moduli of cell membranes with and without (after neuraminidase treatment) the glycocalyx layer. Their results show the glycocalyx on cultured HUVECs has a Young's modulus of approximately 0.39 kPa.

Evidence suggests that the EGL has finite resistance to compression (see Pries et al. [100]). This is supported by the fact that the layer is able to withstand the shear stress exerted by the flowing blood. Furthermore, the EGL recovers its width in about 1 second after being compressed by a passing white blood cell [97] in experiments as already described. As a result of this evidence, and the fact that the EGL can be compressed by static red cells, Secomb et al. suggested that the difference in pressure within the EGL as opposed to the area directly above it is in the order of 20 dyn/cm^2 [101].

Electrochemical Properties of the EGL Electrochemical properties where measured by Vink et al. experimentally [102] [103], whereas theoretical and computational models where constructed by Stace & Damiano [104]. Their assumption was neutrality within the overall vessel or capillary in question, leading to the conclusion that all mobile charged molecules (mainly cations and anions) are of equal opposite charges. However, this cannot be the case within the EGL layer, where the fixed anionic nature of the glycocalyx components can only be neutralised by a comparatively larger concentration of cations. Therefore the anionic concentration within the layer is decreased. This is to say, Stace & Damiano realised early on that the local ionic concentrations near the surface of ECs cannot be equivalent to physiological measurements typically found in blood plasma.

By adapting their model to experimental results of anionic tracer exclusion by the EGL, Stace & Damiano calculated the corresponding charge density of the EGL (the part borne by the GAGs and not the ionic concentration of the medium) between 0.7 and 1.3 mEq/l, which represents approximately 1% of the ionic strength of plasma. The model was subsequently extended by Damiano & Stace [105], focusing on the development of a theoretical model to predict the electrochemical potential gradient across the EGL and the mechanics of the restoring-force mechanism during compression, for example by a leukocyte, and recovery of the EGL.

2.2.4 Physiological Function

Permeability and Revised Starling Principle

One of the first physiological functions that the existence of EGL helped explain was the assumed violation of the Starling Principle, highlighted by Levick [106]. According to the classic Starling Principle, the filtration through the cell membrane is powered by net difference between the hydraulic capillary pressure, generated by the heart and the colloid osmotic pressure due to plasma proteins. Also, according to the Starling principle, the balance between the pressure generated by constriction of blood vessels and the interstitial protein osmotic pressure can force fluid over the membrane into the extracellular space.

It is now accepted that the Starling Principle is violated in most tissues when measurements are made with the latest experimental methods. This, however, does not mean that the Starling Principle is wrong, only that it is incorrectly applied. According to recent models, the pathway of solvents across the micro vessel wall, from the luminal side to the tissue space consists of the endothelial surface glycocalyx and clefts between neighbouring endothelial cells.

Adamson et al. [107] showed from their detailed structural morphology measurements for rat mesentery venules that the clefts between adjacent endothelial cells have uniform width and there are tight junction strands with discontinuous leakages,

or gaps. Illustrated in Figure 2.11 is the schematic view of ultrastructures of the inter-endothelial cleft and the surface glycocalyx.

The Starling principle has been rectified by the Michel-Weinbaum model [108] [109] [106] which proposes that the primary molecular filtering for plasma proteins occurs in the EGL and that the Starling forces are determined, not by the global difference in oncotic pressure between lumen and tissue, but by the local difference in protein concentration across the EGL.

EGL as a protective barrier

The protective nature of the EGL stems from its permeability modulating function, as discussed in the previous section, and therefore also manifests as a barrier against cell adhesion. This was illustrated by Mulivor et al. [110] in experiments where removal of essential glycocalyx elements via heparinase in rat post-capillary venules, induced an increase in leucocyte adhesion.

Increased shedding of syndecan-1 and HS is often associated with damage to the EGL and as a result, these two molecules are used as EGL damage biomarkers [14] [111].

Chappel et al. used a different technique to investigate the protective properties of the EGL by conducting induced ischemia/reperfusion experiments on guinea pig hearts and by application of hydrocortisone treatment prior to the ischemic episode. They concluded that the way in which hydrocortisone protects against ischemia/reperfusion injury is by helping to keep intact the endothelial glycocalyx [16].This was subsequently shown for similar experiments using different EGL preservation agents [12] [111].

Mechanotransduction and the EGL

It has been recently recognised that the EGL serves a critical role in the transmission of WSS to the actin cytoskeleton and in the initiation of nitric oxide (NO) production [112] and intracellular signaling [13], via the mechanism of mechanotrans-

duction, very likely as a primary mechanosensor.

Evidence that supports a major role for the EGL in mechanotransduction originates in experiments in which enzymes were used to selectively degrade specific components of the EGL, followed by a reassessment of function, or, from using bathing solutions without plasma proteins where the EGL is compromised [113] and structural organisation examined in response to WSS.

In experiments by Florian et al. [114], heparinase III was used to selectively deplete the glycocalyx of the HS GAG component of bovine aortic endothelial cells *in vitro*. The NO production rate in response to shearing conditions was determined for heparinase treated and untreated monolayers, and it was observed that heparinase treatment that removed only 46% of the heparan sulfate completely inhibited shear-induced NO production.

In similar experiments, Hecker et al. showed that when intact segments of rabbit femoral arteries were pretreated with neuraminidase, which induces removal of sialic acids from the glycocalyx, shear-induced NO production was inhibited. They also demonstrated that the same enzyme treatment had no effect on another hallmark response of ECs, the shear-induced production of prostacyclin (PGI2) [115]. This study illustrated the fact that there are multiple mechanisms of mechanotransduction and not a single mechanosensor. In another study, the enzyme hyaluronidase was used in isolated canine femoral arteries to degrade the hyaluronic acid (HA) GAG from the EGL layer, and a significant inhibition of shear-induced NO production was demonstrated [116].

Similar findings were observed in a recent study on NO production in microvessels *in vivo* [117]. Yen et al. used fluorescence microscopy to measure NO production in post-capillary venules and arterioles of rat mesentery under reduced (low) and normal (high) flow conditions, with and without enzyme pretreatment to re-

move HS component of the EGL and in the presence of an endothelial NO synthase (eNOS) inhibitor. Their results suggest that the EGL participates in endothelial cell mechanosensing and transduction through its heparan sulfate to activate eNOS.

2.3 Proteoglycans and glycosaminoglycans

Proteoglycans are vital constituents of cellular membranes and extracellular matrices, where they confer tissue mechanical properties and participate in cell-cell interactions [118]. They have been shown to be implicated in neurodegeneration, arthritis, and cancer [119]. Human proteoglycans comprise a core protein plus one or more covalently linked GAG polymers, which often possess functionally important patterns of post-synthetic modification. To this date, obtaining 3D experimental data for proteoglycans is hampered by GAG sequence heterogeneity, conformational flexibility, and molecular size. Experiments yield either low-resolution results (for example using electron microscopy or small-angle X-ray scattering), or are confined to studying isolated elements of the proteoglycan structure (the protein core itself, a small GAG strand) in crystallographic studies.

As well as being the essential building blocks of the EGL, GAGs are important in many other structural and functional biological processes, including blood clotting and injury repair to name a few, and as such, have been widely studied. A key component in several GAGs is the presence of heterogenous sulfation which is known to alter biological recognition and polymer structure. In particular, domains rich in sulfate and iduronic acid maintain connective tissue integrity and mediate signal transduction via multivalent interactions with collagen and growth factors. This section is dedicated to a review of the existing literature concerning the structure and properties of GAGs, and in particular the one most abundant in the EGL: HS alongside its close cousin, heparin.

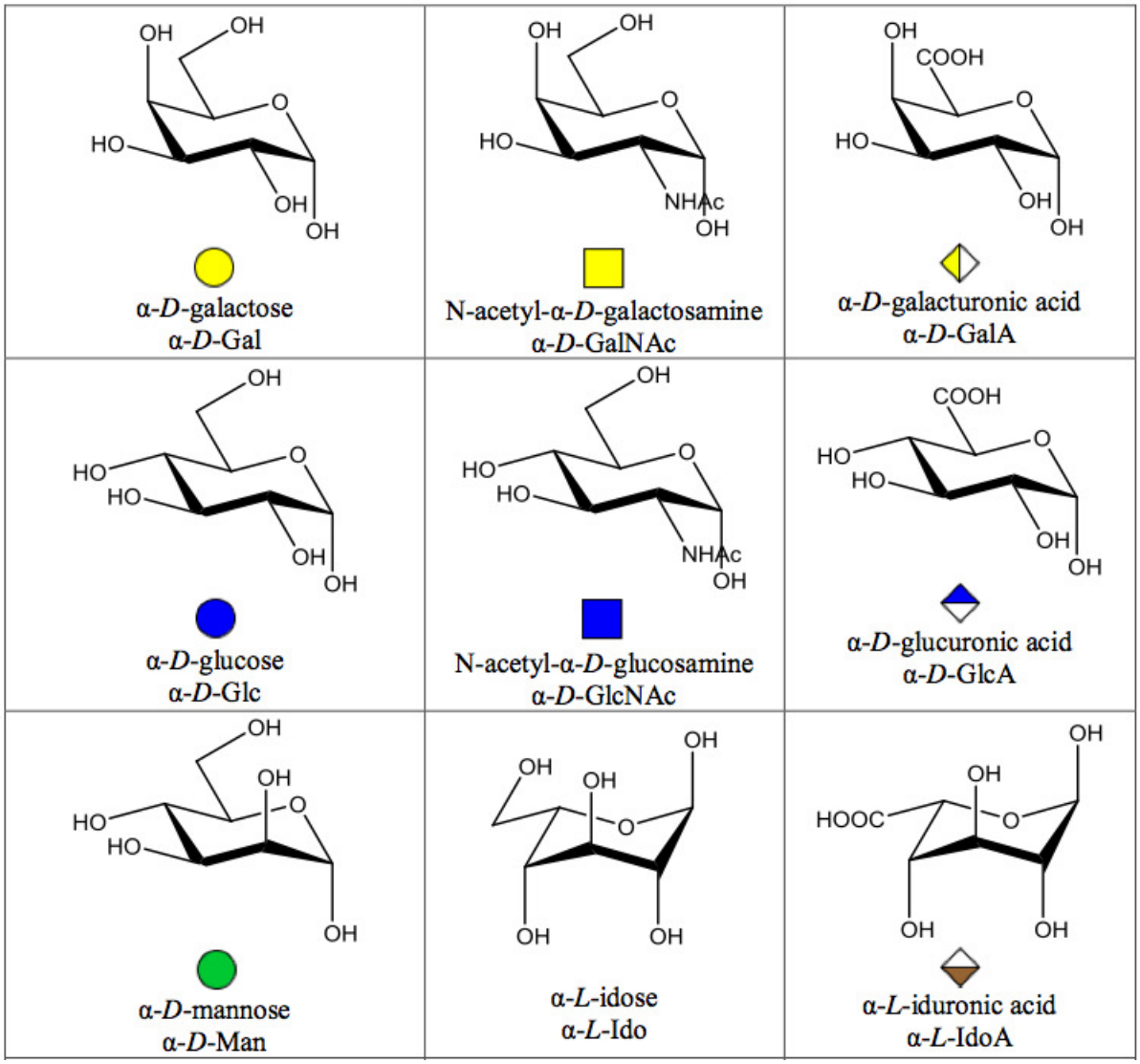


Figure 2.12: ***Examples of monosaccharide residues which form six-membered rings.*** *Shown from top to bottom are the chemical structure, CFG cartoon representation, chemical name, and abbreviated chemical name.*

2.3.1 Structure

GAGs such as HS and heparin are composed of repeating disaccharide units of a hexosamine, and another hexose that is often an anionic residue, such as D-glucuronate (GlcA) or L-Iduronate (IdoA) (see Figure 2.12) [120]. Both heparan sulfate and heparin are biosynthesised, initially as non-sulfated precursor polymers composed of alternating sequences of glucuronic acid ($GlcUA$) and $N-$acetylglucosamine ($GlcNAc$) called heparan, or $N-$acetylheparosan: -GlcUAβ1 $\rightarrow$ 4GlcNAcα1 $\rightarrow$ 4- [121]. In a

subsequent stage, the polymer is enzymatically transformed into complex sulfated derivatives.

2.3.2 Ring Flexing

The ability to accurately model heparin, heparan sulfate and other carbohydrate sequences at length scales ranging from bioactive oligosaccharides to large polysaccharides is complicated by ring flexing (or ring pucker) motions that occur in individual monosaccharides. The pucker of six-membered pyranose rings is proposed to be a defining microsecond-time-scale, 3D signature of numerous chemically distinct monosaccharides and is mechanistically important for molecular recognition and bioactivity (e.g., in heparin-mediated anticoagulation) [20].

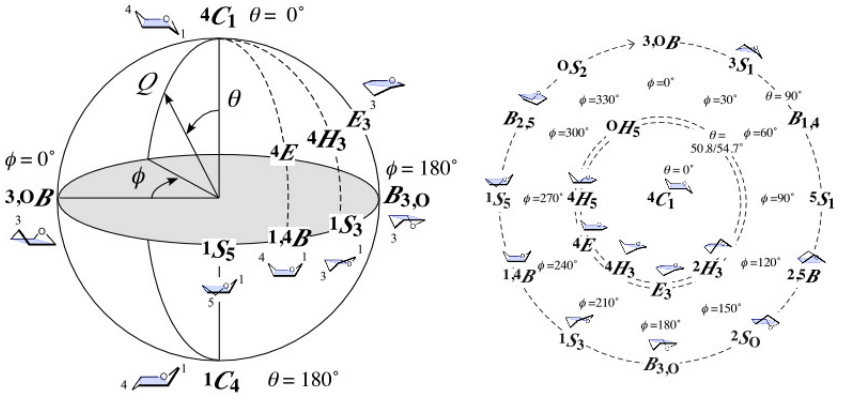


Figure 2.13: ***Spherical polar representation of Cremer and Pople parameters.*** *Left: meridian angle ϕ, azimuthal angle θ, and radius Q [122] and right: Stoddart (or "pseudorotational") diagram (projection of the sphere from the North Pole)*

The description of ring conformation can be either qualitative (typically by assigning conventional names to special ring conformers) or quantitative (employing a parametric measure of the deviation from a reference shape). The first systematic study of this problem came with the group theory analysis of Pickett and Strauss [123] and eventually with the first general definition of ring pucker coordinates of Cremer & Pople [124]. Both Pickett & Strauss and Cremer & Pople approaches describe the conformation of a N-membered ring with 3 parameters, expressed in terms of the elevations of the ring atoms from a reference plane (the

mean ring plane for Cremer & Pople coordinates, and the planar ring conformation for Pickett & Strauss coordinates). These two approaches, proven to be equivalent, represented the first mathematically consistent and quantitative measure of ring conformers for the pucker definition problem. The spherical polar representation of Cremer & Pople parameters, annotated with various ring conformations, can be seen in Figure 2.13. The pyranose ring shape in often highly populated chair conformations ($^1C_4$ and $^4C_1$) can be seen in Figure 2.14). Other popular representations include the boat and skew boat shapes, denoted with the letters B and S respectively.



Figure 2.14: ***Pyranose ring pucker definition:*** $^1C_4$ *and* $^4C_1$ *chairs*

Sulfation

The presence of sulfate groups is a key structural and functional feature of heparin and HS as well as other GAGs. It can be seen in figure 2.15 how all GAGs, with the exception of HA appear in sulfated forms. The N-sulfate substituent is unique to the heparin and HS structures, and as a consequence, we can define heparan and HS as two separate families of N-sulfated GAGs. Both the level of sulfation, as well as sulfation patterns, are highly variable in biologically-sourced GAGs, making sequence analysis difficult. This often limits experimental characterization of GAG structures to composition-based analysis of digested fragments of native GAGs. More detailed analysis, such as by nuclear magnetic resonance (NMR) spectroscopy or crystallography usually employs short isolated, or synthetic oligomers, where the sulfation patterns are well controlled.

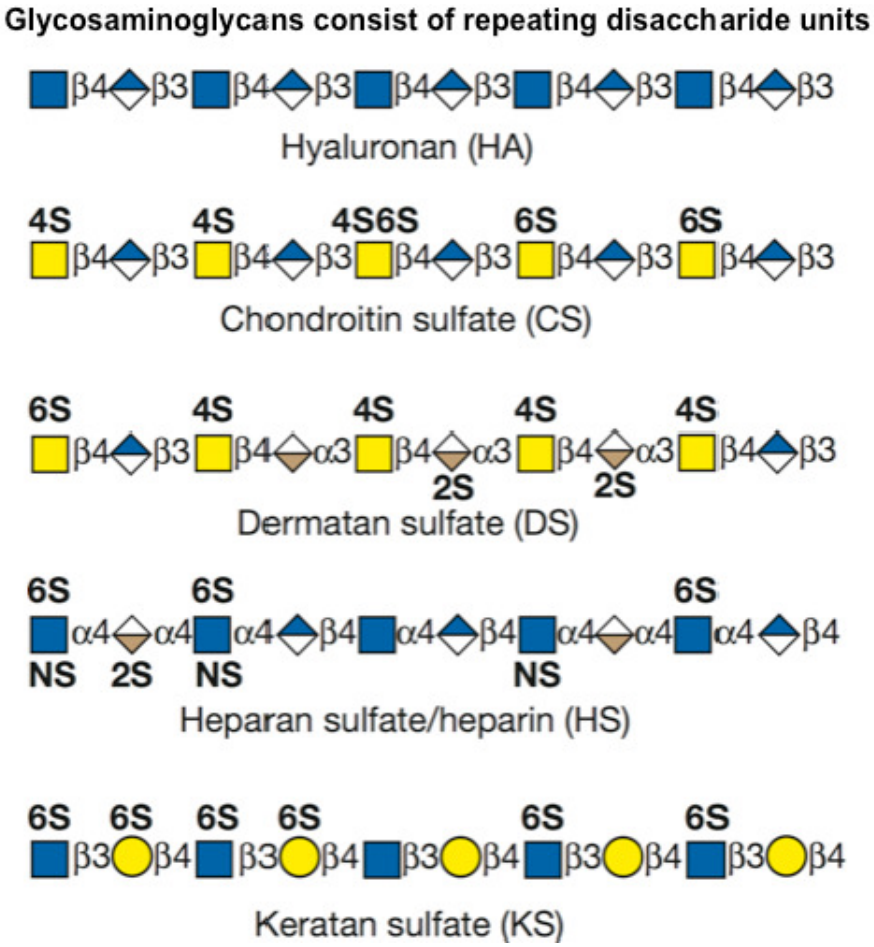


Figure 2.15: ***Glycosaminoglycan structural elements.*** *GAGs consist of repeating disaccharide units composed of an N-acetylated or N-sulfated hexosamine and either a uronic acid (glucuronic acid or iduronic acid) or galactose. Hyaluronan lacks sulfate groups, but the rest of the glycosaminoglycans contain sulfates at various positions. DS is distinguished from CS by the presence of iduronic acid. Keratan sulfates lack uronic acids and instead consist of sulfated galactose and N-acetylglucosamine residues. [125]*

While the sulfation patterns are essential for altering the charge of GAGs, they have also been shown to directly impact the 3-dimensional structure of GAG fragments. NMR spectroscopy of GAG fragments has shown that sulfation patterns can alter the ring pucker associated with non-reducing terminal $\Delta 4,5$-unsaturated uronates [126], [127], [128] and IdoA [129], [130], [131]. Recently published simulations of HA/HS GAGs have shown that IdoA ring pucker transitions (flipping from one hexopyranose ring shape to another) can have a significant impact on the 3D shape of the GAG polymer [20]. Although as noted earlier [132], not all differences in the ring pucker lead to altered overall 3D shape. Whether or not sulfation sig-

nificantly impacts the inter-residue linkage conformations (the glycosidic linkages), or the pucker properties of rings other than IdoA is uncertain. While early modeling studies of heparin disaccharides suggested that the glycosidic linkages were not significantly affected by sulfation pattern, a survey of the protein crystallographic database (PDB) indicates that glycosidic linkages in GAGs do exhibit considerable flexibility. The question as to whether GAG conformation, other than the IdoA ring pucker, is affected by sulfation pattern therefore remains open.

Differences between HS and heparin

The most common disaccharide unit within heparan sulfate is composed of a glucuronic acid (GlcA) linked to N-acetylglucosamine (GlcNAc) typically making up around 50% of the total disaccharide units. In heparin, IdoA(2S)-GlcNS(6S) makes up 85% of heparins from beef lung and about 75% of those from porcine intestinal mucosa. This can become problematic when attempting to categorise hybrid GAGs that contain both "heparin-like" and "HS-like" structures. It has been suggested that a GAG should qualify as heparin only if its content of N-sulfate groups largely exceeds that of N-acetyl groups and the concentration of O-sulfate groups exceeds those of N-sulfate. [133]

2.3.3 Heparan Sulfate Turnover in ECs

In 2008, Potter et al. [134] suggested that, since the EGL is largely absent from static EC cultures, inferences made on the basis of *in vitro* studies with regards to EC surface chemistry should be reconsidered. Indeed, the EGL will only develop in a static culture provided hyaluronan and chondroitin sulfate are added to the cell culture media at hyper-physiological concentrations.

High shear stress has been shown to increase the rate of glycocalyx biosynthesis in the ECs of several tissues, compared to cells experiencing no shear [135] [136] [137]. Giantsos-Adams et al. [138] quantified the increases of heparan sulfate on the

endothelial surface in response to high shear stress and validated the results using fluorescence microscopy. They further presented a time course for the regeneration of HS on the EGL following heparinase cleaving and correlated that recovery course with cell elongation mechanics, demonstrating an active and dynamic glycocalyx present in cultured ECs (HUVECs). Their kinetic model of HS EC turnover is schematically presented in Figure 2.16.

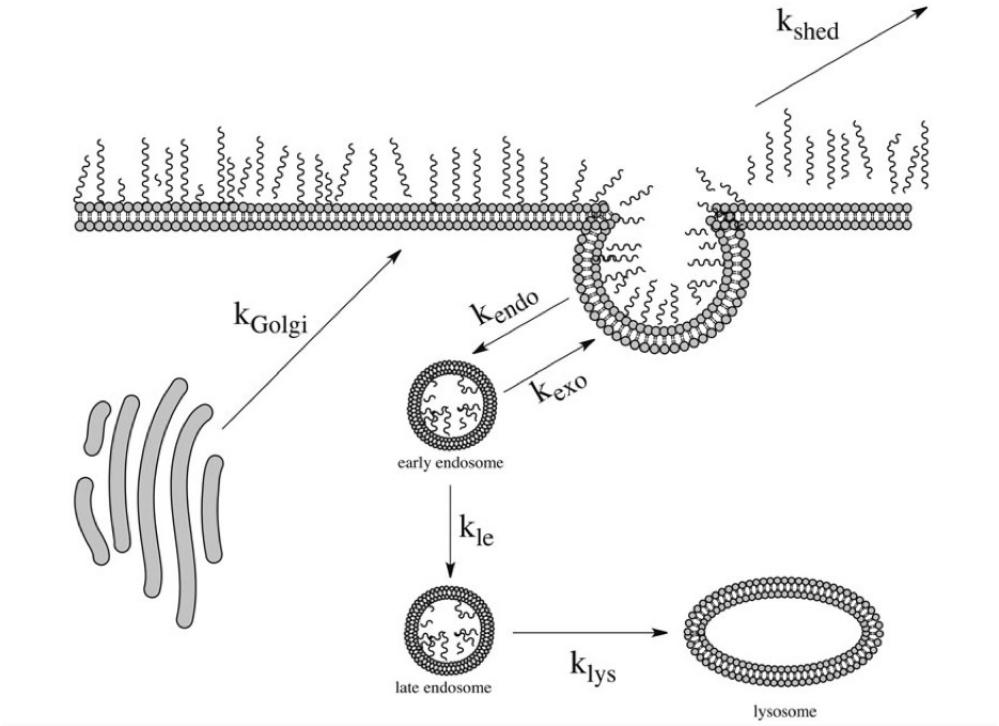


Figure 2.16: ***Kinetic model of heparan sulfate turnover from cell surface.*** *Heparan sulfate concentration on cell surface is modelled as a function of the rate of HS production from the Golgi apparatus (k_{golgi}), endocytosis (k_{endo}) from the cell surface, and exocytosis from recycling vesicles (k_{exo}) and from the late endosome (k_{endole}). [138]*

2.3.4 Properties and Vasoregulatory Function

GAG sulfation patterns have been demonstrated to modulate biological function, for example in the cases of heparan sulfate in growth factor activation and cellular defense [139] [140], chondroitin sulfate growth factor recognition [141], and synthetic heparin/heparan sulfate in anticoagulant activity [142].

In experiments conducted by Kang et al., compositional (via HS enzymatic re-

duction) or effective charge density modification (varying ionic strength) of the EGL facilitated water filtration across the endothelium, enhanced the accumulation of LDLs on the luminal surface and increased the amount of LDL uptake by endothelial cells. Thus, they showed that EGL perturbation in certain arterial sites accelerates flow-dependent concentration polarization of LDLs, which in turn contributes to vulnerability of these arterial sites to atherosclerosis [143].

Based on observations of Na^+ dependence of the flow-induced dilation in resistance arteries [144], Siegel et al. [145] developed a mechanosensory hypothesis involving the role of a sensor-molecule, which they identified as heparan sulfate, based on its viscoelastic and electrochemical properties. According to this hypothesis, proteoheparan sulfate (Siegel et al. refer to HS as well as the attached proteoglycan, in this case assuming a protein of the syndecan family) is present as random coil under 'no flow' conditions, whereby most of the polyanionic sites undergo intramolecular hydrogen bonding. With increasing flow the sensor macromolecule becomes unfolded to a filament structure (Figure 2.17). Siegel et al. propose that in flowing viscous solutions containing molecular coils, these coils will align with the direction of flow.

This proposed mechanism was not, to the best of our knowledge, further explored. The hypothesis fails to explain the exact mechanism by which, given the velocity rates expected near the wall of even the largest vessels in mammals, the flow rate would have such an impact on HS to induce uncoiling. In fact, it is most likely the local charge density due to the ionic strength of plasma that influences the shape of HS and other GAG strands. This is supported by polymer brush theory and experiments, further discussed in the following chapter.

2.4 The multi-scale relevance of the EGL

We have highlighted in this chapter the importance of the EGL and its protective and sensing roles that ensure the function of a healthy vasculature. These roles are

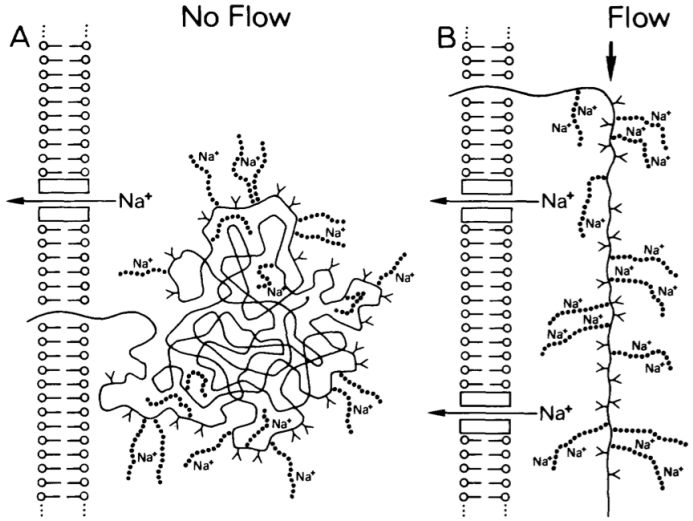


Figure 2.17: ***Proteoheparan Sulfate as a flow sensor at the endothelium-blood interface.*** *(A) The membrane of an endothelial cell with an unspecific, voltage-dependent ion channel, mainly conducting Na^+, is represented schematically. The hydrophobic domain of the protein core of a proteoheparan sulfate macromolecule is anchored in the cell membrane. Under no-flow conditions the viscoelastic polyelectrolyte assumes a random coil form, whereby most of the polyanion sites undergo intramolecular hydrogen bonding. (B) Blood flow causes a conformational change from the random coil state to the directed filament structure state. This conformational transition effects an unfurling of the GAG side chains and is therefore combine with an increase in binding sites for Na^+ ions. Counter-ion migration of Na^+ along the polysaccharide chain is followed by a transmembrane Na^+ influx into the endothelial cell [145].*

experimentally and theoretically examined through the lens of its various properties. For example, properties such as spatial arrangement, thickness, electrochemical properties and GAG component independence when studying the EGL structural stability and its effect on permeability modulation.

In studies of the mechanosensing properties of the EGL, its components, both proteoglycans and GAGs, are often examined in isolation, by targeted enzymatic removal. In this case as well, the electrochemical properties of the EGL are considered important, along with its mechanical properties such as the flexural rigidity and Young's modulus. At the same time, the field of complex carbohydrate chemistry is developing the computational and experimental tools to investigate in detail GAGs and associated proteins, offering the promise of high resolution view of the EGL.

Finally, the EGL is located in all vascular ECs, and location heavily influences

local hydrodynamic properties of the blood flow. Flow induces EGL production and redistribution and has time-dependent effects on the EGL. As a mechanosensor, the EGL responds to flow enabling the propagation of the signal thought the cell, leading to observable remodelling of the cytoskeleton and effects reaching all the way to gene expression (eNOS activation). The effect of the magnitude of WSS and its role in vascular disease remains unclear.

Chapter 3

EGL Modelling

In this chapter we discuss several approaches to modelling the EGL, including cellular-level models that take into account an explicit glycocalyx to some degree of detail. Because of the many different functions the EGL serves, most models mainly focus on one aspect, such as its role in permeability, adhesion or mechanosensing. Other models attempt to capture elements of the structure and mechanical and chemical properties of glycocalyx components, whereas in some studies the glycocalyx is often represented as a polymer brush system. Based on the scale of interest (continuum and coarse grained as opposed to fully atomistic) we present two categories of glycocalyx models. Finally, we also discuss methods of flow implementation in atomic-scale simulations, with a further focus on biological systems.

3.1 Continuum and Coarse-grained Models

3.1.1 EGL Effects on Permeability and Adhesion

Based on permeability measurements and new insights from the revised Starling principle, Sugihara-Seki et al. presented a mathematical model for steady flow across the endothelium in microvessels [146] and used it to calculate the hydraulic conductivity L_p of the total system consisting of the glycocalyx layer and the inter-

endothelial cleft/tight junctions (see schematic: Figure 3.1). Their model suggests much smaller contribution of the surface glycocalyx to the hydraulic conductivity than the previous estimates, primarily owing to substantial spreading of the flow within the glycocalyx.

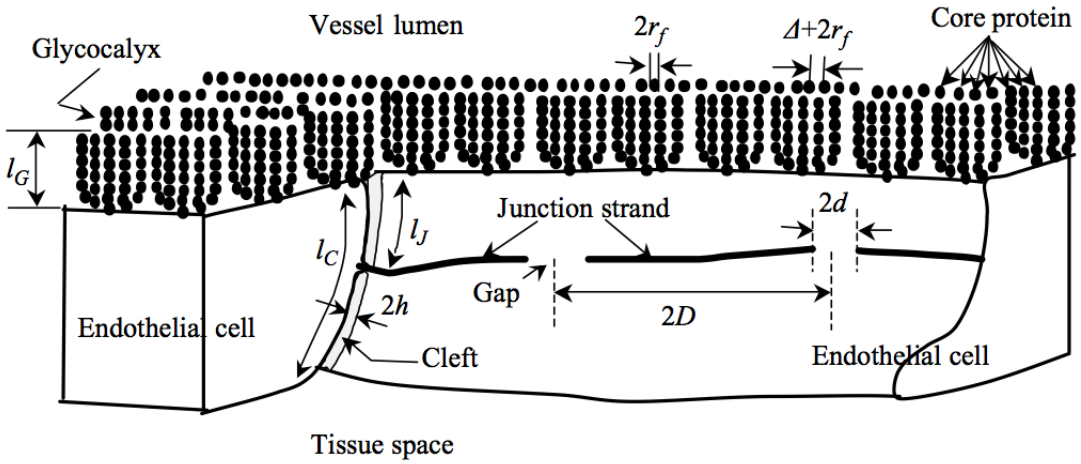


Figure 3.1: ***EC Model by Sugihara et al.:*** *Sketch of an inter-endothelial cleft and surface glycocalyx of a microvessel wall. [146].*

Fu et al. [147] developed an electro-diffusion model to account for the negative charge of the endothelial surface glycocalyx on microvessel permeability. Their model predicts that un-even cation/ion concentration in the glycocalyx due to the presence of negatively charged proteins would result in resisting positively charged proteins more so than negatively charged ones.

Employing a thermodynamic model and Monte Carlo simulations, Agrawal et al. investigated the role of the EGL in nanocarrier-cell adhesion by calculating binding affinities for nanocarriers and the EGL [148]. Their model couples parameters of glycocalyx resistance, flexural rigidity of receptors and the receptor-ligand bond stiffness. Their model was able to quantitatively reproduce the experimental binding affinities with zero fitting to binding data. In a recent study, Sumets et al. [149] developed a boundary-integral representation for biphasic mixture theory and applied it as a model for the EGL. Specifically, their model describes the motion of a rigid particle through a microtube with non-uniform wall shape, filled with viscous New-

tonian fluid and line with a thin poroelastic layer (the EGL representation). This model does not account for electrochemical effects within the EGL and is restricted to a 2-dimensional investigation.

3.1.2 Mechanosensing models

An early 2-dimensional model describing mechanosensing through the EGL was implemented by Secomb et al. as early as 2001. They used a theoretical model to analyse the transmission of fluid shear stress through the EGL from flowing blood to the endothelium. The layer was assumed to be 1 micron in thickness, and to consist of strands extending from the endothelial surface. The tension in the strands per unit membrane area was assumed to be 20 dyn/cm^2 as inferred from experiments described in Chapter 2, Section 2.2.3, to balance the assumed increase $\Delta\pi p$ in colloid osmotic pressure within the layer. The tension was assumed to be uniform along most of the strand length. Both $\Delta\pi p$ and the strand tension were assumed to decrease rapidly to zero at the outer edge of the layer. According to the model, under an applied shear stress, the strands are displaced from the perpendicular to the layer, and transmit shear stress to the cell surface [101].

Recently, Tarbell et al. used a similar model as the one by Secomb et al. as well as Brinkman's equation for each Darcy media layer, in order to estimate both the fluid and solid shear stresses on the surface of embedded cells (not ECs) exposed to interstitial flow within the extracellular matrix. The model took into account a glycocalyx layer, which they are also assumed to possess [150]. Through this model they concluded that solid stresses on suspended cell surfaces can easily be one to two orders of magnitude larger than fluid shear stress on the cell surface. For earlier models of flow within the EGL, see a comprehensive review by Weinbaum [13].

A multicomponent, multicell attempt to model the endothelium was made by Dabagh et al. [151], in the effort to investigate the shear stress amplification experienced by ECs in regions of the apical cell membrane where suspected mechanosensors

such as the glycocalyx layer, actin cortical layer, focal adhesions (proteins located in cell-ECM junctions) and adherens junctions (proteins located at cell-cell junctions), are present. The model predicts an amplification of 250-600-fold in focal adhesions and 175-200-fold for adherence junctions in comparison to apical values for ECs exposed to a mean shear stress of 10 dyn/cm^{-2}. Their model further predicts that the magnitude of the cell-cell junction inclination angle is independent of the cytoskeleton and the glycocalyx.

3.1.3 EGL modelled as a polymer brush

The materials family of polymer brushes have numerous applications (improvement of lubrication properties when the surfaces are exposed to shear [152], protection of the surface against adsorption of nanoparticles or proteins [153], etc.). Polymer brushes are formed from flexible linear macromolecules tethered at one chain end to a solid substrate, forming a dense layer of polymer chains which are more or less stretched in the direction perpendicular to the substrate surface. The types of applications as well as structural similarities point towards a possible understanding of the EGL as a biological polymer brush.

These materials are exposed to shear due to flow of the solvent and therefore the dynamic non-equilibrium response of brushes to external perturbation has been investigated through various theoretical and computational approaches. For an extensive review of the literature on this topic, see Binder et al. [154].

The glygocalyx as polymer brush model was explored experimentally by Lanotte et al. They lined synthetic microcapillaries with a biomimetic polymer brush, "glygocalyx-like", coating in order to study flow resistance as a result of the glycocalyx. Their findings support the hypothesis of the glygocalyx layer being the main factor responsible for the higher flow resistance found in microcapillaries *in vivo* [155].

Deng et al. developed a polymer brush-like coarse-grained model for the EGL

[156] and compared it with a non-biological polymer brush model. They employed the dissipative particle dynamics (DPD) method which is frequently used to model soft matter and fluid flow, over longer time and length scales as opposed to fully atomistic models.

This approach allows a polymer brush model as well as a glycocalyx model to be composed of multiple bead-spring segments or polymer/glycocalyx components (N=36 and N=21 respectively). The tethered bead of the brush-polymer was free to rotate; this was not the case for the EGL model. Grafting for both models was uniform. The EGL bead-spring model included an extra bending energy term to allow for a semi-flexible chain. Flow was imposed via movement of one of the channel walls in the case of the polymer brush (shear-driven flow), whereas in the case of the EGL model, a force on each solvent particle was the driver of flow (pressure driven flow).

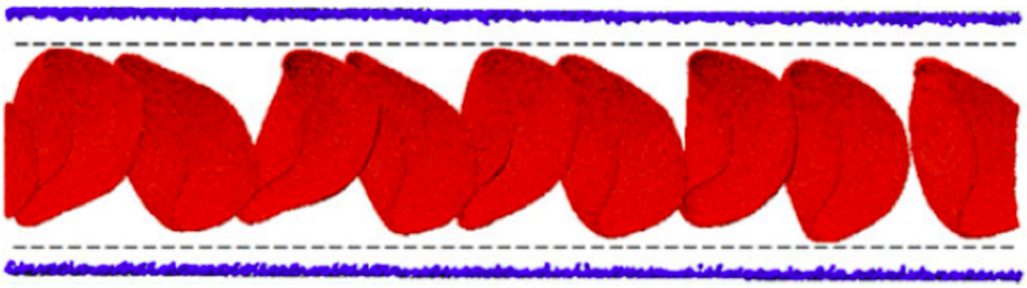


Figure 3.2: ***DPD simulation of blood flow in a tube by Deng et al.:*** *Deformable red blood cells, shown in red, modelled explicitly. Shown in blue are the glycocalyx fibres; while the dashed lines represent the edge of the cell-free layer [156].*

Their model predicts a larger (by a factor of 10) decrease in the height of the EGL as opposed to the height of the polymer brush, attributed to the EGL's flexural rigidity (which was 0 in the case of polymer brushes) due to shear. It is worth re-stating that flow was imposed with a different method for each model.

Dang et al. run further DPD simulations of blood flow in a small tube of diameter D =10 μm, coated with a glygocalyx layer. The blood was modelled as monodisperse, neutrally buoyant, deformable red blood cells (RBCs) in a Newto-

nian fluid. A cell free layer developed as the RBCs migrated towards the tube centre (as shown in Figure 3.2) and there was no RBC - EGL interaction.

3.2 Atomic-scale Models

A representative atomistic model of the endothelial glycocalyx layer would consist of a complex glycocalyx molecule model attached to a bilayer in non-equilibrium (flow) conditions. The ensemble of these elements is essential and has not, to this date, been implemented in molecular simulation experiments. There have been attempts to model specific features of the system separately, as discussed in this section.

An attempt to explicitly model the glycocalyx molecular structure and its interaction with water and surrounding ions was made by Eriksson et al. [157]. They performed classical-mechanical molecular dynamics simulations to study solvation properties of simple cations in explicit water[1]near single and multiple oligosaccharides, as a simplification of the glycocalyx.

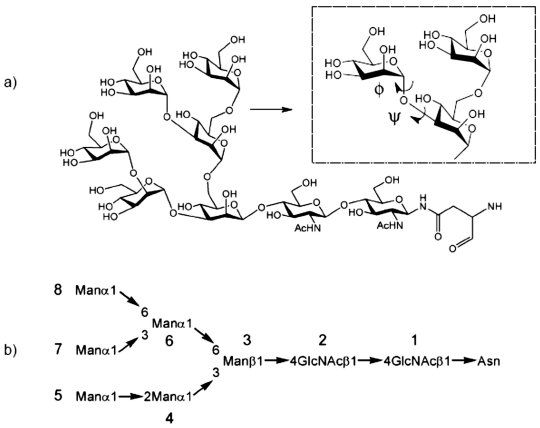


Figure 3.3: ***Structure of the branched asparagine-linked*** N***-glycan*** $Man_6GlcNa_{c_2}$ ***used for modelling studies in Eriksson et al. model:*** *(a) in detail, together with a zoomed-in part showing the two dihedral angles* ϕ *and* ψ *describing the Man(7)(*α*1-3)Man(6) glycosidic linkage and (b) in its abbreviated form with used numbering of monosaccharide residues [(1)-(8)] indicated [157].*

For their modelling studies, Eriksson et al. selected an octasaccharidic branched

[1]Alternative to the computationally cheaper method of using implicit solvent, in which the solvent is simulated as a continuous medium.

N-glycan of the high-mannose type. Its structure can be seen in detail in Figure 3.3. The glycan was tethered to the $z - y$ plane by restraining the coordinates of four asparagine-end atoms. They performed simulations of a single tethered glycan as well as an array of 9 glycans. The 3D configuration and arrangement used in Eriksson et al. [157] can be seen in Figure 3.4.

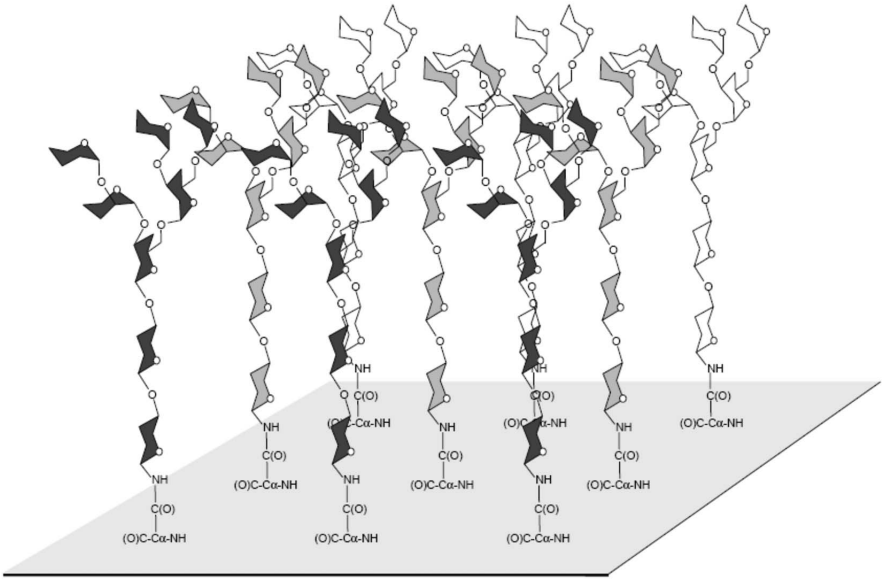


Figure 3.4: ***Glycocalyx Array Set-up by Eriksson et al.:*** *Graphical display of the array of N-linked* $Man_6GlcNa_{c_2}$ *oligosaccharides [157].*

Although in Eriksson et al. focus on the hydration of ions and the relative changes in the presence of the saccharides instead of the bulk of water, their system configuration is a good starting point for the modeling of the glycocalyx for the purpose of this project as well. Indeed, we have the same methodology, N-glycan model and molecular modelling software for both the simulation and the processing of the results, as we shall discuss in Chapters 4 and 5.

In recent years, the most detailed work on modelling GAGs and proteoglycans, with applications beyond the accurate modelling of the glycocalyx but covering the broad spectrum of structures and functionality, has been made and refined by Sattelle et al. Starting with atomistic models of the heparanome (heparin and HS), a coarse-grained model is developed based on detailed parameters derived from the fully atomistic and explicit water model [20]. The aim is to explore large GAG

structures in realistic time-scales, including the full cellular membrane environment. A schematic illustrating the ambition driving this approach is shown in Figure 3.5.

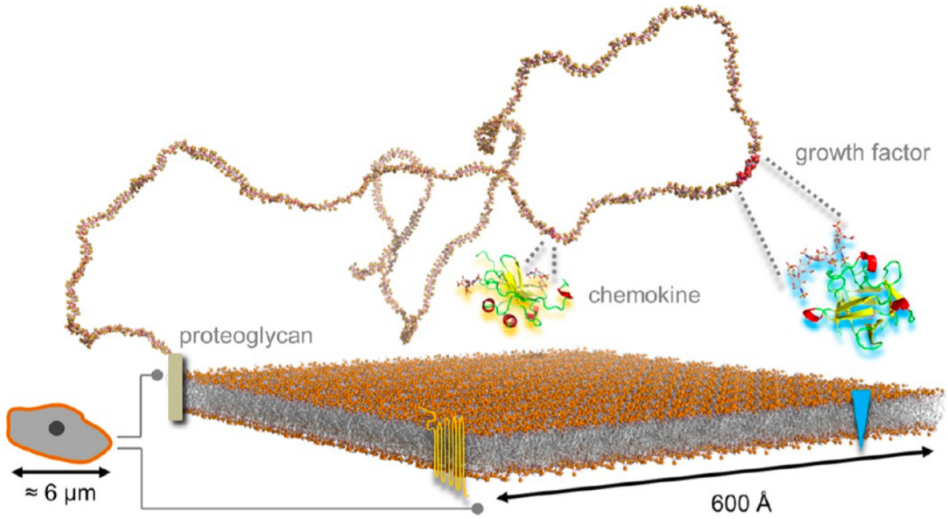


Figure 3.5: ***Schematic mammalian extracellular environment including an example heparan sulfate chain containing 800 residues and extracted from a 5 μs coarse-grained simulation:*** *Mammalian cell surfaces exhibit large sulfated carbohydrates (e.g., heparan, chondroitin, dermatan, and keratan sulfates), which stem from transmembrane proteoglycans and interact with numerous proteins via their sulfate-decorated domains. A lipid bilayer and HS binding proteins, a chemokine and basic fibroblast growth factor, are shown (not to scale). The physical dimensions of a typical mammalian cell are illustrated (left). [20].*

This approach was recently further expanded to cover other GAGs, such as HA, dermatan sulfate and chondroitin sulfate [158]. It is the first attempt to construct a realistic GAG model that can accurately characterise long-chain properties, a feat that has thus far remained elusive for both experimental and computational approaches.

The first, (and to our knowledge the only) attempt at modelling a fully atomistic glycocalyx system, alongside the transmembrane protein (Syndecan-4), GAGs (represented by HS) embedded within a lipid bilayer was recently made by Cruz-Chu et al. [159]. They employed a mixture of implicit and explicit solvent approaches, as well as the implementation of flow, to equilibrate and study parts of the system as well as all the components in place simultaneously. Due to the size and complexity of the experimental set-up, simulation times were less than or equal to a maximum

of 5 nanoseconds, when using explicit solvent. This time scale is not large enough to produce converged dynamics, as the glycan will not have managed to explore even the most frequently occurring conformations. The fully atomistic system developed by Cruz-Chu et al. can be seen in Figure 3.6.

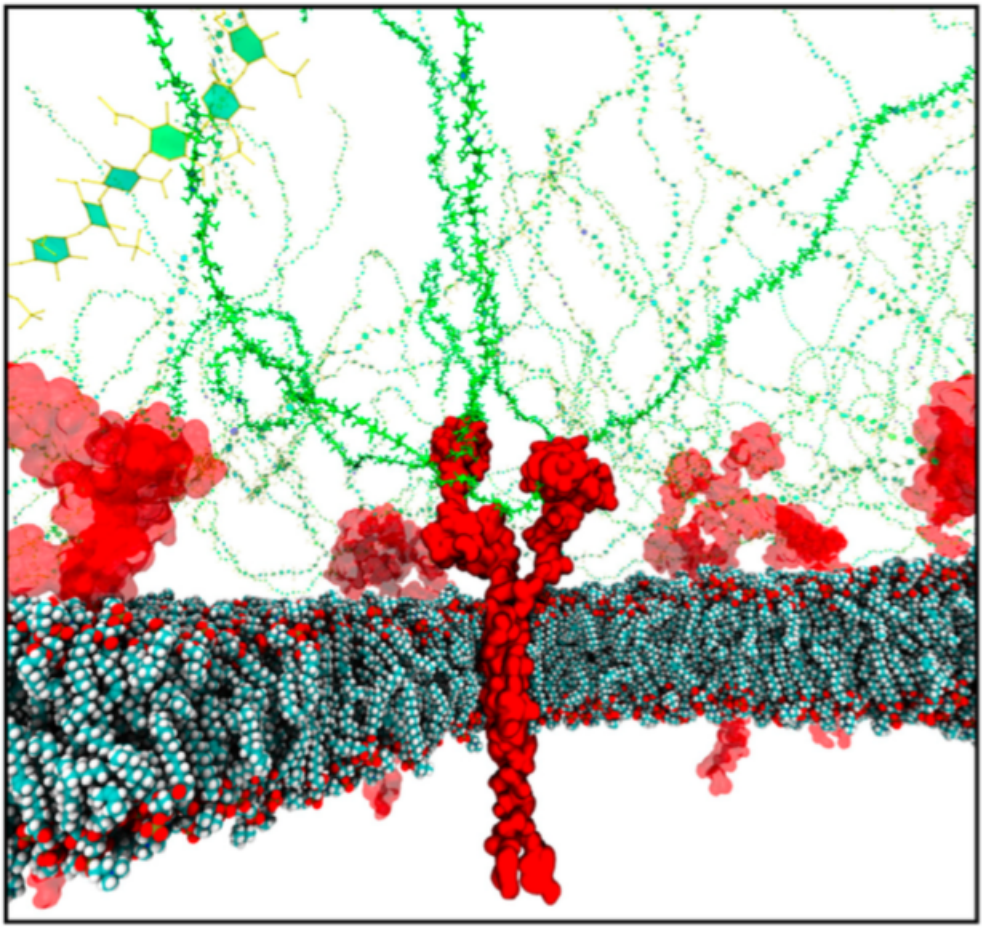


Figure 3.6: ***All-atom model of the glycocalyx.*** *All-atom model of the glycocalyx. (Green licorice representation) Sugar chains. (Red surfaces) Transmembrane protein; (van der Waals beads) lipid bilayer. Water and ions are not shown [159].*

Their results reveal the folding of proteoglycan ectodomain and the extended conformation of heparan sulfate chains. Furthermore, they studied the glycocalyx response under WSS and its role as a flypaper for binding fibroblast growth factors (FGFs). Their simulations show that the glycocalyx increases the effective concentration of FGFs, leading to FGF oligomerisation, and acts as a lever to transfer mechanical stimulus into the cytoplasmic side of endothelial cells. However, we consider the flow regime they implemented as highly unrealistic, and discuss the

implications in the following section.

3.3 Flow in Molecular Simulation

3.3.1 Flow Implementation methods

Several methods have been developed for imposing flow in MD simulations. Cieplak et al. [160] used a pressure gradient in order to simulate flow and study the boundary conditions at a fluid-solid interface. Temperature gradient-induced flow was employed by Shiomi [161] et al. who investigated the influence of the non-equilibrium thermal environment on a water cluster within a single wall carbon nanotube. Chemical potential gradients have also been implemented when two domains of different chemical potential are separated by a permeable medium [162] or micropores [163].

An early method for implementing Couette or Poiseuille flow for systems of simple fluids over surfaces with atomic structure is to impose an acceleration ("gravity"-like forcing) towards the direction of the flow [164] [165]. For the study of the flow of methane in a carbon nanotube, Sokhan et al. [166] used an acceleration of 0.5×10^{12} ms^{-2} to attain a mean velocity of 57 ms^{-1}. However, this procedure may result in low frequency oscillations of the mean flow for two-dimensional Poiseuille flow of methane between graphite surfaces [167].

To maximize the mean flow oscillation frequency for a system of water forced through single-wall carbon nanotubes, Kotsalis et al. [168] proposed an alternative adaptive scheme in which the imposed acceleration is computed dynamically to secure the desired mean flow. Another variation of the external acceleration method is to apply the gravitational force only on the intlet of the flow channel to drive the flow [169].

Recently, To et al. [170] designed a novel pressure driven flow method. By modifying the inlet and outlet boundary conditions while preserving the difference in squared velocity between the ingoing and outgoing molecules of the simulation domain, periodic boundary conditions are not violated. Their method was compared and deemed in agreement with analytical Navier-Stokes solutions for systems in plane Poiseuille flows. Thus, they developed a method which could be used to simulate pressure driven flows and modelling compressibility effects.

3.3.2 Applications to Biological Systems

Lipid bilayers are the main structural elements of all cell membranes and have therefore been extensively studied in different contexts. A recent study of bilayer deformation due to flowing liquid was made by Hanasaki et al. [171]. It is a coarse-grained MD simulation which does not take into account any type of glycocalyx model and focuses only on the deformation of a bilayer directly exposed to water flow. Coarse-graining was deemed necessary in order to achieve computational times relevant to the response of cell membranes to flow shear.

In the Hanasaki study, shear rates of 13 ns^{-1} led to clear undulations and eventual rupture of the membrane. The bilayers appeared to exhibit instabilities that reveal both a liquid and solid behavior. These results are consistent with experimental studies [172] in cells that are not equipped with a glycocalyx structure as they are not exposed to liquid flow *in vivo* (they are not ECs) and therefore, extreme values of shear stress lead to necrosis.

The effect of flow-induced shear on blood clotting was investigated by Zou [174] and Chen [173] at the molecular level, in two studies that measured the effect of flow on the Iba β-switch glycoprotein, which has been found to undergo β-hairpin folding, thus initiating a cascade of biomechanical signalling that leads to coagulation (for a schematic, see Figure 3.7). In both studies, flow was generated by applying a constant force to the oxygen water molecules along the x-direction. However, the

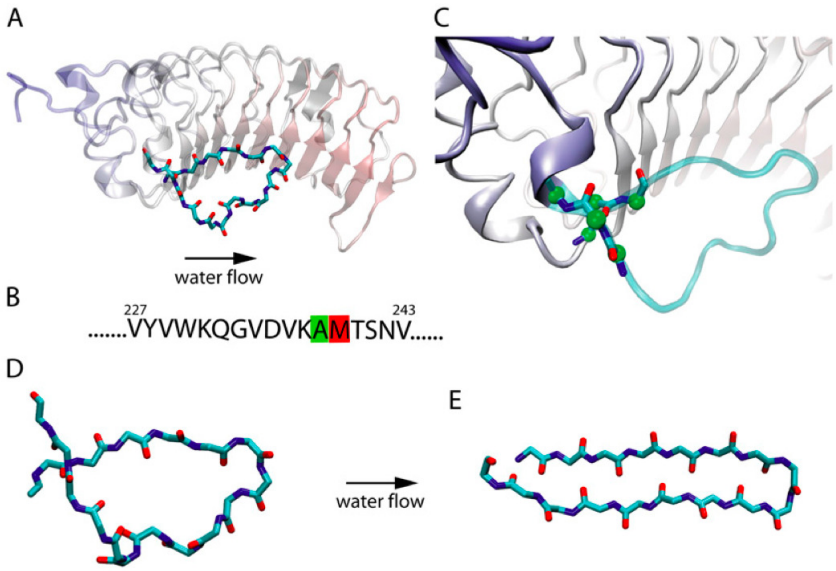


Figure 3.7: ***Schematic of beta-switch glycoprotein:*** *β-switch region of GPIb. (A) α-subunit of GPIb, a platelet surface receptor glycoprotein involved in tethering of blood platelets. Highlighted in solid licorice representation is the actual β-switch, while the remainder of the protein subunit is shown in transparent cartoon representation. (B) Amino-acid sequence of the b-switch with the two amino acids involved in loss and gain of function mutants (see text) highlighted in green and red, respectively. (C) Connection region between b-switch and GPIbα. Backbone of the residues connecting the b-switch to the GPIbα is highlighted in solid licorice representation. Green spheres indicate carbon atoms constrained in the simulation to their crystallographic positions. (D and E) Schematic presentation of the flow-induced transition from a flexible loop β-switch (D) to a β-hairpin β-switch (E) [173].*

magnitude of the achieved flow rate (up to $50m/s$) is highly unrealistic given that such high blood flow velocities are not realistic for even the most severe of wounds.

In the recent glycocalyx modelling study described in the previous section, Cruz-Chu et al. [159] applied a water velocity of up to $10m/s$, which induced a shear stress of up to 0.8 MPa. Such a magnitude of flow and shear stress would be unrealistic, even for the centre of large vessels, more so for near-wall conditions. The authors also fail to investigate the impact of shear in the conformational properties of the glycans themselves as well as the ring-shape for each component of HS. It is likely that high flow values would induce non-physiological deformation on the overall glycan topology and hexose rings.

3.4 Feasibility of complete atomic scale models

The focus of this thesis is exclusively on the EGL in the atomic scale, and this is reflected in the four following chapters. However, we started our investigation of the EGL at the biologically relevant scale of vascular physiology in Chapter 2 and reviewed several continuous and coarse-grained models in the present chapter.

Despite advances in algorithms and computation, an all-encompassing, fully atomistic representation of the EGL remains elusive and will continue to do so for the foreseeable future. We therefore believe that atomistic glycocalyx models should focus on a few particular aspects of its structure and function. In this way they can be well designed to converge for a set of pre-defined characteristics and only then can they render meaningful and useful results that can be validated and incorporated in large-scale models.

3.5 Overarching aims of the thesis

The work presented in chapters 5 and 6 was undertaken in accordance with the following primary aims:

- **Aim 1:** To determine whether and how, physiological flow can effect the properties of the GAG components of the EGL.
- **Aim 2:** To investigate whether and how physiological flow can effect the anchoring glycoprotein and whether shear stress can be transmitted via the glycoprotein to the cytoskeleton.
- **Aim 3:** To assess the impact of suflation, and therefore charge, on the compressibility of the EGL
- **Aim 4:** To investigate novel ways of calculating the effect of sulfation based on individual building blocks of the EGL.

In accordance with aims 1 and 2 as presented above, a fully atomistic glycocalyx in physiological flow model was constructed and MD simulations took place in order to implement the following tasks:

- **Task 1:** Add physiological flow to existing atomistic glycocalyx model and evaluate the resulting velocity profile.
- **Task 2:** Establish the validity of the above model by performing fully converged simulations and comparing to gold-standard experimentally derived glycan conformational properties.
- **Task 3:** Investigate the orientation and conformational properties of the glycocalyx under the effect of flow and compare with a diffusion-only model.
- **Task 4:** Investigate the conformational properties of the anchoring glycoprotein (represented by the terminal amino-acid connecting the glycan component of the glycocalyx) and compare with the diffusion-only model as above.

In accordance with aims 3 and 4 as presented above, a fully atomistic glycocalyx as represented by sulfated and unsulfated HS glycans was constructed and MD simulations took place in order to implement the following tasks:

- **Task 1:** Establish the validity of the above model by performing fully converged simulations and comparing to gold-standard experimentally derived glycan conformational properties.
- **Task 2:** Using a novel enhanced sampling method, calculate the free energy of closely packed monomer components of HS as a function of distance between their centre of mass.
- **Task 3:** Using the same enhanced sampling method, calculate the free energy of a system of closely packed GAGs as a function of density (represented here as distance between two HS strands). Compare between sulfated and unsulfated GAG regions.

- **Task 4:** Investigate the effect of close proximity on the conformational properties of HS and its constituent monomers.

- **Task 5:** Construct and evaluate a model for the calculation of free energy of sulfated and unsulfated strands of HS based on the types of monomers it is composed of.

Chapter 4

Molecular Dynamics Modelling

This chapter focuses on the MD method for atomistic simulations and provides a brief overview of statistical mechanics and methods typically used to validate computational models using experimental results. The computational tools used in this work are also presented in this chapter, alongside certain technical and algorithmic underpinnings. The chapter concludes with a more detailed discussion of an enhanced sampling method, the adaptive biasing force.

4.1 Classical Molecular Dynamics

MD simulation is a technique for computing the equilibrium and transport properties of a many-body system. In *classical* MD, the nuclear motion of the constituent particles obeys the laws of classical mechanics (i.e., quantum effects are excluded) [175]. In the case of atomistic MD, a molecule is represented as a mechanical system in which the particles (atoms) are point masses which exhibit bonded interactions of a harmonic nature. The molecule can experience rotation, translation and vibration, exploring conformations in space as a response to the sum of inter- and intramolecular forces collectively acting upon it.

4.1.1 Force Fields

Force fields enable the prediction of molecular structures and potential energies using empirical parameters to describe electrostatic and van der Waals interactions, bond and angle vibrations, and internal torsion rotations (Eq.(4.1)). These empirical potentials are parameterised based on regressions performed on both experimental data and quantum mechanical calculations, so that they describe accurately the interatomic interactions and system thermodynamics over a broad range of conditions [176] [177]. Some force fields are developed to be applicable to a wide range of molecules, while others describe particular molecular classes such as proteins, nucleic acids, lipids or carbohydrates.

Forces are expressed as a sum of harmonic-like terms for bond-length and bond-angle deviations from reference values (Eqs. 4.2 4.3); trigonometric torsional terms to account for internal rotation (Eq. 4.4) and non-bonded van der Waals (Eq.4.6) and electrostatic potentials (Eq.4.5).

$$\begin{aligned} V(r_1, ..., r_N) = & \sum_{i<j} V_{Bonds}(r_{ij}) + \sum_{i<j<k} V_{Angles}(\theta_{ijk}) \\ & + \sum_{i<j<k<h} V_{Torsions}(\phi_{ijkh}) \\ & + \sum_{i<j} V_{Coulomb}(r_{ij}) + \sum_{i<j} V_{VdW}(r_{ij}) \end{aligned} \tag{4.1}$$

where

$$V_{Bonds}(r_{ij}) = \frac{1}{2} k^b_{ij} (r_{ij} - r^0_{ij})^2 \tag{4.2}$$

$$V_{Angles}(\theta_{ijk}) = \frac{1}{2} k^\theta_{ijk} (\theta_{ijk} - \theta^0_{ijk})^2 \tag{4.3}$$

$$V_{Torsions}(\phi_{ijkh}) = \frac{1}{2} k_\phi (1 + cos(n\phi + \gamma)) \tag{4.4}$$

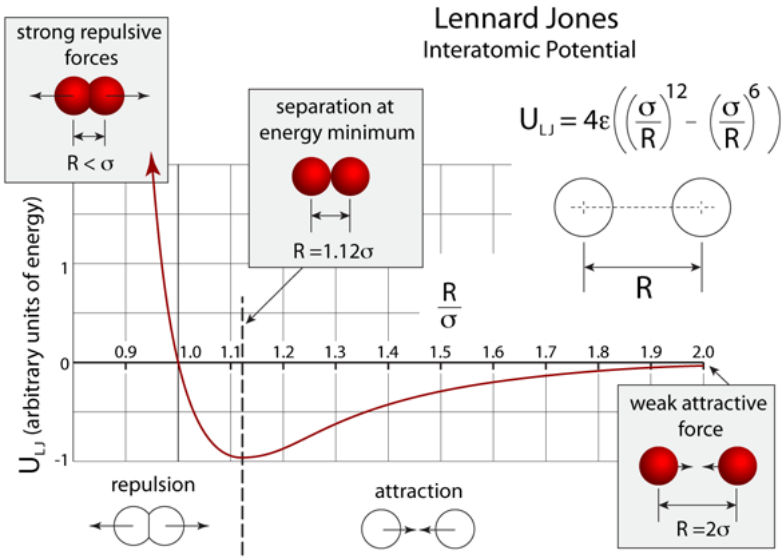


Figure 4.1: ***Lennard-Jones potential.*** *This graph depicts the Lennard-Jones potential function indicating regions of attraction and repulsion. At low temperatures atoms achieve the minimum potential energy at the bottom of the potential curve. When the atomic separations are to the left of the minimum the atoms repel, otherwise they attract one another [178].*

$$V_{Coulomb}(r_{ij}) = \frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{r_{ij}} \tag{4.5}$$

$$V_{VdW}(r_{ij}) = 4 \in_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{4.6}$$

where r_{ij} the distance between the pairs of atoms i and j, q_i and q_j the electric charge of atoms i and j respectively and ϵ_0 is the electrical permittivity of space.

The expression for the Van der Waals potential in Eq.(4.6) is given by an empirical equation called the Lennard-Jones potential [179] [180]. It is a model that describes pairwise interactions between neutral atoms, consisting of a long-range attractive (due to van der Waals, or dispersion forces) and a short-range repulsive component, approximating the Pauli repulsion well due to overlapping orbitals. The Lennard-Jones well depth ($\in_{ij}$) and distance at which the inter-particle potential is zero (σ_{ij}) characterise the non-electrostatic interaction between two atoms (see Figure 4.1).

Most classical force fields share the functional form of (4.1), whereas they differ in the philosophy and approach behind the determination of parameter values of the individual terms (Eqs. (4.2) - (4.6)) as well as the detail in which they attempt to describe interatomic interactions. This is particularly evident in molecules such as solvents, and in particular, water, where several potential functions have been implemented. The atomistic modelling of water is particularly important in simulations of biological systems, as it is the solvent that is ever-present in all physical, biological and chemical interactions.

The most commonly used water models are: the SPC [181], TIP3P and TIP4P [182] which became more widely used from 1983 onwards. The TIP3P and SPC models follow the minimalist TIPS (transferable inter-molecular potentials) form [183] of having just three interaction sites centered on the nuclei. Each site has a partial charge for computing the intermolecular Coulomb energy, and the only additional term is a Lennard-Jones interaction between oxygen molecules. For the TIP4P model, the site of negative charge is moved off the oxygen to a point 0.15 along the bisector of the H-O-H angle (this extra site often referred to as "dummy atom"). More recently 5-site models have also been developed, such as TIP5P [184]. In this case, the negative charge of the dummy atoms is representing the lone oxygen pairs in a tetrahedral-like geometry. Although more accurate, the use of 4 and 5-site models is avoided if possible in most applications as they are computationally expensive.

4.1.2 Statistical Mechanics

Many-body statistical systems are so complex that one cannot obtain all the information to completely characterise them. This is why statistical mechanics deals with ensemble averages. A statistical ensemble is an idealisation consisting of a large (some times infinite) number of copies (microscopic states) of a system, considered all at once, each of them representing a possible macroscopic state that the real

system might assume, given known macroscopic properties, such as temperature, energy, volume, etc.

For the canonical ensemble, in which the temperature T and number of particles N_m are fixed, the equilibrium average of some quantity X is expressed in terms of phase-space integrals, involving the potential energy $U(\mathbf{r}_1, ...\mathbf{r}_{Nm})$,

$$< X >= \frac{\int X(\mathbf{r}_1, ...\mathbf{r}_{Nm}) e^{-\beta U(\mathbf{r}_1, ...\mathbf{r}_{Nm})} \, d\mathbf{r}_1 ...\mathbf{r}_{Nm}}{\int e^{-\beta U(\mathbf{r}_1, ...\mathbf{r}_{Nm})} \, d\mathbf{r}_1 ...\mathbf{r}_{Nm}} \tag{4.7}$$

where $\mathbf{r}_i \mid i = 1, ...N_m$ are the coordinates, $\beta = 1/k_B T$, T is temperature and k_B is the Boltzmann constant. This average corresponds to a series of measurements over an ensemble of independent systems. The denominator at the right side of the equation is the canonical partition function Z, which for a system that finds itself in exact states (microstates) of total energy E_s can be expressed as:

$$Z = \sum_s e^{-\beta E_s} \tag{4.8}$$

The partition function can be related to thermodynamic properties because it encodes the statistical properties of the system. The probability P_s that a system occupies a micro state s is

$$P_s = \frac{1}{Z} e^{-\beta E_s} \tag{4.9}$$

thus, given Equation 4.8, the partition function acts as a normalising constant:

$$\sum_s P_s = \frac{1}{Z} \sum_s e^{-\beta E_s} = \frac{1}{Z} Z = 1 \tag{4.10}$$

Molecular dynamics simulation produces averages of the form

$$< X >= \frac{1}{M} \sum_{\mu=1}^{M} X_\mu(\mathbf{r}_1, ...\mathbf{r}_{Nm}) \tag{4.11}$$

over a series of M measurements made as the system evolves.

Assuming that the sampling of the MD algorithm is sufficiently thorough to capture completely the typical behaviour of the system, meaning that the most probable configurations are sampled, the two kinds of averaging described by the equations (4.7) and (4.11) will be identical.

The ergodic hypothesis relates the ensemble average to measurements carried out for single equilibrium systems during the course of its natural evolution [185]. It states that all accessible macrostates are equiprobable over a long period of time.

For the purpose of this work, the Isothermal - Isobaric (NPT) Ensemble was used. In NPT simulations, as hinted by its name, the number of particles (N), pressure (P) and temperature (T) of the system remain constant. Its characteristic state function is the Gibbs Free Energy:

$$G(N, P, T) = -k_B T \ln \Delta(N, P, T) \tag{4.12}$$

where $\Delta(N, P, T)$ is the partition function for the isothermal - isobaric ensemble, and can be expressed as a weighted sum of the partition function of the canonical ensemble described by equation (4.8).

4.1.3 Computational Considerations

Minimisation Unless simulations begin with a pre-equilibrated system, usually a method of energy minimisation needs to be applied prior to running an MD algorithm. Energy minimisation (or energy optimisation, or geometry optimisation) is a process of finding a spatial arrangement of the system's particles which, according to the force-field parameters, minimises the potential energy of the system. A frequently utilised algorithm for minimisation is the conjugate gradient algorithm [186].

Conjugate gradient makes use of the previous history of minimisation steps as the current gradient to determine the next step. The step-size can be adjusted during

the steps of the procedure using a line search algorithm. The method converges to the minimum energy in $O(n)$ steps for quadratic energy surfaces, where N is the number of degrees of freedom [187].

Cut-off and switching functions

Given the huge proportion of the computational cost associated with the calculation of interatomic forces, potential functions are truncated at some cutoff radius r_c rather than continuing to calculate attractive interaction of the long tail as $r \to \infty$. Truncation serves to reduce the number of pairwise interactions acting on each atom and therefore can decrease the computational complexity of force evaluation [185] in the case of van der Waals and electrostatic interactions. Switching functions are often applied at the point of truncation, in order to "ease" the transition to 0 van der Waals potential.

Periodic Boundary Conditions To eliminate surface interaction of water molecules, periodic boundary conditions are frequently used in molecular simulation. This method involves surrounding the system under study with identical virtual unit cells (see fig. 4.2). The atoms in the surrounding virtual systems interact with the atoms in the real system. Potential forces between pairs are calculated for the closest instances to one another.

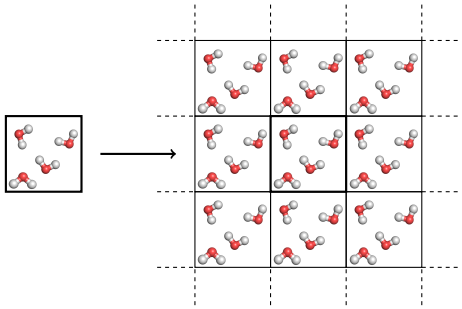


Figure 4.2: ***Periodic boundary conditions.*** *Example of water molecules in a box with applied boundary conditions in 2-D.*

Integrator

The numerical integrator used for this work is the velocity Verlet integrator [188] [189]. In the numerical analysis literature, the Verlet method is also knows as the "explicit central difference method".

According to the implementation of the velocity Verlet algorithm, given positions x_n and velocities v_n at time t_n, the positions x_{n+1} and velocities v_{n+1}

$$x_{n+1} = x_n + v_n \Delta t + \frac{1}{2}\alpha_n (\Delta t)^2 \tag{4.13}$$

$$v_{n+1} = v_n + \frac{1}{2}(\alpha_{n+1} + \alpha_n)(\Delta t) \tag{4.14}$$

where α_n the acceleration at time t_n.

Since the major computational bottleneck in molecular dynamics simulation is the calculation of forces and energy, the actual speed of the integrator of the equations of motion is not the most important criterion in choosing this component of the MD algorithm. Instead, the accuracy and ease of implementation tend to be more important deciding factors.

The method of integration does, however, have an indirect impact on the speed of calculation by setting the length of the time step, which in turn, determines the number of time steps required to produce a given amount of simulation time. Selection of the appropriate time step that saves calculation time but produces accurate results always depends on the specifics of the system one attempts to simulate. In systems with widely varying time scales, it can be advantageous to use different time steps for the various length scales that make up the system [176].

Parallelization

In the practical implementation of molecular dynamics, parallel computing is almost always used due to the volume of computation required. Molecular dynamics

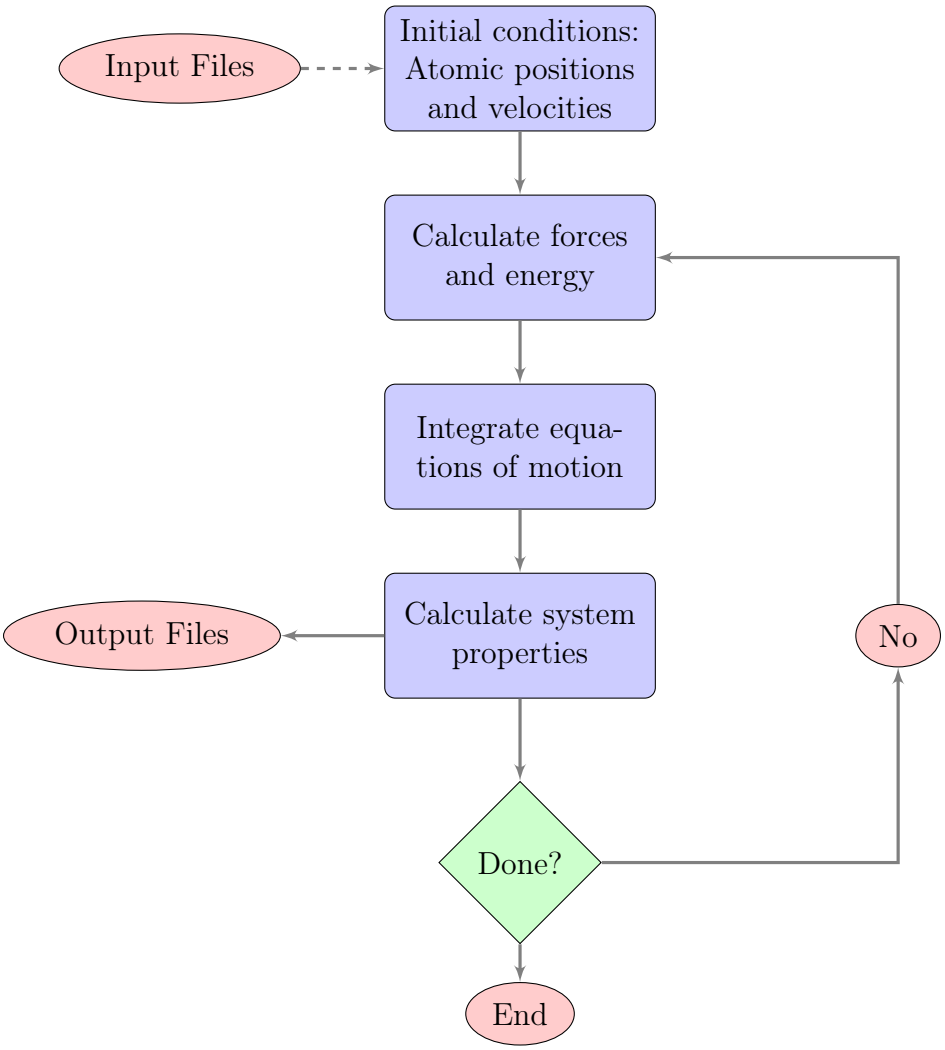


Figure 4.3: ***Molecular dynamics algorithm flowchart.*** *Each block in the diagram represents a significant amount of algorithmic implementation and computation, which could be broken down into a number of smaller steps in a more detailed block diagram.*

is well suited to parallel computation, as at each time step, forces for each molecule can be calculated independantly and thus can be performed simultaneously.

The three different approaches to parallelising MD simulations with short range forces are:

- *Atom decomposition:* Where subsets of atoms are assigned to each processor.
- *Force decomposition:* Where subsets of individual force calculations are assigned to each processor.

- *Spatial Decomposition:* Where atoms within a subdomain of the simulation space are assigned to each processor.

Most of the available simulation packages implement a variety of the above methods. Long-range forces pose extra algorithmic difficulties in parallelised systems and require sophisticated techniques of message passing and memory access between processing units. More information on the particular approaches to speed and accuracy of computations is given in the next section.

Electrostatics

The Ewald summation algorithm [190] is a method for computing full electrostatics for simulation systems with periodic boundary conditions and is very often used to manage the expensive computation of long-range electrostatics in molecular dynamics platforms. The Ewald summation algorithm separates the interaction potential $\phi(r)$ in two terms: a short-ranged part $\phi_{sr}(r)$ whose sum converges quickly in real space and a long-ranged part $\phi_{lr}(r)$, whose sum converges quickly in Fourier space:

$$\phi(r) = \phi_{sr}(r) + \phi_{lr}(r) \tag{4.15}$$

Typically, the Particle-Mesh Ewald (PME) [191] method, a fast $(N \log N)$ numerical method is employed to compute the Ewald sum. In this method, two summations are performed: a direct sum for short-range energy E_{sr} in real space and a summation in Fourier space of the long-range part:

$$E_{sr} = \sum_{i,j} \phi_{sr}(r_i - r_j) \tag{4.16}$$

$$E_{lr} = \sum_{k} \widetilde{\phi}_{lr}(k) |\widetilde{\rho}(k)|^2 \tag{4.17}$$

where $\widetilde{\phi}_{lr}$ and $\widetilde{\rho}$ represent the Fourier transforms of the potential and the charge

density. Because of fast convergence, truncation is performed with little loss of accuracy and great time-gains. In order to evaluate the charge density Fourier transform $\widetilde{\rho}(k)$, one uses Fast Fourier Transform method, such as the Cooley-Tukey algorithm [192].

PME provides physically correct electrostatics (without artefacts due to truncation) and allows a smaller short-range cut-off to be used. When using PME, the cut-off distance signifies the range within which pair-wise electrostatic interactions will be directly calculated.

Collective variables

A recent development in molecular dynamics simulations, is that it is often useful to reduce the great number of degrees of freedom into a few parameters which can be either analysed individually, or manipulated in order to alter the dynamics in a controlled manner. These have been called "order parameters", "collective variables", "surrogate reaction coordinates", and many other terms (also see the discussion in the subsequent section on the reaction coordinates in this chapter). For the purpose of this work, the term "collective variable" is used, and it indicates any differentiable function of atomic Cartesian coordinates, $\boldsymbol{x}_i$, with i ranging between 1 and N, the total number of atoms.

4.2 Implementation Specifics

Selection of the appropriate software is a very important step, when embarking into creating a new MD model. The most important decision factor is the system one will attempt to simulate, which involves taking into account its complexity and particularities. Identifying the appropriate computational tools for the system narrows down the list considerably. The next step is to consider other factors, such as, software maintenance and support, usability and user base.

4.2.1 NAMD

The software selected for this project is NAMD. It was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign [193]. NAMD is a parallel molecular dynamics code designed for high-performance simulations of large biomolecular systems. It has the capability to work with several potential (force field) functions such as CHARMM [194] (which is the default and most "native"), AMBER [195] [196] and GROMOS [197]. Surrounding NAMD is a powerful open source community alongside the main research group that develops the new releases. It is sophisticated, user friendly, constantly updated and with an active mailing list that is quick to respond to any problems the members are dealing with.

In order to improve performance, several techniques are implemented in NAMD. Parallelisation is implemented using a combination of force and spatial decomposition [198]. In this particular implementation of spatial decomposition, the simulation space is split into cubes, referred to as "patches", whose dimensions are slightly larger than the cutoff radius. Thus, atoms in one patch need to interact only with their 26 neighbouring cubes. For each pair of neighbouring patches, a non-bonded force computation object is assigned, which can be independently mapped to any processor. This approach solves the issue created by applying only spatial decomposition, which can be greatly affected by size and density of the simulation space, thus creating load imbalances.

For electrostatics calculations, NAMD employs the smooth PME method [199]. A multi-timestep method is used for the calculation of forces, which are approximated as a sequence of impulses and divided in short- and long-range, the latter being calculated less frequently. A cut-off parameter determines the distance at which switching to long-range interactions occurs.

For the present MD application, water flow needed to be incorporated in the simulation. This is not a usual addition and therefore, quite underdeveloped for most MD simulation platforms. NAMD, however, has implemented quite a clear and sensible way to add extra forces to the simulation, using TCL language-based scripting. Accounting for this and all of the above factors in mind, NAMD was the straightforward software of choice.

4.2.2 The Glycam Force Field

The latest release of GLYCAM06 force field parameters was obtained from the Research Group of Professor R.J. Woods (University of Georgia, Athens GA). These include all the essential parameters for input in the tLeaP/sLEaP program (See section 4.2.4) which is needed to create the hydrocarbon structure and to solvate it [196]. Parameterisation of GLYCAM06 employed training and test sets of nearly 100 molecules from the chemical families of hydrocarbons, alcohols, ethers, amides, esters, carboxylates, molecules of mixed functional groups as well as simple ring systems related to cyclic carbohydrates. The parameters identified are sufficient for a complete characterisation of glycans and glycoconjugates and they include atomic partial charges, bond lengths, angle and torsion parameters for each particular functional group. With GLYCAM06 one can, at the moment, create input files for hydrocarbons lipids and glycoproteins, compatible with several water models. This makes it an excellent tool for the spatial and temporal characterisation of this particularly flexible category of molecules.

4.2.3 Integrator and time step

The velocity Verlet integrator is, as described in the previous section, an explicit integrator and it comes implemented in NAMD. In MD, the use of an implicit solver can lead to a faster implementation by increasing the time step without compromising simulation stability [200]. However, the maximum time step used throughout

this work was $2fs$ as required by the GLYCAM06 force field and provided that hydrogen bonds are constrained. In an all atom simulation one is limited by the vibrational frequencies of the heavy atoms. Unless coarse graining is used for approximating this, short time steps need to be used, negating the use of an implicit solver.

4.2.4 Supporting Tools

AmberTools

AmberTools consists of several independently developed packages that work well by themselves, and with the Amber simulation software itself. Since we had to use the GLYCAM force field which is compatible with the Amber potential function and format, some of the tools were very useful. The ones that were needed for this project were tLEaP [201] and ptraj [202].

tLEaP helps set-up the input files that contain the molecular structure along with the potential function parameters needed to calculate forces. It exports information regarding the coordinates of each atom and their relationship to each other along with information on the type of atom and type of interaction they can have with their surroundings. tLEaP can be used in order to build the molecular structure of the sample. In a second stage it is used again to build the final simulation input files for the molecule array. This process is explained in Figure 4.4.

The ptraj program reads trajectory output files from MD simulations and can manipulate them in several ways as well as analyze them to calculate values such as radial distribution functions, root mean square deviation, distances in the course of time, dihedral angles etc. It proved very useful as it was the main tool for post-processing in this work.

VMD

Images of molecular structure were generated using VMD [203]. Visual Molecular Dynamics software (or VMD) provides a wide range of molecular representations, coloring styles, transparency and material properties. In addition it contains several simulation analysis tools and can be extended with multiple plugins. A command line interface is also provided via the TCL language. Alongside NAMD, VMD has been developed by the Theoretical and Computational Biophysics group at the University of Illinois and the Beckman Institute and the two are designed to work very well together.

4.2.5 Supercomputing resources

Supercomputer resources were necessary to carry out the work presented in the two following chapters. All production runs were conducted in the University College London (UCL) research computing platform "Legion" and in particular on nodes with the following specifications: Dell C6220 nodes, dual processor, eight cores per processor at 4GB RAM per core (64GB per node) and 1TB of local disk storage per node. At most 64 cores (4 nodes) at a time were used for simulations in the present work.

Early simulation runs as well as (less computationally expensive) post-processing were also conducted at the facilities of the Complex Carbohydrate Research Centre in the University of Georgia in Athens GA (USA).

4.3 Work Flow

In order to prepare, run and analyse simulations of the type pertaining to this work, we created a workflow consisting of all of the above described programs (Figure 4.4.

Preparation

All oligosaccharide sequences were generated according to the all-atom GLYCAM06 [196] parameter set for oligosaccharides and glycoproteins, using the GLYCAM web tool [204] and the tLEap component of AmberTools for further modifications where necessary.

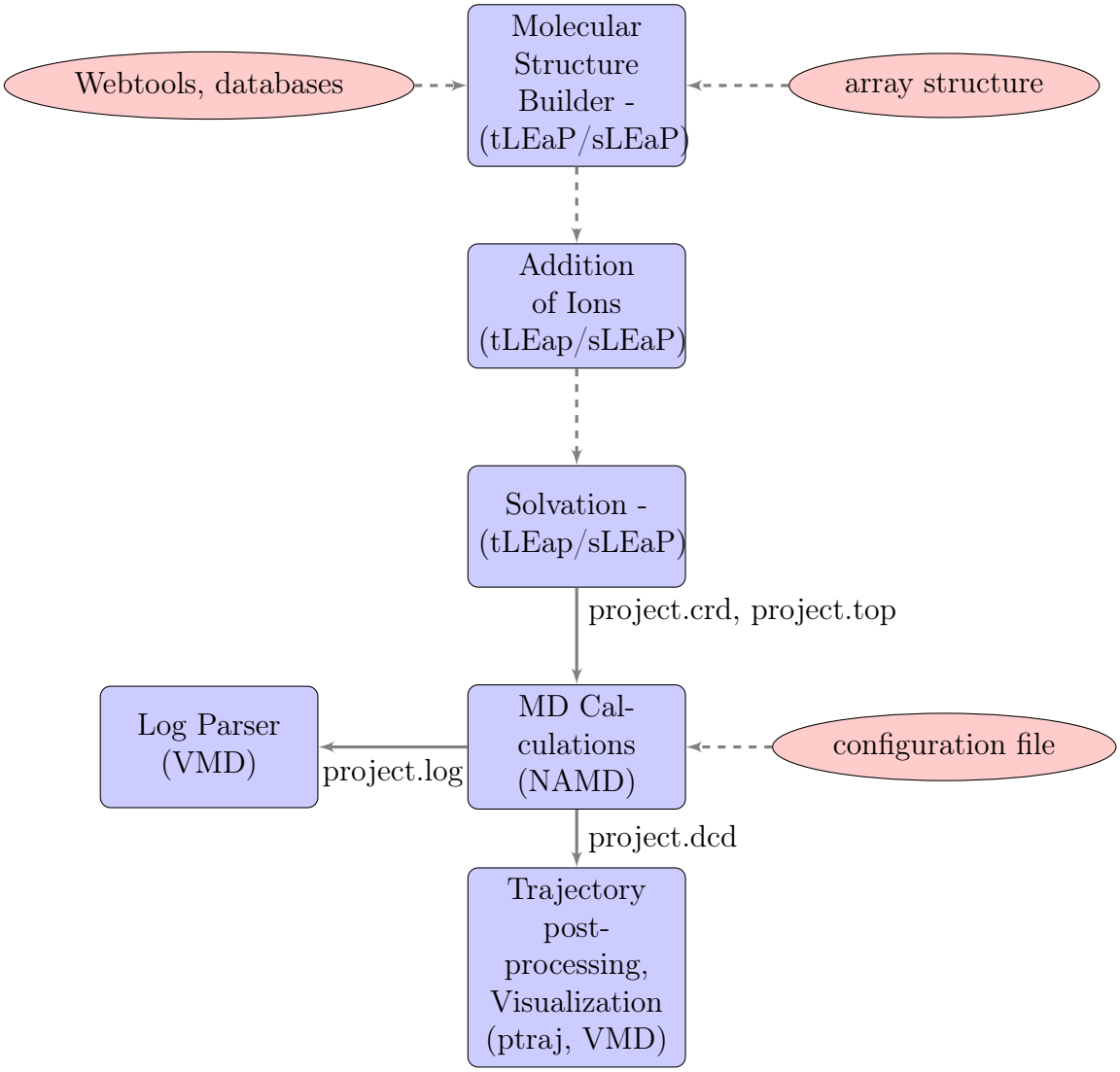


Figure 4.4: ***Work flow diagram.*** *Outline of the various stages in the process of setting up, simulating and analysing the system.*

Simulation Run

Once the molecule is set up, the simulation parameters are defined, in a simple

TCL executable configuration file. Parameters include experimental conditions, such as temperature and pressure as well as computational options for integration and dynamics, as well as added external forces. Once the configuration file is read by NAMD the simulation runs for a period of (real) time which depends on the number of steps specified and the size of the system and its properties. The work is divided in four separate stages with different parameters:

- *Minimization:* This is not a molecular dynamics step, but every MD simulation begins with a minimisation protocol, in order to avoid high energy starting configurations. The conjugate gradient/line search algorithm is used at this stage, as described in section 4.1.3.

- *Heating:* In the heating stage, molecular dynamics simulation steps take place and the temperature, starting off at the near-zero point it was left at minimization, is incrementally reassigned (by means of increasing velocities) to reach the desired one required for the equilibration stage.

- *Equilibration:* This is a core MD part of the simulation which takes place until the simulation reaches a state beyond its initialisation transients (which in MD is a constant variance of the thermodynamics properties measured).

- *Production:* The production run follows after equilibration has been achieved. In terms of MD algorithm it's exactly the same as the equilibration step, unless water flow is incorporated. Desired properties are measured only in the production stage.

4.4 Validation

It is important to always validate MD simulations using experimentally observable properties that can be calculated from simulation results, usually involving the

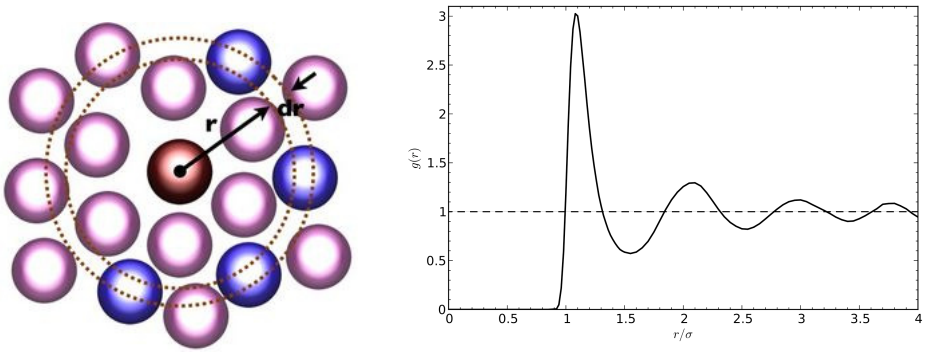
resulting trajectories. Experiments that validate MD results include X-ray scattering, crystallography and nuclear magnetic resonance (NMR) spectroscopy. The corresponding attributes that provide the link between experiment and simulation are: the radial distribution function (RDF) and coordination numbers, dihedral angle representation between linkages (presented also as Ramachandran plots) and nuclear Overhauser effect (NOE) distances. For all such results to be valid, convergence needs to be achieved. Typically in MD simulations, convergence is said to be achieved when all observed states have been sampled at least twice during the course of a simulation.

4.4.1 RDS and coordination number

In statistical mechanics, the RDF (or pair correlation function), describes how the atomic density varies as a function of the distance from one particular atom. For an equilibrated system, the RDF may be inferred from the positions of the molecules averaged over all sampled configurations [205].

The RDF for a simple liquid is measured by simply calculating the ratio between the average number density at a distance from any given atom and the density at the same distance from an atom in an ideal gas at the same overall density [175]. In practice this is calculated up to a certain cut-off of the radius and in specified space intervals (see Figure 4.5).

By calculating the RDF for certain types of atoms, one can study properties such as ion hydration and anion-cation pairing and interactions during the course of a simulation. For example, in Figure 4.5a, blue spheres could represent cations, whereas pink spheres are water molecules and the red spheres anions. In that case, the radius of the first hydration shell of the anion (and likewise the second, etc) can be directly calculated as an averaged property from the simulation. In a similar



(a) **(b)**

Figure 4.5: ***Pair correlation function:*** *(a) Schematic representation of the first and second pairwise relations for a single atom. (b) RDF for a Lennard-Jones fluid with an average number density of* $\rho = 0.844$.

manner, pair-wise anion-cation interactions can be calculated by taking into account the distances between them.

The RDF is of interest also because neutron and X-ray scattering experiments on simple fluids yield information about it. In ion containing water solutions, the structure of the first hydration shell around ions is more or less flexible and can take different conformations. In a first approximation, it can be quantified by the coordination number, i.e., the average number of water molecules in direct contact with the ions. The coordination numbers can be calculated by integrating the ion-water oxygen RDFs from zero up to the first minimum.

Coordination numbers provide an important link between experiment and simulations. However, difficulties involved in extracting radial distribution functions and coordination numbers from scattering data have resulted in measured coordination numbers in similar ranges for different types of ions. An alternative source for structural ion hydration data for force field validation is simulations using *ab initio* approaches [206].

4.4.2 Torsion angles and Ramachandran plots

Recognising that an amino acid's backbone can define a plane, Ramachandran [207] was able to define the relative orientation of two adjacent amino acids by three degrees of freedom, the torsion angles between the planes. This greatly reduced the number of terms required for defining the occupied conformational space of peptide linkages, resulting in the easily interpretable two-dimensional plots (see Figure 4.6b).

The backbone dihedral angles of proteins have been defined as: ϕ (involving the backbone atoms $C'-N-C^{\alpha}-C'$), ψ (involving the backbone atoms $N-C^{\alpha}-C'-N$) and ω (involving the backbone atoms $C^{\alpha}-C'-N-C^{\alpha}$). Thus, ϕ controls the $C'-C'$ distance, ψ controls the $N-N$ distance and ω controls the $C^{\alpha}-C^{\alpha}$ distance.

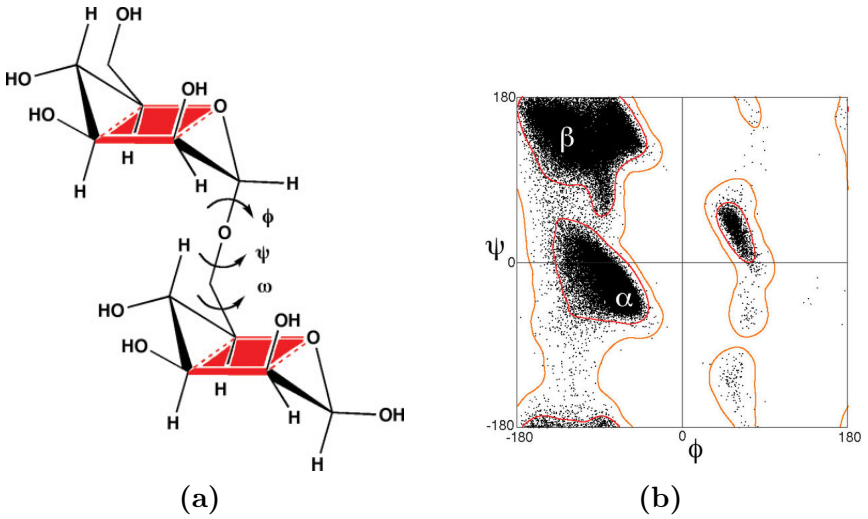


Figure 4.6: ***Ramachandran Plots:*** *(a) Graphical representation of the planes formed by carbohydrates, and the phi (ϕ), psi (ψ), and omega (ω) angles that can be used to provide the relative orientation of the two planes [208]. (b) Ramachandran plot for the general case in peptides, where symbols α and β indicate the two highly populated conformations (data from Lovell et al. [209]).*

For many decades already the method of plotting the ϕ angle versus the ψ angle for amino acid peptide linkages to reveal the coupling between these two angles has been extensively used for protein characterisation. These plots are well understood, with specific regions of allowed and disallowed conformational areas displayed. [210]

Ramachandran plots are mainly used for structure validation of molecules. This is achieved by comparison with experimental data either from X-ray crystallography and/or NMR derived information for the averaged dihedral angles between the linkages. The latter is particularly useful as sampling times for MD are comparable to NMR results for solution conformations and excellent for validation of the particular structure, provided there is experimental data available [211] [212] [213].

Experimentally characterised 3-dimensional (3D) structures of carbohydrates such as di-, poly-, and oligo-saccharides, are predominantly determined by fitting a set of model structures to experimental NMR observables [214] [215]. Analogous to amino acids, the representation of a carbohydrate can be reduced to a plane that is formed by four ring atoms (see Fig. 4.6a). The relative orientation between two adjoined carbohydrates is defined by the angles between the planes, which can be described by the glycosidic angles phi (ϕ) and psi (ψ). An additional angle, omega (ω), is required for $1 \rightarrow 6$ linkages. The coupling between these angles can be observed by plotting ϕ versus ψ versus ω. These carbohydrate Ramachandran-type plots display allowed and disallowed conformational space [208].

4.4.3 Nuclear Overhauser effect distances

In typical NMR experiments, distances are determined from spectra of magnetically susceptible isotopes, typically hydrogen and carbon[13]. Distances between two hydrogens of less than 0.5 nanometers can be identified by correlation with the nuclear Overhauser effect (NOE) intensities. The experimental NOE intensity and the distance between hydrogens are related by calculating both the internal motions within the molecule and its tumbling within the solvent. The slow tumbling rates in large molecules have a nominal contribution to the NOE intensity and therefore, often, only internal motions are used to approximate distances. The internal motions, or fluctuations in distances, are directly related to the intensity by $1/r^6$ where r is the

distance between two protons. NOEs represent an ensemble average of all states present when collecting the NMR spectra. This makes NMR data ideal for comparison to the ensemble of structures identified by a well-sampled molecular dynamics simulation. Distances between hydrogen can be directly calculated and averaged over the MD resulting trajectories. For the purposes of this work, differences between experimental NOE distances and NOE averaged distances from MD of > 0.5 Å were considered significant.

4.5 The Adaptive Biasing Force

The adaptive biasing force (ABF) method is aimed at improving the efficiency of molecular dynamics simulations in which the potential energy surface is sampled ineffectively due to free energy barriers. Such barriers appear as bottlenecks in the dynamics of certain coordinates that describe transitions between physically important states. They also cause the system to become trapped in some states for durations exceeding the time scale of the simulation. In this section, we explore the theoretical underpinnings of this method which was used in MD simulations presented in Chapter 6. ABF has been implemented in NAMD since version 2.6 and has been in the course of time improved and used extensively by practitioners. It is currently offered as one of many collective variable methods.

4.5.1 The reaction coordinate

It is frequently the case that the progress of some chemical, mechanical, or thermodynamics process can be followed by following the evolution of a small subset of generalised coordinates in a system. When generalised coordinates are used in this manner, they are typically referred to as reaction coordinates, transition coordinates, collective variables, or order parameters, often depending on the context and type of system.

In MD simulations, it is often necessary to calculate free energies along an order parameter or on the surface spanned by a few order parameters. Strictly speaking, "order parameter" is a term used in phase transition theory to indicate a thermodynamic variable scaled to zero in one phase and nonzero in another phase.

In the context of the methods described in this chapter, we will use the term "reaction coordinate" to refer to any chosen metric along which the free energy is calculated, whether or not it can be defined as an order parameter. In MD simulations the reaction coordinate is often called a "collective variable".

4.5.2 Thermodynamic Integration for free energy calculations

Different methods exist in the literature that employ various techniques in order to efficiently calculate free energies. One such widely used method is *thermodynamic integration* (TI) [216]. In TI, instead of directly calculating the free energy as a function of the chosen reaction coordinate ξ, its derivative $dA/d\xi$ is computed, and subsequently integrated to yield the free energy, up to a constant.

Most often in free energy calculations, TI is based on unconstrained dynamics, that allow for an optimal and unrestricted exploration of the phase space. The forces acting on ξ are calculated and averaged during the course of simulations over appropriately selected bins and the free energy is reconstructed after the simulation by integration [217]. The term "unconstrained" denotes that no restrictions are imposed on the evolution of the reaction coordinates.

4.5.3 Sampling using the ABF method

The unconstrained TI method does not offer any advantages over classic Boltzmann sampling, unless it is coupled with a enhanced sampling technique along the reaction coordinate ξ, of which there are many (metadynamics [218], umbrella sampling [219], ABF [220], etc). The shared feature of the above techniques is an adaptive adjustment of some quantity that eventually leads to uniform sampling of the chosen

reaction coordinate.

The principle of ABF is to perform unconstrained TI, in which the bias applied is history-dependent, meaning that it is designed to cancel out the running estimate of the local free energy gradient at each binned value of the reaction coordinate. Concretely, the mean force along the reaction coordinate, is estimated from a running average in the appropriate bin. At that point, an external force that opposes the mean force is applied, the net result being, after a brief equilibration, that the average force acting on the reaction coordinate is close to zero and the system undergoes diffusion-like motion along the order parameter, rendering the sampling of the reaction coordinate uniform. The adaptive biasing force is applied along the reaction coordinate as follows:

$$F^{ABF} = \nabla_x \tilde{A} = - < F_\xi >_\xi \nabla_x \xi \tag{4.18}$$

Here, $\tilde{A}$ denotes the current estimate of the free energy and $< F_\xi >_\xi$ the current average of F_ξ, the force acting on the reaction coordinate.

This is a technique designed to "free" the reaction coordinate to move in the entirety of coordinate space, bypassing high energy barriers. It is an alternative to Boltzmann sampling, which does not allow for a proper exploration of the phase space for complex chemical and biological systems, thus yielding statistical averages that depend strongly on initial conditions.

It is important to make the distinction between TI on unconstrained dynamics and ABF. Each is an independent component of an approach to calculating free energies. TI is the method used to calculate free energies, whereas ABF is used to improve the sampling of the reaction coordinate. Together, these two methods form a highly efficient approach to calculating free energies.

4.6 MD modelling of the EGL

In this chapter we covered all methodological aspects of using MD in this work. The implementation specifics will be presented separately in each of the two following chapters. Chapter 5 describes an MD model of the EGL as a tethered glycan in physiological flow, driven by an added force on all water oxygen atoms. In Chapter 6, enhanced sampling is implemented using the ABF method, in order to investigate the interaction of two charged glycans in close proximity, another aspect of the EGL morphology.

Chapter 5

Case study: Tethered Glycan Under Physiological Flow

This study employs atomic-scale molecular simulations with the aim of investigating the conformational and orientation properties of highly flexible oligosaccharide components of the glycocalyx and their suitability as transduction molecules under hydrodynamic loading. Fluid flow was shown to have nearly no effect on the conformation populations explored by the oligosaccharide, in comparison to static (diffusion) conditions. However, the glycan exhibited a significant orientation change, when compared to simple diffusion, aligning itself with the flow direction. It is the tethered end of the glycan, an asparagine amino-acid, which experienced conformational changes as a result of this flow-induced bias. Our results suggest that shear flow through the layer can have an impact on the conformational properties of saccharide-decorated transmembrane proteins, thus acting as a mechanosensor.

5.1 Flow in the EGL

Molecular dynamics (MD) simulations have been extensively implemented in the study of biochemical pathways of cells mainly in the areas of genomics, proteomics [221], and glycomics studies [222]. The glycocalyx consistency occurring on cell

surfaces is cell-type specific, as well as characteristic of the developmental stage of the cell. However, a series of glycan structure types is common to most eukaryotic cell surface environments. Among the group of glycoproteins, the so-called N-glycoproteins carry the N-glycans bound to the sidechain of an asparagine amino acid of the protein via an N-glycosidic linkage to N-acetylglucosamine, which forms the nonreducing end in all N-glycans.

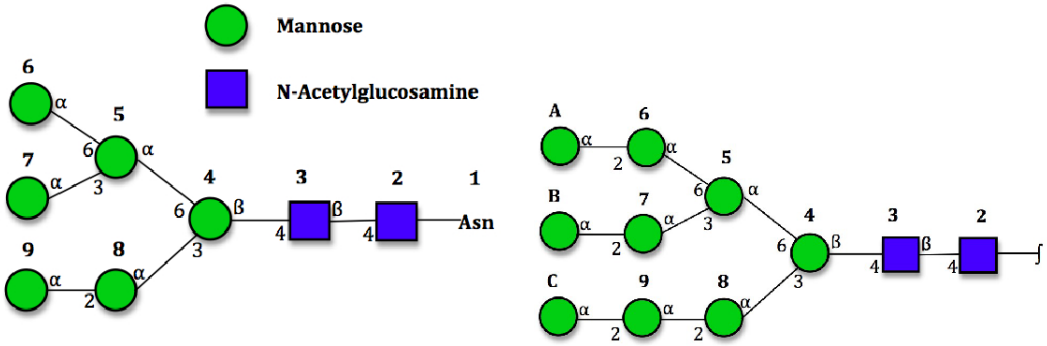
Eriksson et al. [157] developed an EGL model based on branched N-glycan octasaccharides of the high-mannose type with the purpose of investigating alterations in ion hydration and solvation properties with regards to oligosaccharide density in the MD simulations. This was a first attempt at an all-atom EGL model under static conditions ('static' in this sense referring to equivalent static culture conditions but specifically meaning conditions of diffusion). Using their high-mannose octasaccharide as a prototype for the EGL we aim to answer two questions related to the effect of flow: 1) what is the change in the glycan conformation and orientation; and 2) what effect does flow have on the amino acid which tethers the glycan to the glycoprotein anchor? We expand on the Eriksson et al. model of a well-characterised glycan to develop a method with the aim of subsequently applying the lessons to the larger EGL proteoglycans and their associated negatively charged long-chain polymers.

As it will become more clear in the upcoming results section, the velocity of the imposed flow is high, and therefore the simulations are meant to represent large vessels, such as the aorta, femoral, carotid etc.

5.2 System Set-up

The simulation setup described below followed generally that of Eriksson et al. [157]. Due to the significant computational demands the flow implementation adds to the system, we only considered a single glycan configuration, as opposed to an array of

glycans. This limitation is partially alleviated by the lateral periodic boundary conditions imposed.The asparagine-linked N-glycan $Man_6GlcNAc_2$ is shown in Figure 5.1a. Classical mechanical molecular dynamics (MD) simulations were performed on systems consisting of an isolated $Man_6GlcNAc_2$ unit in 0.15 M NaCl aqueous solution.

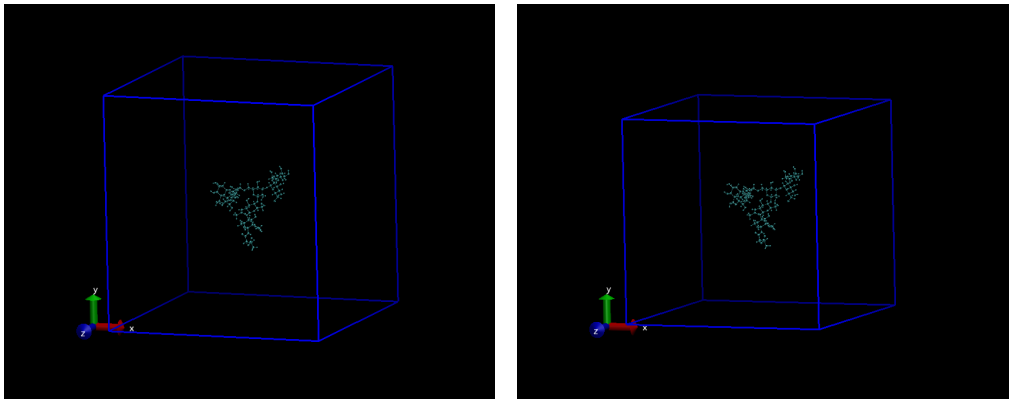


(a) $Man_6GlcNAc_2$ **(b)** $Man_9GlcNAc_2$

Figure 5.1: Schematic representation of glycans *(a)* $Man_6GlcNAc_2$. *(b)* $Man_9GlcNAc_2$ *in NMR study.*

The N-glycan was solvated in a large TIP3P [182] water box with initial dimensions: 62 Å × 66 Å × 61 Å and comprised of 6133 water molecules, 17 Na^+ ions and 17 Cl^- ions (Figure 5.2). This setup models an isolated glycan, separated from its periodic images by approximately 60 Å, vertical to the imaginary surface spanned by the xz plane of the simulation box, with its asparagine end at the base. All MD simulations were performed with NAMD 2.9 [193]. The oligosaccharide sequence was generated according to the all-atom GLYCAM06h [196] parameter set for oligosaccharides and glycoproteins using the web site interface [204].

During the simulation, a constant temperature of 300K was maintained by weak coupling to an external bath, using the Berendsen thermostat [223]. Use of this particular thermostat was essential in combination with flow, because of the particular algorithm used for scaling. Using this method, the temperature of the system is corrected such that the deviation from the assigned temperature exponentially



(a) *Box at initial dimensions* **(b)** *Box dimensions at 100ns*

Figure 5.2: Glycan in water box *(a)* $Man_6GlcNAc_2$ *inside a box with initial dimensions. (b)* $Man_6GlcNAc_2$ *in box dimensions at 100ns. The volume shrinks to accommodate the constant pressure constraint.*

decays with some constant τ:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \tag{5.1}$$

The re-scaling constant used was $0.2ps^{-1}$. Because the re-scaling happens for each coupled atom irrespectively of its velocity, the Berensden thermostat allows for the preservation of low/high as well as directionally biased regions of flow. By contrast, a thermostat that rescales the velocities by adding random noise from a normal distribution would smooth out the effect of flow.

A constant pressure of $1atm$ was maintained with isotropic position scaling and a pressure relaxation time of $100fs$. All bond lengths for hydrogens were constrained to their equilibrium values through application of the SHAKE algorithm [224]. The dynamics simulations were not in any way constrained to reproduce experimental observables.

The cut-off distance, common for electrostatic and van der Waals calculations, was set to 12.0 Å. All pairs connected directly through a 3rd atom were excluded from non-bonded calculations whereas pairs connected directly through 2 intermediate atoms were not scaled, in accordance with the force field documentation (GLY-

CAM06). The non-bonded interactions pair list was generated for distances over 13.5 Å.

The timestep was set to $2fs$. Short-range non-bonded interactions were calculated in every timestep. Full electrostatic calculations were performed every 2 timesteps and the atom list reassignment cycle was 10 timesteps. Particle Mesh Ewald [199] was employed for long range electrostatic interactions.

Parameters were identical for both diffusion and flow simulations. The water flow was implemented as an additional force in the direction of the x-axis applied to the oxygen water atoms. This method is similar to the one used by Koplik et al. [164] [165] and later by Sokhan et al. [166] and Kotsalis et al. [168]. The value of the force for the simulation results presented here was $5 \times 10^{-4} kcal/mol$ Å (approx. $0.35 \times 10^{-2} pN$ per atom) and it was applied in every timestep ($2fs$). In order to simulate the wall shear stress (WSS) effect of the flow, the glycan was tethered to its equilibration coordinates by its asparagine end, using harmonic cartesian restraints (fixed atoms: C, C_α, C_β, C_γ, N) with a restraint barrier of 10 $kcal/molA^2$.

Coordinates and velocities were saved every $0.4ps$, giving 500,000 data frames of both flow and diffusion trajectories to be analysed by the ptraj module of the Amber molecular dynamics package [225]. Two-dimensional torsion angle distribution heat maps where generated using Boltzmann population weighting [226]. Images of molecular structure were generated using VMD [203]. The simulation real time for $200ns$ of diffusion simulations using 64 cores was 11 days. Using identical computational resources, flow simulations were three times more expensive at 33 days for $200ns$.

5.3 Results and Discussion

We have carried out two conformational analyses of $Man_6GlcNAc_2$, using molecular dynamics calculations in diffusion as well as water flow conditions. Restraints

where imposed on the aglycon asparagine residue in order to achieve tethering of the glycan. Figure 5.3 contains snapshots of the simulation during the second half of the evolution of the system (100-200ns in 20ns intervals). We can observe the glycan's range of motion which, due to the 5-atom restraint conditions imposed, can at most align itself perpendicular to the flow, but is never able to align itself to it. However, the proportion of frames where the glycan is in aligned in favour to the flow is increased in comparison to diffusion (where the same restraints are also applied).

As part of the analysis, the molecular dynamics trajectories were used to back-calculate certain NMR parameters, such as inter-proton NOEs. The findings were further compared with results from experimental NMR measurements of $Man_9GlcNAc_2$ [227], shown in Figure 5.1b. Residues A, B and C are absent in the glycan used for the computational model (Figure 5.1a). However it provides experimental results that can be used to approximate the behaviour of $Man_6GlcNAc_2$, as the structures, including the branching pattern, are almost identical.

Investigation of ion hydration properties as well as ion-ion interactions via calculation of radial distribution functions, for the case of diffusion, yielded results identical to those of Eriksson et al. as reported, using the GLYCAM04 [228] and AMBER ff99 [195] force fields. We also looked at the effect of flow in the ion distribution about the solute, resulting from the biasing force of the water on the ions themselves as well as the reorientation of the glycan. Per sugar-residue radial distribution calculations revealed no major differences between static and flow conditions.

5.3.1 Ion Hydration

Here we present ion hydration properties, corresponding to the ones investigated by Eriksson et al. [157] for the same system. They are characteristic of the simple diffusion system and serve as for validation of the simulation.

One of the main differences between Eriksson et al. and this work is the use of

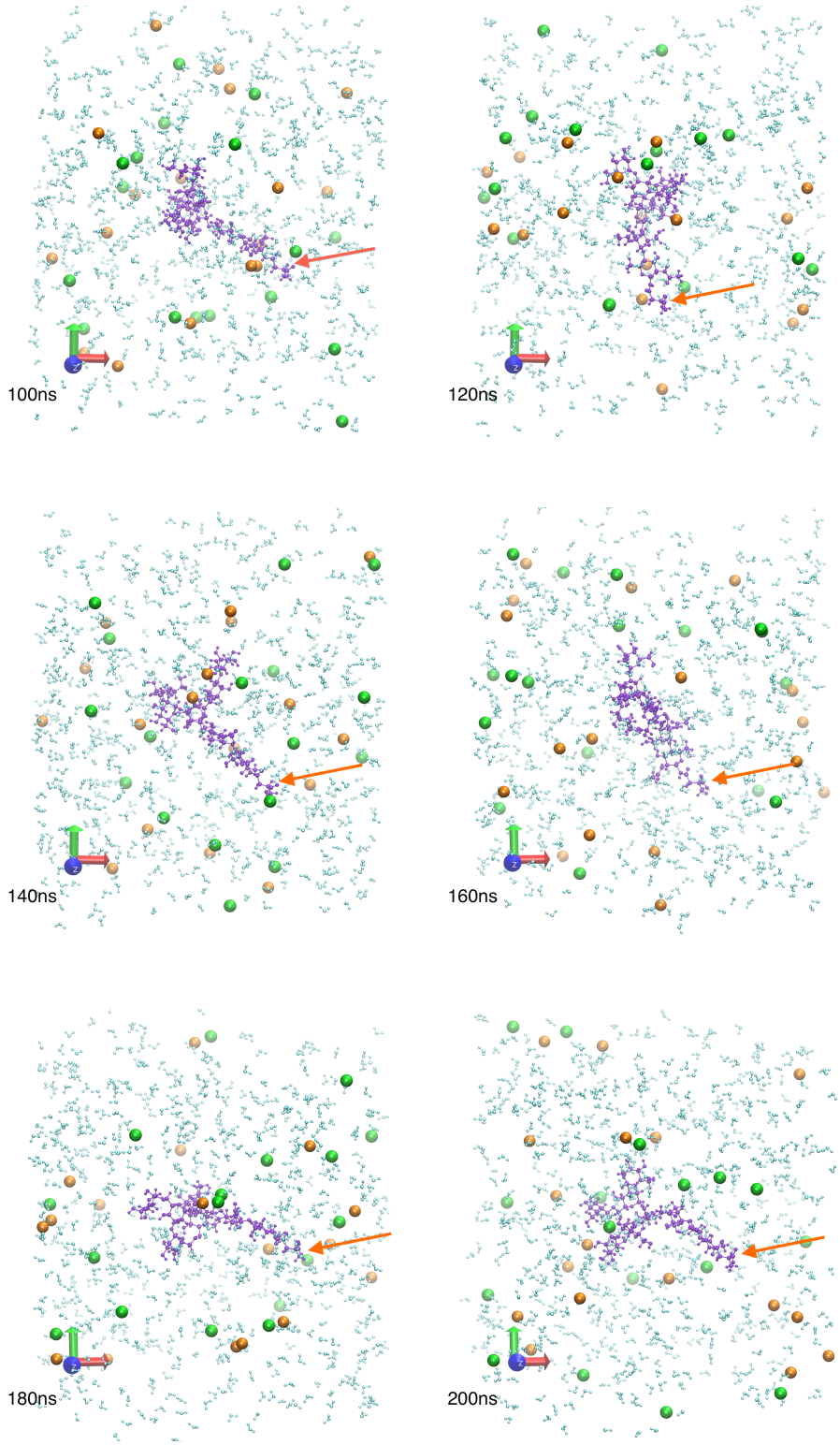


Figure 5.3: Snapshots of the flow trajectory centred around the $Man_6GlcNAc_2$ centre of mass *Showing the glycan (purple), Na^+ (orange) and Cl^- (green) ions as well as 1/20th of water molecules in the simulation box (light blue). The red arrow indicates the tethering location on the aglycon. The axis in the bottom left corner indicates the x (red) y (green) and z (blue) direction. Flow is imposed parallel to the x axis.*

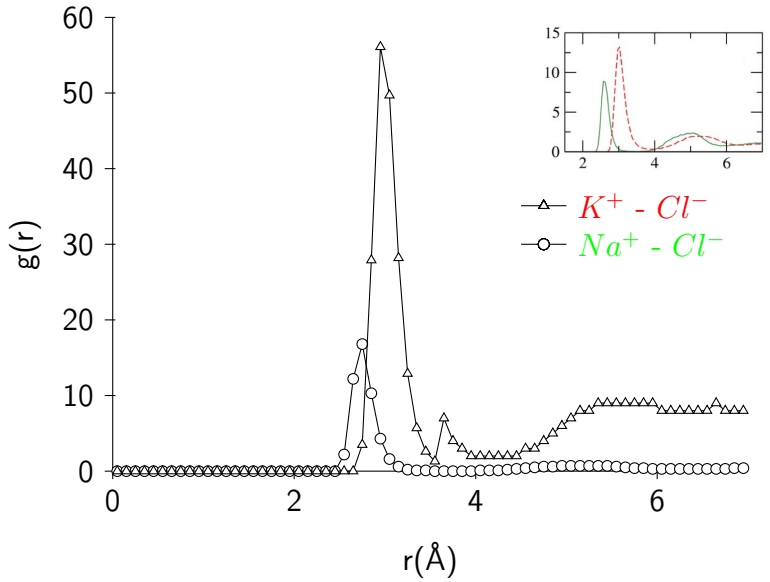


Figure 5.4: ***Ion Interactions in presence of carbohydrate.*** *Cl^- radial distribution functions of Na^+ and K^+. Results from Eriksson et al. [157] are shown in the top right corner (Na^+: green solid line and K^+: red dashed line)*

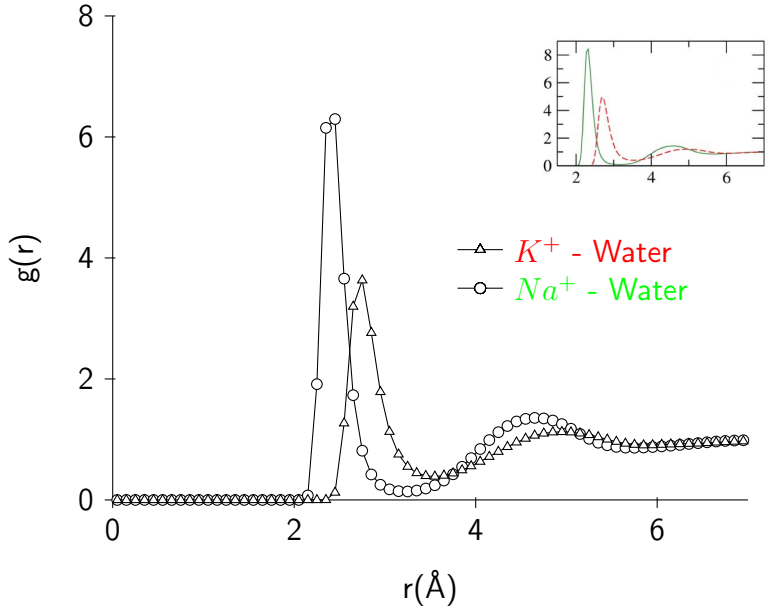


Figure 5.5: ***Water-Ion Interactions in presence of carbohydrate.*** *Water-oxygen radial distribution functions of Na^+ and K^+. Results from Eriksson et al. [157] are shown in the top right corner (Na^+: green solid line and K^+: red dashed line)*

different force field parameters. Eriksson et al. used ion parameters from Beglov and Roux [229] whereas parameters from the AMBER [195] [196] force field were used

in the present work. They also performed reference simulations with ion parameters from the AMBER force field which resulted in a much higher tendency for ion-pairing in general.

In particular, ion-pairing was enhanced approximately fivefold for K^+, as observed in the results of the present work (Figure 5.4). Results indicating higher tendency for ion-pairing are therefore consistent with the Eriksson et al. observations using AMBER parameters and the difference in force field parameters explains the inconsistency between the graph and the embedded image in Figure 5.4.

The relative heights and total areas of cation hydration (Na^+ vs K^+) RDFs are similar to Eriksson et al. In particular K^+ appears to be more flexible in terms of its first hydration shell configuration in comparison to Na^+. However, our system, possibly as a result of difference in the force field ion parameters does not overestimate the first hydration shell of the RDF as much as the Eriksson et al system, compared to the *ab initio* calculated coordination numbers [206].

5.3.2 Calculation of imposed flow

The water-bulk velocity in the direction of the x-axis, as imposed by the added force, was calculated by averaging over instantaneous velocities in the trajectory.

$$v_{x(bulk)} = \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} \frac{v_{xij}}{n}}{m} \tag{5.2}$$

where $n = 6133$, the number of water molecules, $m = 500000$, the number of frames and v_{xi} the instantaneous velocity of water molecule i in the time step at which the frame is saved.

According to Equation 5.2 the water-bulk velocity on the x-axis for all reported results was $v_{x(bulk)} = 0.50 \pm 0.15 m/s$, for the added force discussed.

We believe that the flow rate, although still on the upper limit of what is expected of near-wall velocity even in large vessels [230], is more relevant to the problem

than rates reached in recent studies, where bulk velocities of 10-50 m/s are not uncommon [173] [174] [159]. In larger arteries which can be of several centimetres in diameter, mean blood velocities reach 0.2 m/s and peak velocities can go as high as 1 m/s. As a reference, the mean middle cerebral artery blood velocity has a mean value of 0.52 m/s as reported in studies on athletes [231] at rest. However it is recognised that near-wall velocities are expected to be considerably lower.

The flow imposed is considered continuous rather than pulsatile, given the time-scale of the MD simulation which small compared to typical oscillatory flow frequencies. However the imposed flow velocity is indeed oscillatory, however at a much higher frequency than pulsatile blood velocities.

In order to investigate the relationship between the force imposed per oxygen atom and the resulting flow, simulations of the exact same system were performed for the same number of time steps but with 10% less force imposed. As this is not a straightforward process (many forces operate on each individual atoms, including the effect of the thermostat), the resulting flow was not proportionally low. In fact, the calculated slower flow was 60% of the original: $v_{x(bulk_slow)} = 0.30 \pm 0.14 m/s$.

The magnitude of the force used to drive the flow was calculated using trial and error, allowing the system to run for a few nanoseconds at a time to estimate the average velocity generated. The method used to impose flow is computationally expensive and therefore does not allow for long trial and error simulations, which makes it difficult to accurately estimate the resulting velocities. This is the reason why only two different regimes where compared in full.

The aim of this exercise was to provide guidelines for anyone intending to simulate this system using flow again. The resistance to the flow contributed by the glycan, and therefore its role in regulating the velocity, is very much system-specific. For example, in the absence of a glycan (pure water simulation), the added force causes the water velocity to rise to the point of violating the minimum image convention, thus stopping the simulation.

5.3.3 Effect of the tethered glycan on the water velocity profile

The water velocity profile across the y-axis was calculated by averaging the individual water oxygen velocities v_x of overlapping 7 Å -wide slabs of the simulation box (Figure 5.6). Despite the flexibility of the molecule and heat motion of the water which generates a significant amount of noise, it can be observed that, as expected, along the length of the glycan's range of movement the velocity begins to drop, particularly as the shape of the molecule becomes increasingly dendric, and in accordance with the most populated configurations that we have discussed. The glycan is able to stretch to a maximum of 33 Å after which the velocity begins to rise again and, at the furthest point from the glycan, it reaches its peak. Such a depiction encapsulates several mechanisms influencing the impact of the oligosaccharide moiety on the flow: effect of highly populated conformations, maximum glycan extension, and coupled molecule-water motion are simultaneously captured and result in the velocity distribution shown in the figure.

Although the velocity profile is similar for both flow regimes, it can be observed that the reduced flow regime produces a higher drop within the glycol's range of movement in comparison to the faster flow. This indicates that by further reducing the force which drives the flow, very quickly the velocity on the x-axis could become unsustainable, as the force will not be enough to compensate for the "friction" generated by the non-bonded interactions of the glycan and water and the equilibrating effect of the thermostat.

5.3.4 Flow-induced orientation and conformational changes on the glycan

Figure 5.7 contains a schematic representation of two different kinds of linkages, namely the 1,3 linkage between residues 8 and 4 and the 1,6 linkage between residues

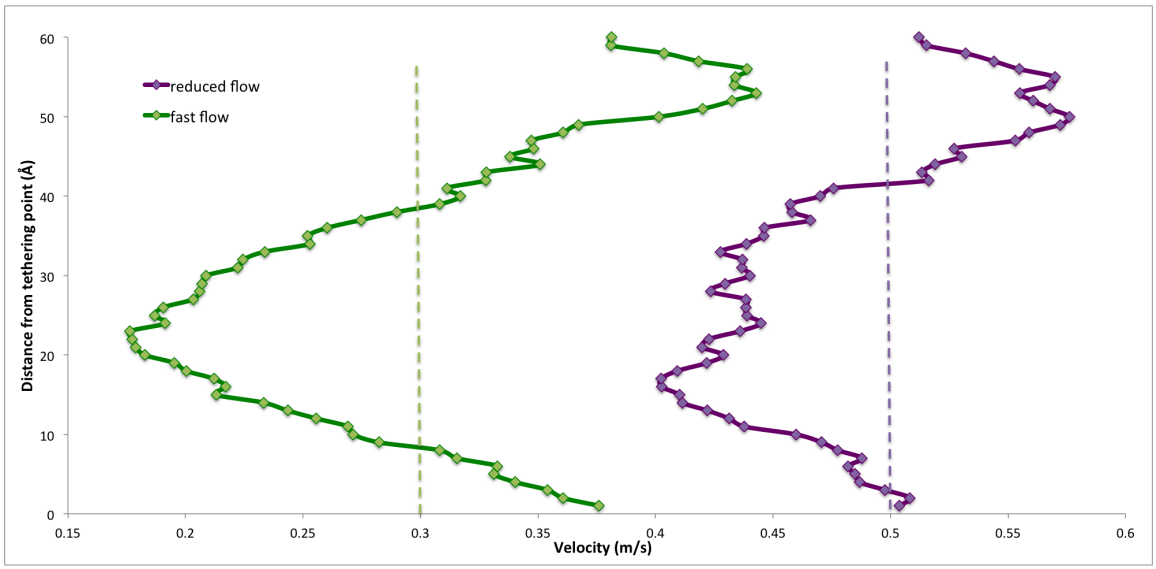


Figure 5.6: Velocity profiles across the y-axis. *The dotted lines mark the average velocity for the normal and reduced flow. Solid lines indicate average velocity for that segment of the y-axis and the embedded schematic of the stretched conformation is to scale. The schematic denotes the highest possible level the constrained glycan can reach. Due to periodicity this velocity profile is repeated above and below the end-points of the y-axis (there is no boundary effect).*

5 and 4. In Figure 5.8 we present plots of torsion angles versus simulation time of $200ns$ for all glycosidic linkages in $Man_6GlcNAc_2$. They are essential in extracting information on the adequacy of sampling of conformational states for the glycan in solution. All plots correspond to the diffusion trajectory torsion angles calculated using the following definition: for (1-x) linkages, where $x = \{1, 2, 3, 4\}$: $\phi = H_1 - C_1 - O - C'_x$ and $\psi = C_1 - O - C'_x - H'_x$, and for the (1-6) linkages: $\phi = H_1 - C_1 - O - C'_6$, $\psi = C_1 - O - C'_6 - C'_5$ and $\omega = O - C'_6 - C'_5 - O'_5$.

Corresponding results for flow do not significantly differ, with a few exceptions that will be discussed later in this section. The 1,2 and 1,3 linkages (Figure 5.8e - 5.8g) principally occupied the $-60°\phi$ angle and $0°\psi$. The 1,4 linkages (Figure 5.8c - 5.8d), remained at $+60°\phi$ and $0°\psi$ throughout the simulations.

The more flexible 1,6-glycosidic linkages (Figure 5.8a - 5.8b) show one state for ϕ, at $-60°$, and two states for ψ, mostly $180°$ with some $60°$. The ω angle in both linkages shows that the $-60/+60°$ states are dominant though 6-5 is a mixture of

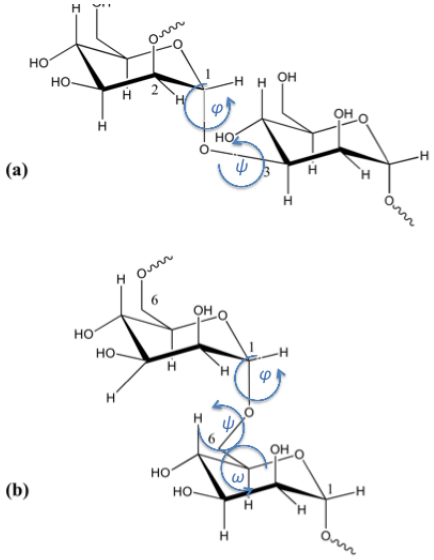


Figure 5.7: Detailed schematic of glycosidic linkages *(a) 1,3 linkage between residues 8 and 4 and (b) 1,6 linkage between residues 5 and 4.*

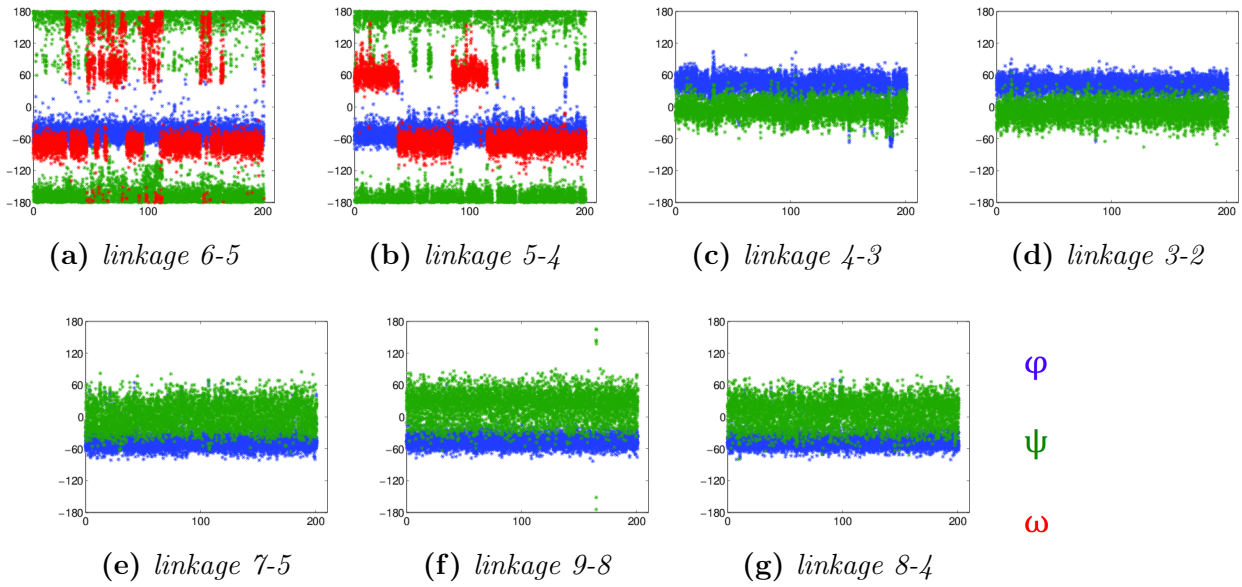


Figure 5.8: Plots of torsion angle versus simulation time for the glycosidic linkages (a) 6-5, (b) 5-4, (c) 4-3, (d) 3-2, (e) 7-5, (f) 9-8 and (g) 8-4 in $\mathbf{Man_6GlcNAc_2}$. *Colours used in all graphs: blue for ϕ, green for ψ, and red for ω. The data were obtained from 200ns of unrestrained MD simulations in water.*

180°, 60° and −60° while 5-4 only has $-60/+60°$. Interestingly, the only difference between diffusion and flow comes from the 5-4 linkage, specifically the ψ torsion angle (Figure 5.9), which occupies a new −60° state, not explored at all in diffusion. This

state is also not observed in the slower version of the flow. One possible explanation is that this new state, exclusive to the faster flow, has a low occurrence frequency, and that the flow provides the necessary energy for more frequent transitions.

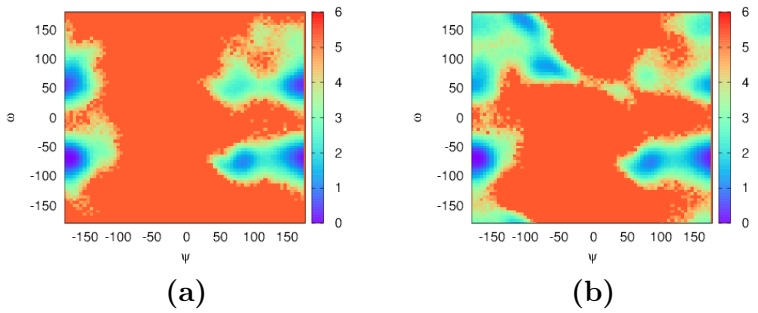
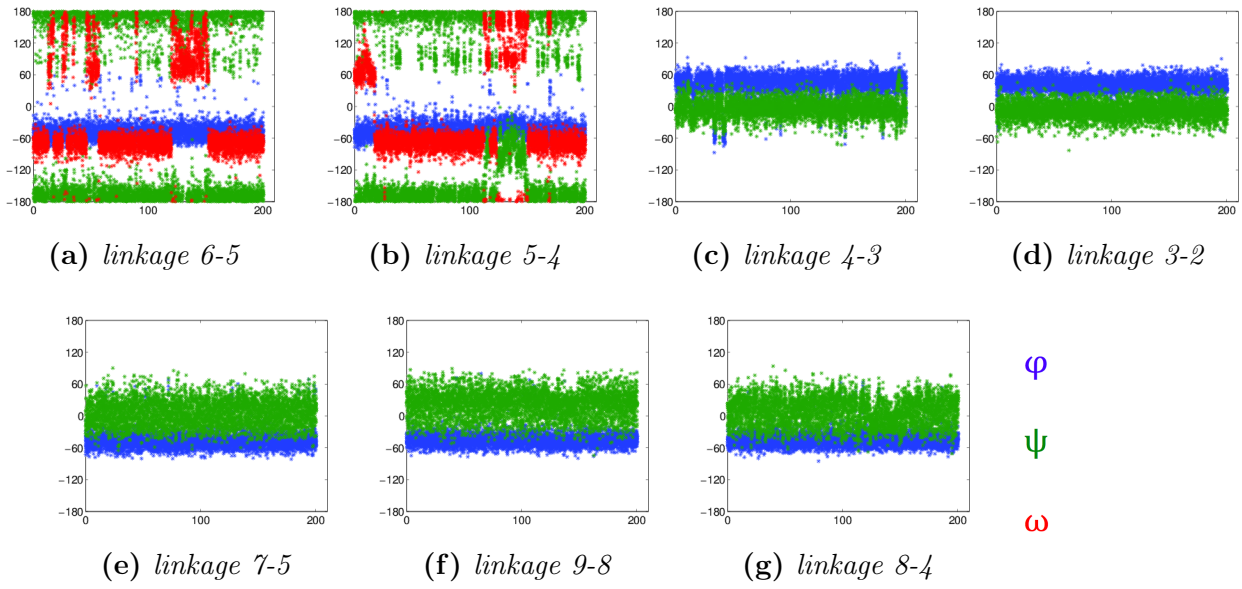


Figure 5.9: Boltzmann energy heat maps for torsion angle ψ vs torsion angle ω in the 5-4 linkage. *(a) In diffusion and (b) in flow. Energy colour bar in kcal.*

All torsion angles in flow over the simulation time are shown in Figure 5.10. Here we can observe (sub-Figure 5.10b) that the -60 ψ angle state for the 5-4 linkage occurs and becomes sustained only once in the trajectory (compared to diffusion, Figure 5.8b). If we take the condition for convergence to be the observation of each state at least twice in the trajectory, then the frequency of this rare state cannot be determined based on these results, and in fact may be, under diffusion, in the order of microseconds.

Table 5.1 contains the calculated and experimentally observed NOEs for the asparagine linked $Man_6GlcNAc_2$ and $Man_9GlcNAc_2$, respectively. It is important to highlight that, although $Man_6GlcNAc_2$ and $Man_9GlcNAc_2$ are similar, the presence of a terminating residue on each of the three branches could alter the branching linkage conformations. Nonetheless, the NOE distance values calculated from the diffusion trajectory are in good agreement with experimentally derived ones. Discrepancies of about 1.0 are observed in the $\alpha1\rightarrow2$ H_1/H_3' distance between residues 9 and 8, the $\alpha1\rightarrow3$ H_1/H_1' and H_1/H_4' distances between residues 8 and 4 and the $\alpha1\rightarrow3$ H_1/H_4' distance between residues 7 and 5. In all cases the NOE distances



(a) *linkage 6-5* **(b)** *linkage 5-4* **(c)** *linkage 4-3* **(d)** *linkage 3-2*

(e) *linkage 7-5* **(f)** *linkage 9-8* **(g)** *linkage 8-4*

Figure 5.10: Plots of torsion angle versus simulation time for the glycosidic linkages in flow (a) 6-5, (b) 5-4, (c) 4-3, (d) 3-2, (e) 7-5, (f) 9-8 and (g) 8-4 in $\mathbf{Man_6GlcNAc_2}$. *Colours used in all graphs: blue for ϕ, green for ψ, and red for ω. The data were obtained from 200ns of unrestrained MD simulations in water.*

appear increased, potentially due to insufficient sampling, that is, the possibility that conformations with short proton-proton distances were either not sampled or under-sampled. We show in the following analysis performed on the 8-4 and 9-8 linkages that the latter is not the case.

In the case of the 8-4 glycosidic linkage, for which analysis of the simulation trajectory results in short NOE distance in comparison to experimental values given in Table 5.1, only one conformation has been sampled (Figure 5.8g). The Boltzmann energy distribution of the ϕ vs ψ angle in diffusion (Figure 5.11a), when examined together with the average H_1 - H'_4 proton distance map (Figure 5.11b), indicates no sampling bias towards long distances. This is also true in the case of flow. The behaviour of the 9-8 glycosidic linkage is similar when examining the average H_1 - H'_3 proton distance map (Figure 5.12b) alongside the Boltzmann energy distribution in diffusion (Figure 5.12a).

The NOE distance values calculated from the flow trajectory did not present

Table 5.1: Calculated inter-proton inter-residue distances in $Man_6GlcNAc_2$ compared to experimental observations in $Man_9GlcNAc_2$ [227]. Distances derived from overlapping peaks are indicated in parentheses.

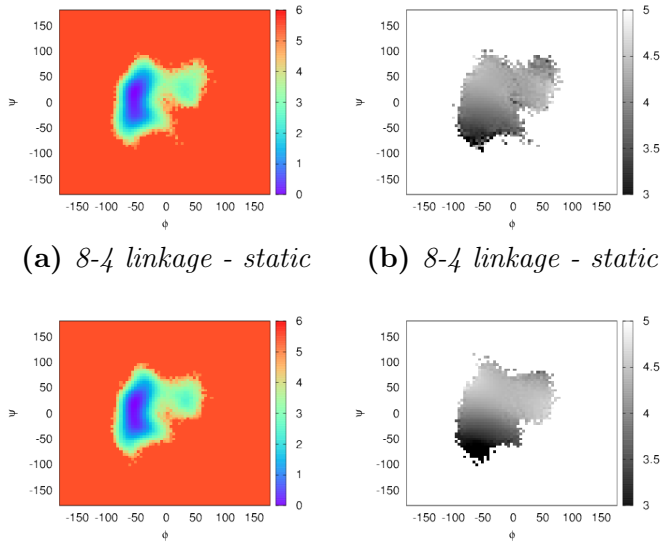
Linkage	Proton pair	Distance NMR (Å)	Distance MD diffusion (Å)	Distance MD flow (Å)
$(\alpha 1 \rightarrow 2)$				
$9-8$	H_1/H_1'	obscured	3.19	3.20
	H_1/H_2'	(2.10)	2.28	2.28
	H_1/H_3'	3.23	4.28	4.28
	H_1/H_4'	no-NOE	4.40	4.41
	H_5/H_1'	(2.74)	2.45	2.44
$(\alpha 1 \rightarrow 3)$				
$8-4$	H_1/H_1'	3.30	4.61	4.62
	H_1/H_2'	3.11	3.38	3.40
	H_1/H_3'	short	2.27	2.27
	H_1/H_4'	2.98	3.95	3.96
	H_5/H_1'	no-NOE	4.88	4.86
	H_5/H_2'	obscured	2.53	2.53
$7-5$	H_1/H_1'	no-NOE	no-NOE	no-NOE
	H_1/H_2'	2.99	3.54	3.57
	H_1/H_3'	2.09	2.30	2.30
	H_1/H_4'	2.90	3.89	3.89
	H_5/H_1'	no-NOE	4.95	4.91
	H_5/H_2'	2.69	2.51	2.50
$(\alpha 1 \rightarrow 6)$				
$5-4$	H_1/H_6'	2.75	2.86	2.68
	H_1/H_6'	2.41	2.42	2.47
	H_1/H_5'	3.27	3.57	3.94
$6-5$	H_1/H_6'	(2.40)	2.75	2.73
	H_1/H_6'	2.14	2.44	2.45
	H_1/H_5'	obscured	4.03	3.85

statistically significant differences in comparison to diffusion, as both sets of values qualitatively describe the same molecular structure. Furthermore, there were no significant differences between this case and the "slow" flow regime, which is why relevant results are not included in the table.

Given the above analysis, and assuming that corresponding NOE data between $Man_6GlcNAc_2$ and $Man_9GlcNAc_2$ should match, we can dismiss the explanation of under-sampling a population of conformations. Also, the assumption that the tumbling of the glycan has no effect on the NOE calculation may also result in

some variation, in agreement with experiment. This is explored in a recent study on sucrose published by Chalmers et al. [232] where they argue that, for long-range NOEs in flexible molecules, long-range distances will be off due to the internal motion of the glycan. The method they propose takes these conformational changes into account as well as the distances, and although it is certainly more correct, it does requires exceptionally long simulations that would have been unfeasible for the purposes of the current work.

However, it is also possible that short-distance conformations were not at all sampled in the MD experiment. Given the timescale of NMR (ms) in comparison to MD (ns) and the fact that we did not constrain the starting conformation based on the NMR experiment, this is, indeed, a plausible explanation.

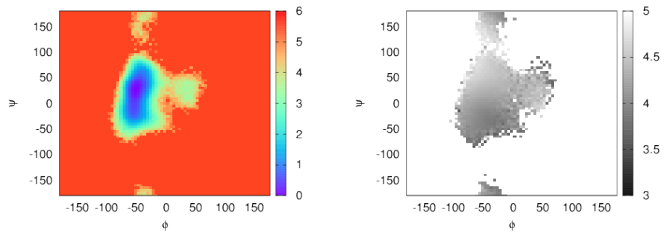


(a) *8-4 linkage - static* **(b)** *8-4 linkage - static*

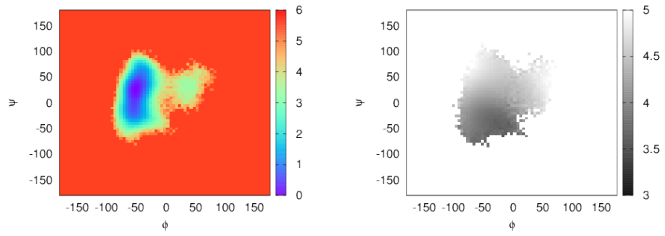
(c) *8-4 linkage - flow* **(d)** *8-4 linkage - flow*

Figure 5.11: Distance and Boltzmann energy heat maps for the 8-4 linkages. *Ramachandran plot of ϕ vs ψ angles, colour-coded according to Boltzmann energy distribution for (a) in static conditions, (c) in flow conditions. Linkage Ramachandran plot of ϕ vs ψ angles, grey-scaled according to distance between protons H_1 and H_4' for (b) static conditions and (d) flow conditions. All colour bars in kcal, greyscale bar in Å .*

For completeness, in Figures 5.13 and 5.14, we present the remaining Ramachandran plots of ϕ vs ψ angles, colour-coded according to Boltzmann energy distribution for the 3-2, 4-3, 5-4, 6-5 and 7-5 linkages in diffusion and flow as well as plots of ψ



(a) *9-8 linkage - static* **(b)** *9-8 linkage - static*



(c) *9-8 linkage - flow* **(d)** *9-8 linkage - flow*

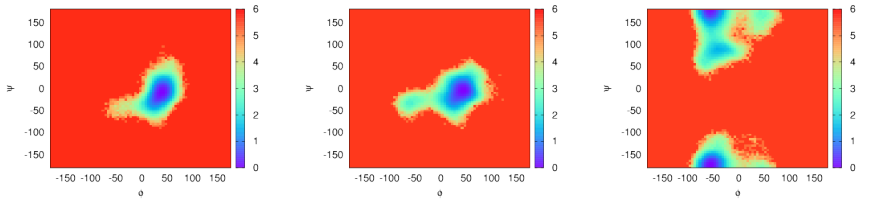
Figure 5.12: Distance and Boltzmann energy heat maps for the 9-8 linkages. *Ramachandran plot of ϕ vs ψ angles, colour-coded according to Boltzmann energy distribution for (a) in static conditions, (c) in flow conditions. Linkage Ramachandran plot of ϕ vs ψ angles, grey-scaled according to distance between protons H_1 and H_4' for (b) static conditions and (d) flow conditions. All colour bars in kcal, greyscale bar in Å .*

vs ω angles for the 6-5 linkage.

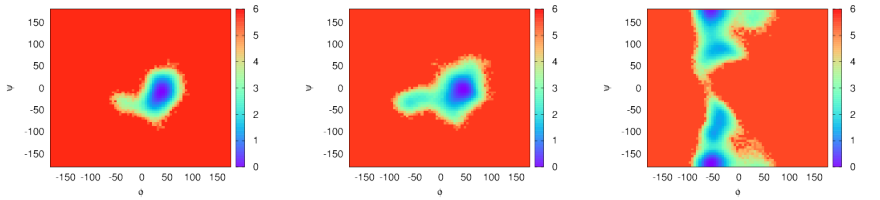
5.3.5 Effect of flow on the amino acid

So far in our analysis, results point towards minor or insignificant changes in the glycan conformation, as characterised by NOE distances and glycosidic linkage torsion angles. There is, however, an alteration on the orientation of the glycan which is attributed to the flow. In this section we investigate whether this information would translate to the transmembrane protein, which is relatively stiff and immobile, therefore unable to align itself with the flow in the same way the flexible glycan does.

Two conformational states were identified which varied between the flow and diffusion simulations. The conformation of these two states were entirely dependent on the orientation of the glycan relative to the amino acid aglycone, defined by the torsion angles according to the following atoms' positions ϕ: C_β, C_γ, N, C_1', ψ: C_γ,

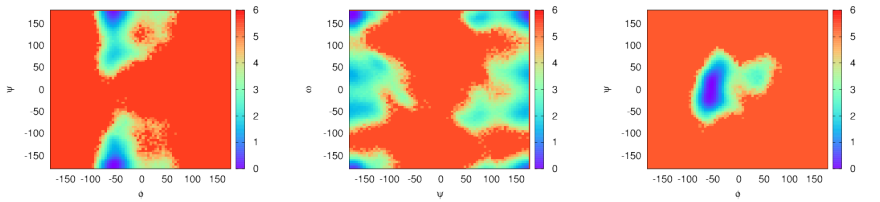


(a) *3-2 linkage - static* **(b)** *4-3 linkage - static* **(c)** *5-4 linkage - static*

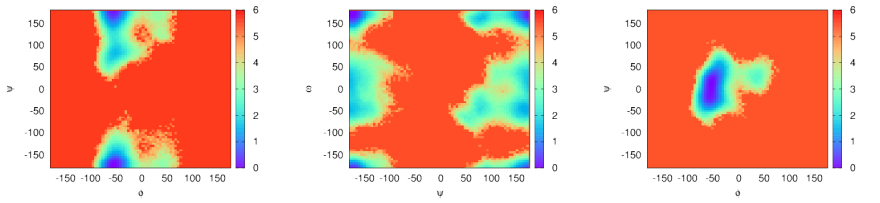


(d) *3-2 linkage - flow* **(e)** *4-3 linkage - flow* **(f)** *5-4 linkage - flow*

Figure 5.13: Boltzmann energy heat maps for the 3-2, 4-3 and 5-4 linkages. *Ramachandran plot of ϕ vs ψ angles, colour-coded according to Boltzmann energy distribution for the 3-2 linkage in (a) static conditions and (d) flow conditions, the 4-3 linkage in (b) static conditions and (e) in flow conditions and the 5-4 linkage in (c) static conditions and (f) flow conditions. Also shown is the Ramachandran plot ψ vs ω for the 6-5 linkage. All colour bars in kcal.*



(a) *6-5 linkage - static* **(b)** *6-5 linkage - static* **(c)** *7-5 linkage - static*



(d) *6-5 linkage - flow* **(e)** *6-5 linkage - flow* **(f)** *7-5 linkage - flow*

Figure 5.14: Boltzmann energy heat maps for the 6-5 and 7-5 linkages. *Ramachandran plot of ϕ vs ψ angles, colour-coded according to Boltzmann energy distribution for the 6-5 linkage in (a) static conditions and (d) flow conditions. Ramachandran plot of ψ vs ω angles, colour-coded according to Boltzmann energy distribution for the 6-5 linkage in (b) static conditions and (e) flow conditions. Ramachandran plot of ϕ vs ψ angles, colour-coded according to Boltzmann energy distribution for the 7-5 linkage in (c) static conditions and (f) flow conditions. All colour bars in kcal.*

N, C_1', C_2' and χ_2: C_α, C_β, C_γ, N for the asparagine amino acid (C_1', C_2' correspond to residue number 2. C_α, C_β and C_γ are all restrained atoms, whereas the remaining are not.). The effect that this linkage has on the orientation of the glycan is shown in Figure 5.15 using the average ϕ, ψ and χ_2 angles for each of these two states.

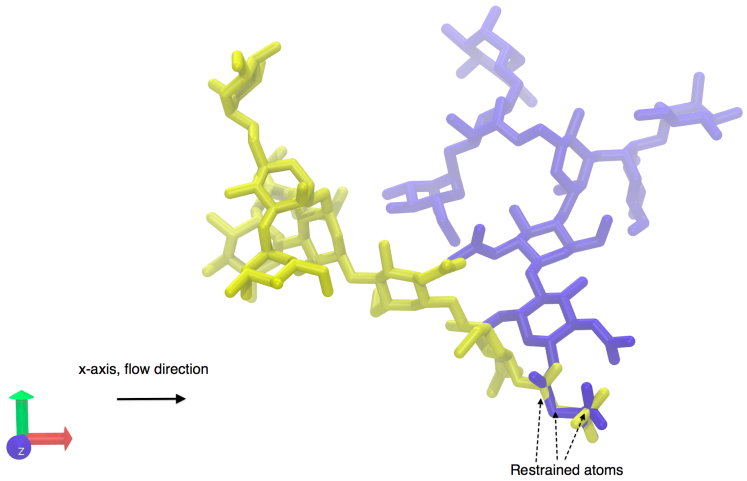
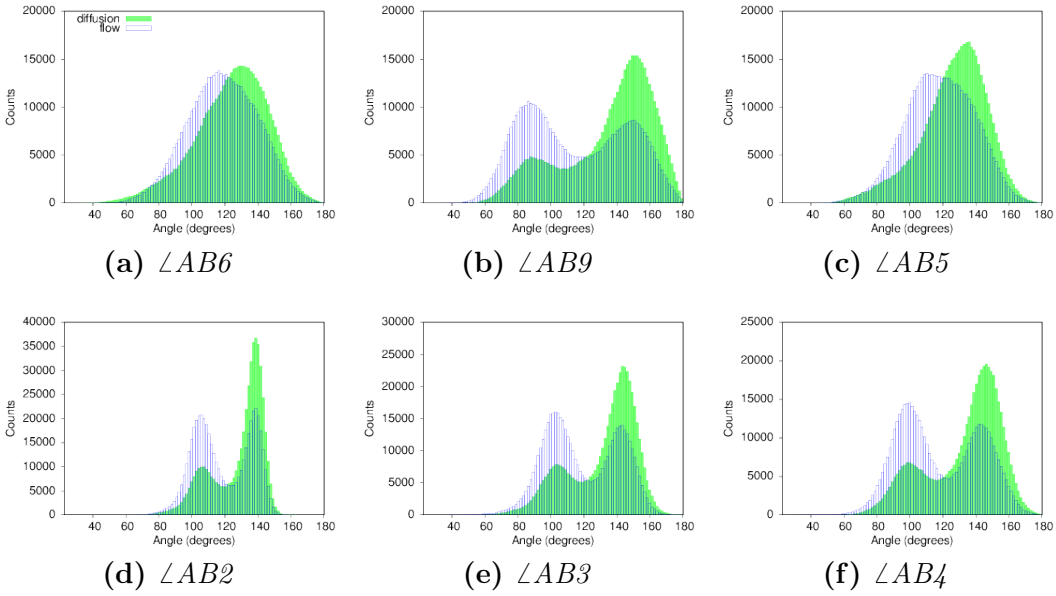


Figure 5.15: A representation of the two dominant states in diffusion and flow. *The yellow and blue molecules represent the two distinct highly populated states favoured by the glycan. While both states are observed in flow and diffusion, the yellow molecule's conformation was dominant in diffusion whereas the blue molecule conformation was dominant in flow conditions.*

In order to quantify the orientational changes flow introduces to our system, we have compared characteristic angles formed between two fixed points on the asparagine end and the individual branches of the glycan in the diffusion and flow trajectories (Figure 5.16). All angles are of the form $\angle ABx$, where A and B are the restrained atoms C_α and C_β respectively and x is the centre of mass of the residue. In both flow and diffusion, the sugar residue appears in two principal states, as evidenced by the peaks on Figures 5.16b, 5.16d, 5.16e and 5.16f.

Flow affects the angle distribution in two ways: By inverting the principal state populations, making smaller angle values the dominant state, as seen for residues 2, 3, 5 and 9, which represent one branch of the glycan (Figures 5.16b, 5.16d, 5.16e



(a) $\angle AB6$ **(b)** $\angle AB9$ **(c)** $\angle AB5$

(d) $\angle AB2$ **(e)** $\angle AB3$ **(f)** $\angle AB4$

Figure 5.16: Calculated angles characterising the orientation and bending of the glycan. All angles are of the form ∠ ABx, where A and B are the restrained atoms C_α and C_β respectively and point x corresponds to the centre of mass of the residues $(a)6$**,** $(b)9$**,** $(c)5$**,** $(d)2$**,** $(e)3$ **and** $(f)4$**.**

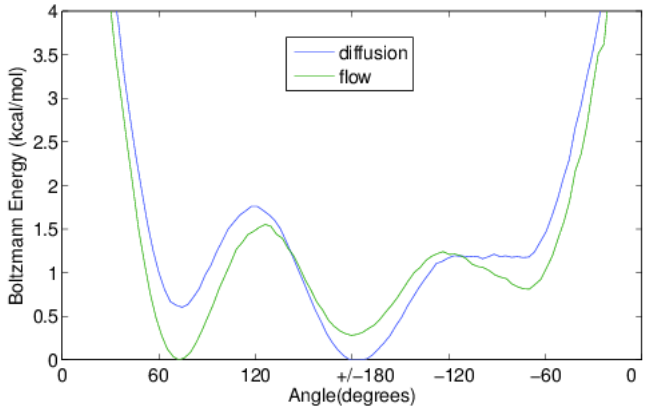


Figure 5.17: Boltzmann energy plots of the χ_2 angle populations for the three aglycon (asparagine) conformational states. *The energy shift between diffusion and flow is pronounced especially for the* $+60$ *and* $+/-180$ *angles.*

and 5.16f) and also by shifting the population towards the lower angle values in the case of residues 5 and 6, on the other branch (Figures 5.16c and 5.16a) and this dominance switches between the two states. This transition favours the orientation of the glycan which is most closely aligned with the flow vector as demonstrated in

Figure 5.15.

To illustrate the effect of flow on the tethered asparagine end of the glycan, we plot Boltzmann energy χ_2 angle populations (Figure 5.17). Of the three distinct highly populated states, the χ_2 value of $\pm$180 degrees is dominant in diffusion, a result consistent with crystallographic experiments [233], whereas in flow, the dominant state shifts to the lower χ_2 values around 60 degrees. The overall shift in energy is in the order of 1 kcal (+0.6 kcal for 60 over 180 degrees in diffusion and -0.3 kcal in flow). There is a third state in the -60 degrees which is not significantly populated in either case. For the case of reduced flow, the increased energy around 60 degrees is still observed, however the relative energy at $\pm$180 degrees remains the same.

5.4 Summary of findings

Using conformational analysis we illustrated that, within the time-frame of the simulations carried out, flow does not significantly alter the conformation space explored by the glycan, as indicated by glycosidic torsion angles and proton distance calculations. In the case of one linkage exploring additional ψ and ω angles, this could be a result of enhanced sampling the flow provides, rather than actual change in conformation. Calculated results derived from the simulation trajectories are accordance with expected experimental values for both diffusion and flow conditions. Indeed, that is to be expected, given the relative magnitude of velocity attributed to flow in relation to Brownian motion.

As a result of tethering, the orientation of the glycan is limited to a specific range. Differences in tethering would produce a smaller or larger range of movement and could allow the glycan to align its orientation with the flow over time. This introduced "resistance" to alignment with flow gives rise to the difference in angle distribution, as measured from the asparagine end to the branches of the glycan.

We have also shown that flow translates to a population shift for torsion angles in the asparagine end, in order to facilitate flow alignment.

In the following chapter, the electrochemical aspect of the EGL components, absent from the simple flow model, is investigated in detail.

Chapter 6

Case study: GAG interactions in close proximity

In this study atomic-scale molecular simulations were employed with the aim of calculating free energy as a function of the density of packed GAG structures, using thermodynamic integration and the ABF enhanced sampling algorithm. The glycan structures investigated were chosen to resemble HS and heparin-like structures. The main aim of the simulations was to calculate PMF curves resulting from the simulation at regular intervals of the chosen reaction coordinate, which is a distance metric used as the surrogate for GAG packing density. This methodology was applied to monomer pairs as well as tetramer pairs with the aim to investigate the potential of using the resulting monomer PMF curves to predict tetramer PMF curves. Further comparisons were made with a naive electrostatic prediction model.

6.1 Charge density in the EGL

As already discussed in Chapter 2, the GAGs which form the structural basis of the EGL exhibit a huge variation in their composition and biochemical complexity. A representative atomic scale model which would contain all the main elements of the EGL in the correct proportions, shapes and sizes is theoretically possible but,

in practice, would be too complex for MD simulations, particularly where explicit solvent is required. Although processing power has come a long way, MD simulations based on such a representation would not be able to run for adequate time to converge in any meaningful way, given current computational capabilities.

In the present work, the focus was turned on certain aspects of the EGL in the atomic scale in order to determine what can be deduced through detailed high resolution models. The hypothesis tested in this chapter in particular is that, in the abstract, the glycan parts of GAGs most significantly influence EGL compressibility via their charge content.

The EGL volume is characterised by inhomogeneity of charge density, owning to the distribution of sulfated and unsulfated GAG regions, and a varying fibre density, depending on the distance of the long GAG strands from the parent proteoglycan, as well as the presence of other biomolecules in the layer and the magnitude of shear acting upon it. Indeed, as discussed in Chapter 3 this depiction of the EGL as a biopolymer brush is not uncommon, as it provides a useful model for studying its behaviour, which, as in brushes of this kind, is mainly characterised by two variables: charge density and the local ionic strength of the medium.

For an EGL layer in a given charge density, a certain force is required to compress the charged GAG strands to closer packing, which will depend on the overall free energy of the system as a function of fibre density. The free energy of compressing the EGL volume therefore depends on the charge density of the layer, itself depending on the packing density (not to be confused with the density of the solvent which itself depends on the volume of the system and number of molecules) of the GAGs and the ionic strength of the medium. In this case, the packing density of the GAGs can be viewed as a "reaction coordinate" against which the system's change in free energy can be calculated. Alternatively the ionic strength can also be used as a reaction coordinate, but it is a more complex choice as it would require the removal and addition of ions to the system.

Although flow is not directly imposed in the simulations of this study, it is implied that the higher density of the EGL strands occurs as a result of the circumferential as well as shear stress applied on the vessel via the blood flow. However, compression of the EGL is also achieved through other factors such as the passing of a white blood cell in small diameter capillaries, the presence of a stent etc. Therefore, investigating the effect of compression of the EGL in this way is of a broader scope and not limited to flow of any particular kind (continuous or pulsatile).

Since the study of the entire, or even significant and representative samples, of the EGL cannot be adequately modelled in detail, this investigation begins with the simplest case of two GAG strands. In this set-up, the free energy of the system is calculated across a distance metric between the GAGs, which acts as a surrogate metric for density. By definition, a one-strand system set-up would represent the case of zero packing density. The distance metric can be defined by various distances. Some examples include: The distance between two atoms, the distance of the centre of mass as defined by groups of atoms, optionally along a given direction (a vector), or a given axis of the cartesian space, etc.

The ABF derivation for any arbitrary reaction coordinate has also been proven to work for the case of multi-coordinate (multi-parameter) ABF calculations. This would allow for example the calculation of a more sophisticated measure of density based on the definition of two distance-related reaction coordinates. Although this approach is demonstrated to work in several examples [220], it is not often used, due to the complications it poses on multi-variable thermodynamic integration. For this reason the approach presented in this chapter was limited to one reaction coordinate (a distance metric) as discussed in the following sections.

6.2 System Set-up

6.2.1 Monosaccharide simulations

Several systems consisting of pairs of monomer and tetramer glycans were investigated. Overall, 6 unique monomer pair configurations were set up. The details for this system are presented in this section. Details for 2 tetramer systems are presented in the following section.

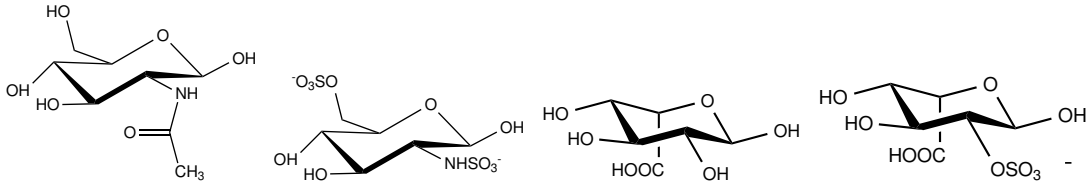
The simplest building block of the saccharide parts in GAGs are monosaccharides. Although, in the literature, dimers (a combination of two monomers) are often considered as the primary units, monomers were chosen for the purpose of this first computational study, as they are simpler units to work with.

As discussed at length in Chapter 2, HS and heparin are closely related and differ mainly in the proportion of various types of constituent elements. For this experiment four of the most frequently occurring elements of heparin were used, as shown in Figure 6.1. However, there is no reason why this process could not have worked with monomers associated predominately with heparan sulfate. Indeed, what this research will investigate is whether the main driver behind inter-molecular interactions in a dense layer of GAGs is a result mainly of sulfating patterns rather than more nuanced elements of glycan structure. For the purpose of consistency, the molecular structures for the remainder of this chapter will be considered and referred to as HS component elements.

In order to measure the effect of sulfation on the free energy of closely packed GAGs, ABF-driven simulations (see Chapter 4 for details) were set up with a distance metric used as order parameter for the reaction, which in this case is simply the change in separation between the two glycans.

The metric chosen was the distance between the centres of mass between the 6 ring atoms for each glycan. Overall, four types of 6-member ring glycans were investigated: N-Acetylglucosamine (GlcNAc), its twice sulfated (N and O) form

(GlcNS(6S)), L-Iduronic acid (IdoA) and its sulfated form (IdoA2S).



(a) *GlcNAc* **(b)** *GlcNS6S* **(c)** *IdoA* **(d)** *IdoA2S*

Figure 6.1: ***Four of the most prevalent heparin constituent monomers:*** *(a) N-Acetylglucosamine (GlcNAc), (b) its twice sulfated (N and O) form (GlcNS(6S)), (c) L-Iduronic acid (IdoA) and (d) its sulfated form (IdoA2S).*

Classical mechanical MD simulations were performed on systems each consisting of two monomer HS component units. Each pair was solvated in a TIP3P [182] cubic water box. Na^+ ions were added as needed to neutralise the resulting aqueous solution. This resulted in each experiment containing a different concentration of ions and ionic strength which in each case, differs from typical ion concentration in the blood serum.

It would be extremely difficult to include all types of ions necessary to recreate a typical volume of blood and a dramatically larger box would be required, owning to the low concentration of certain elements. The decision taken was to simply neutralise the solutions, with the knowledge that the difference in ionic strength can influence comparisons between monomer results.

The configurations and initial dimensions of all combinations studied are outlined in Table 6.1.

Table 6.1: Initial configurations of monomer simulations

Monomer pairs	Na^+ ions	water molecules	Box side (Å)
$GlcNAc - GlcNAc$	0	2762	46
$GlcNS(6S) - GlcNS(6S)$	4	3085	47
$IdoA - IdoA$	2	2542	44
$IdoA2S - IdoA2S$	4	2672	45
$GlcNAc - IdoA$	1	3281	47
$GlcNS(6S) - IdoA2S$	4	3140	47

There are more combinations to be made by mixing sulfated and non-sulfated monomers. However, the focus for this analysis was placed on the combinations outlined in Table 6.1, given that these are the monomers found in the tetramer systems, which as described in the next session. In the case of simulations involving two pairs of the same type and as a result of symmetry in the system and in the way ABF is applied, each simulation is effectively two independent calculations.

As with the flow simulations outlined in Chapter 5, all MD simulations described were performed with NAMD 2.9 [193]. The monomer and tetramer sequences were generated according to the all-atom GLYCAM06h [196] [234] parameter set for oligosaccharides and glycoproteins using the web site interface [204] and the tLEap module [201] of the ptraj in the Ambertools suite [202].

For the purpose of these MD simulations, the Berendsen [223] thermostat that was used in the flow simulations as described in Chapter 5 was not suitable. This is because it allows for "pockets" of very high and low velocity, since atomic velocities are scaled independently of each other. This was a welcome result when the desired outcome was the addition of flow in the system. However, for a system where close packing of molecules is required, this type of thermostat can lead to instability, due to localised high velocity pockets. Instead, Langevin Dynamics were applied to control the temperature of the system. At each time step Δt, the Langevin thermostat changes the equation of motion so that the change in momenta is:

$$\Delta p_i = (\frac{\partial \phi(\mathbf{q})}{\partial \mathbf{q_i}} - \gamma \mathbf{p_i} + \delta p)\Delta t \tag{6.1}$$

where $\gamma \mathbf{p_i}$ damp the momenta and δp is a Gaussian distributed random number with probability:

$$\rho(\delta p) = \frac{1}{\sqrt{2\pi}\sigma} \exp -\frac{|\delta p|^2}{2\sigma^2} \tag{6.2}$$

and standard deviation $\sigma^2 = 2\gamma m_i kT$.

With the exemption of hydrogen, all atoms were coupled to the Langevin thermostat with a damping coefficient γ of 1.

A Nose-Hoover Langevin piston was used for pressure control with the desired pressure set to $1atm$. The Langevin piston Nose-Hoover method in NAMD is a combination of the Nose-Hoover constant pressure method as described by Martyna et al. [235], with piston fluctuation control implemented using Langevin dynamics as in Feller et al. [236]. The barostat oscillation time period was set to $100fs$ and the damping time was $50fs$.

All bond lengths for hydrogens were constrained to their equilibrium values through application of the SHAKE algorithm [224]. This allowed for the simulation timestep to be set to $2fs$, having adjusted for the fast hydrogen bond oscillation. The dynamics simulations were not in any way constrained to reproduce experimental observables.

The cut-off distance, common for electrostatic and van der Waals calculations was set to 12.0 Å. All pairs connected directly through a 3rd atom were excluded from non-bonded calculations whereas pairs connected directly through 2 intermediate atoms were not scaled, in accordance with the force field documentation (GLYCAM06).The non-bonded interactions pair list was generated for distances over 13.5 Å.

Short-range non-bonded interactions were calculated in every timestep. Full electrostatic calculations were performed every 2 timesteps and the atom list reassignment cycle was 10 timesteps. PME [199] was employed for long range electrostatic interactions.

Coordinates were saved every $0.5ps$, giving 100,000 data frames for each monomer pair simulation. The trajectory of the reaction coordinate was also saved every $0.5ps$. Full ABF samples were collected for 1000 timesteps prior to force application. The resolution of ABF reaction coordinate was 0.1 Å and the space investigated was between a minimum distance of 3.5 and maximum of 12.0 Å.

The protocol followed for all monomer simulations was the following: Minimization for $10,000$ steps, heating for $25ps$, equilibration for $2ns$ and ABF production runs for $100ns$.

It is a methodological caveat when applying ABF that the algorithm will not thoroughly sample the reaction coordinate as it will invest more time on high-energy space. A way to overcome this is splitting the reaction coordinate space into windows and applying ABF to them independently, merging the results at the end of all simulations. This method, also referred to as stratification, was applied to one of the six monomer simulations ($GlcNAc - GlcNAc$) in order to test the robustness of the single window simulations.

The range of the reaction coordinate was split into four windows: 3.5 to 5 Å, 5 to 7 Å, 7 to 9 Å and 9 to 12 Å. $100ns$ of production run was then performed for each window, giving a total of $400ns$.

This was a much more expensive way of simulating monomers and in order to evaluate it, the resulting PMF curves were compared to the non-stratified method for the full range of the reaction coordinate (for further details see section 6.4: Results and Discussion). As illustrated in the results section, no significant changes were observed between the stratified ABF and non-stratified ABF application. The single-window method was therefore consistent with a more robust multi-window method while also reducing computational expense by 75%. It was therefore deemed an unnecessary amount of computational time to apply on the remaining monomer simulations.

6.2.2 Tetrasaccharide simulations

The sulfation effects on two longer GAG segments were investigated for the specific case of two identical interacting tetrasaccharides, one with a fully-sulfated pattern and one without any sulfation. A schematic representation of the two types of tetramers (comprised of the monomer components as presented above) is given in

Figure 6.2.

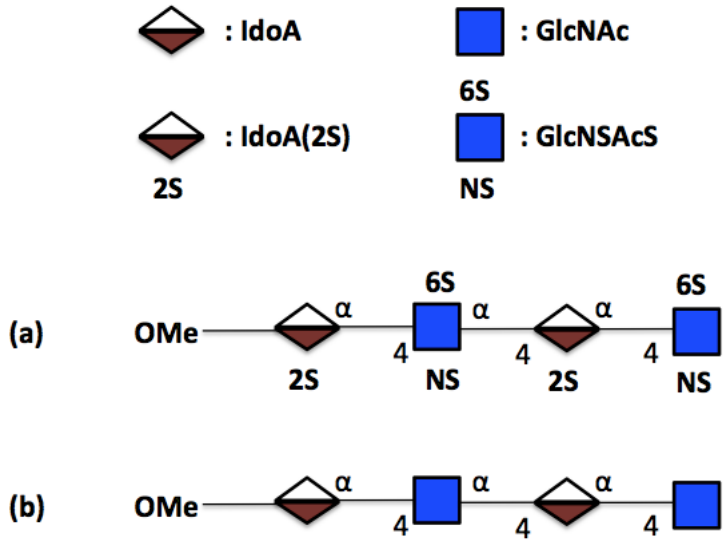


Figure 6.2: Schematic representation of the two tetramers *a) Sulfated b) Unsulfated*

Given the absence of any experimental measurements or indications with regards to GAG density in the literature, we proceeded with constructing a model of close GAG proximity, measured in up to tens of Å rather than nanometers. This model is therefore most applicable to a dense, presumably the deepest, region of the EGL, assuming a non-uniform density which decreases away from the EC membrane, an assumption consistent with EGL imaging evidence. As discussed in Chapter 2, there is some evidence of glycocalyx "fibre" density, as well as quasi-symmetry in experimental observations [93] [94].

Several coarse grained and continuous models have indeed taken a fibre density of 1 every 20 nm and a square lattice of fibres as their basis. However, this density does not explicitly refer to GAG fibres, and is most likely indicative of the proteoglycan spacing. For the purpose of this work, we assume that, for certain regions within the EGL, GAGs are significantly more densely packed than proteoglycans.

The configuration and protocol of the tetramer pairs simulations was mostly identical to that of monomer pairs. The resolution of ABF reaction coordinate was

again 0.1 Å , however the space investigated was between a minimum distance of 4 and maximum of 16.0 Å. Table 6.2 contains the full specification of the starting configurations.

Table 6.2: Initial configurations of tetramer simulations

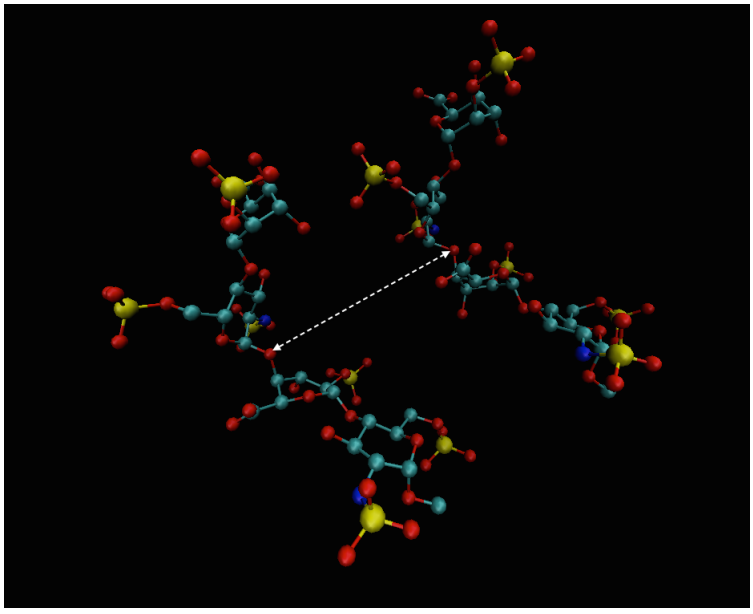
Tetramer pairs	Na^+ ions	water molecules	Box side (Å)
$Sulfated - Sulfated$	16	7804	64
$Unsulfated - Unsulfated$	4	6540	60

One major point of departure in the set-up of the tetramer simulations was the usage of the window method for the ABF calculations. In the case of the tetramers, and in particular the more charged, sulfated one, the force required to bring the pair together was so elevated, that close-range distances took up most of the allocated timesteps for calculation.

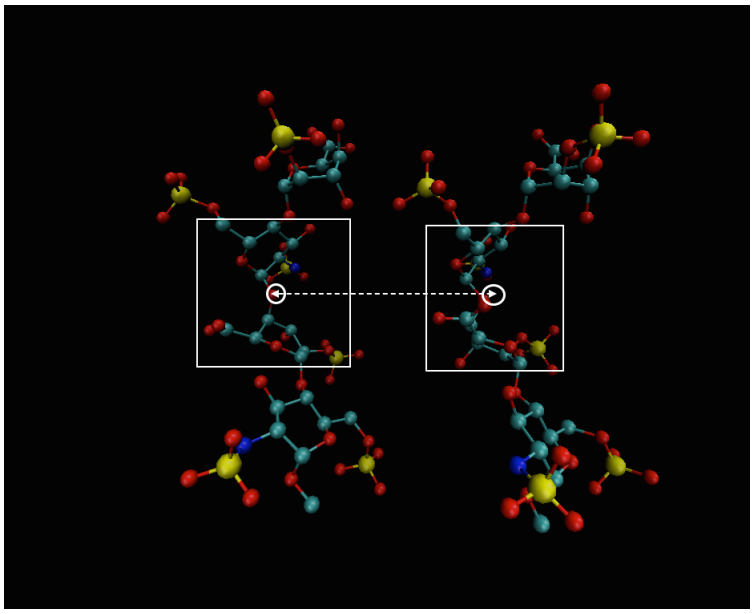
As in the case of the monomer pair described above, the range of the reaction coordinate was split into four windows in order to allow for sufficient sampling of the full range of the reaction coordinate: 4 to 7 Å, 7 to 10 Å, 10 to 13 Å and 13 to 16 Å. $100ns$ of production run was then performed for each window, giving a total of $400ns$.

Another major difference between the monomer and tetramer simulations is that in the case of longer molecules, the choice of a distance metric is not a straightforward process. In fact, it can be argued that a multidimensional application of ABF is required for this process. However, multidimensional ABF leads to complicated integrations and harder to interpret PMF curves. As a middle ground, two distance metrics were chosen and simulations based on them were run independently. Their definitions are shown on Figure 6.3.

The simplest metric used was the actual distance between the two middle glycosidic oxygens (O2O distance), between the $GlcNAc$ and $IdoA$ and between $GlcNS6S$ and $IdoA2S$ for the unsulfated and sulfated versions respectively (Figure 6.3a).



(a) *Glycosidic oxygen distance*



(b) *Centre of mass distance*

Figure 6.3: ***Definition of distance metrics.*** *The sulfated version of the tetramers is used for illustration purposes. Ions and water molecules are not shown for reasons of visual clarity.*

The second metric was defined as the centre of mass (CoM distance) as calculated between the main ring atoms of the two middle residues: $C1, C2, C3, C4, O5, C5$ (Figure 6.3b).

Given all of the above, two simulations were performed on each set of tetramers,

giving four sets of results and a total of 1.6 μs worth of simulation time. Together with $600ns$ for each monomer simulation and a further $400ns$ for the stratified GlcNAc-GlcNAc calculation, the total simulation time amounted to $2.6\mu s$.

6.3 Prediction of tetramer PMF through monomer interaction

6.3.1 PMF model

Calculation of individual PMF responses for each monomer was made in the aim of building a predictive model that can be extended to systems of arbitrary size. The main assumption made is that the principal driving force that determines the interaction of (in the simple case) two HS segments is a result of pair-wise interactions between the monomers. In order to test this hypothesis, the reproduction of the PMF curves of both the sulfated and unsulfated tetrasaccharide was attempted using trajectories generated from both distance metrics. The prediction was based solely on pairwise distances between monomer residues, as recorded during the simulation trajectory.

The first step of implementing this model is fitting functions onto the monomer PMF curves. The method used is a simple linear interpolation and linear extrapolation for outside-of-range values, since the tetramers were studied for larger distances than the monomer pairs. Extrapolation for values of distance larger than 12 Å were kept constant at $PMF_{monomer}$ at 12 Å.

Once the PMF curves are fitted, they can be used to calculate the total PMF of the system based on pairwise interactions. The tetramer trajectory is split into 100 bins, based on the distance coordinate, as defined for each case (O2O, CoM). For each bin, the average PMF is calculated using a simple additive linear combination of all monomer PMFs:

$$PMF_{tetramer,calc}(i) = \frac{\sum_{j=1}^{n}\sum_{k=1}^{m} PMF_{monomer}(j,k)}{n} \tag{6.3}$$

where m equals the number of monomer pair combinations, in this case $4^2 = 16$ and corresponds to the appropriate fitted curve and n equals the number of observations per bin and corresponds to a monomer pair distance grouped within that bin. As with the original PMFs derived from the simulations, absolute values are not significant and it is only the shape of the curves and point differences that are of note. In order to compare with the tetramer simulation derived PMF, the minimum value of the calculated curves was set to 0.

6.3.2 Electrostatic model

A naive electrostatic model was also used to calculate tetramer interaction dynamics based on monomer distances. The methodology resembles the one described in the previous section, with the exception that instead of fitted monomer curves, the relationship used is the Coulomb potential as expressed between each monomer (see Equation 6.4). An assumption was made that each monomer's centre of mass carries the overall residue charge.

$$PMF_{coulomb} = \frac{1}{4\pi\epsilon_w}\frac{q_1 q_2}{r} \tag{6.4}$$

where ϵ_w is the electrical permittivity of water, defined as $80e_0$ and q_1, q_2 the charge born by each residue in the calculation (in the case of the sulfated molecule, the charge is always $2q_e$, q_e being the elementary charge, due to the di-sulfation of the glucosamine and the original charge of IdoA).

The total PMF is again calculated for each bin through a similar methodology as the one described by Equation 6.3:

$$PMF_{tetramer,elec}(i) = \frac{\sum_{j=1}^{n}\sum_{k=1}^{m} PMF_{coulomb}(j,k)}{n} \tag{6.5}$$

where $PMF_{coulomb}$ is given by Equation 6.4. In the case of the neutral tetrasaccharide, given that only the IdoA residue has a charge, this calculation did not take place.

6.4 Results and Discussion

As previously mentioned, simulations involving a pair of the same type are effectively two independent calculations. in the following sections, where appropriate (for example in the case of heat-map or scatter plots), results for each identical monomer will be presented as combined. This will also be the case for corresponding monomer residues in the tetramer calculations, given that the two interacting tetrasaccharides are identical.

6.4.1 Conformation analysis

This section contains results on conformation analysis of monomer shapes both in the monomer simulations and also as constituent parts of tetramers in tetramer simulations. The internal motion of the tetramers is further investigated via conformations adopted by their glycosidic linkage angles.

Monomer ring pucker

Cremer-Pople ring pucker parameters θ, ϕ and Q [124] were used to quantify the pyranose ring geometry for both glycans in each of the monomer calculations. The angle θ was defined based on the ring atoms $C1 - C2 - C3 - C4 - O5 - C5$. All parameters for monomer combinations of the same type are shown in Figures 6.4 - 6.7. There were no observable differences in the case of mixed monosaccharide pairs.

Calculation of pucker parameters was deemed important for two reasons. Firstly, in order to determine whether the application of biasing force in the ABF methodology results in a different pucker behaviour to the one expected. This would suggest

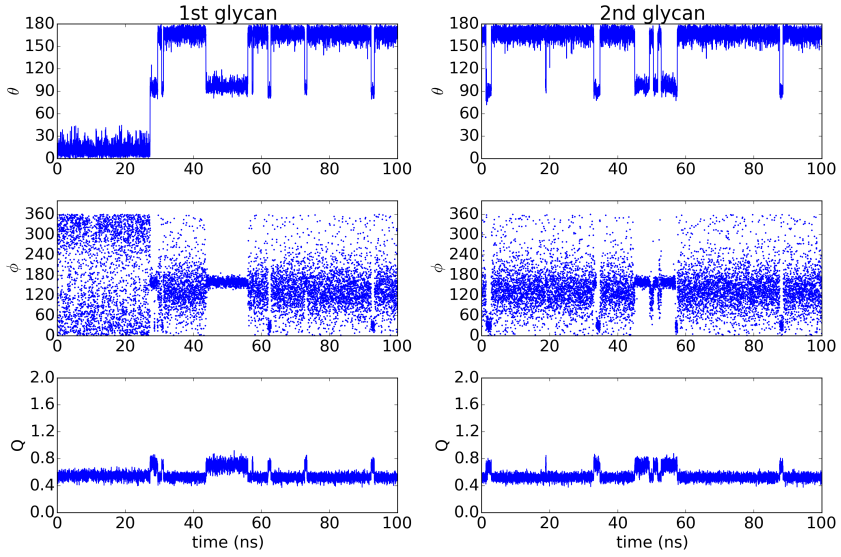


Figure 6.4: ***Cremer and Pople parameters for the IdoA2S - IdoA2S monomer pair.*** *From the top - down: θ, ϕ and Q.*

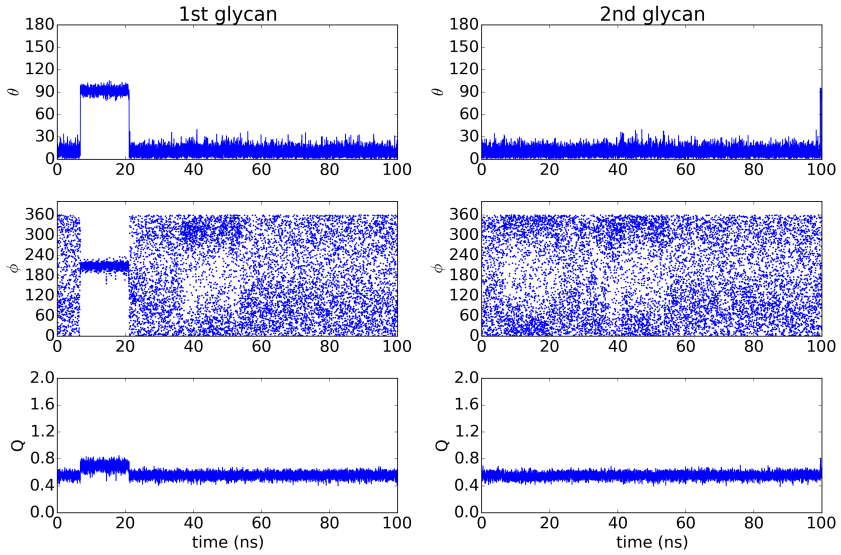


Figure 6.5: ***Cremer and Pople parameters for the GlcNS6S - GlcNS6S monomer pair.*** *From the top - down: θ, ϕ and Q.*

the application of excessive force and therefore abnormally close proximity of the monomers. No out-of-bounds values were observed in any of the simulations.

The second reason for calculating pucker parameters was in order to observe whether the proximity of the monomers would affect the rate of ring flexing. This was not as straightforward to answer with simulations limited to $100ns$, given that

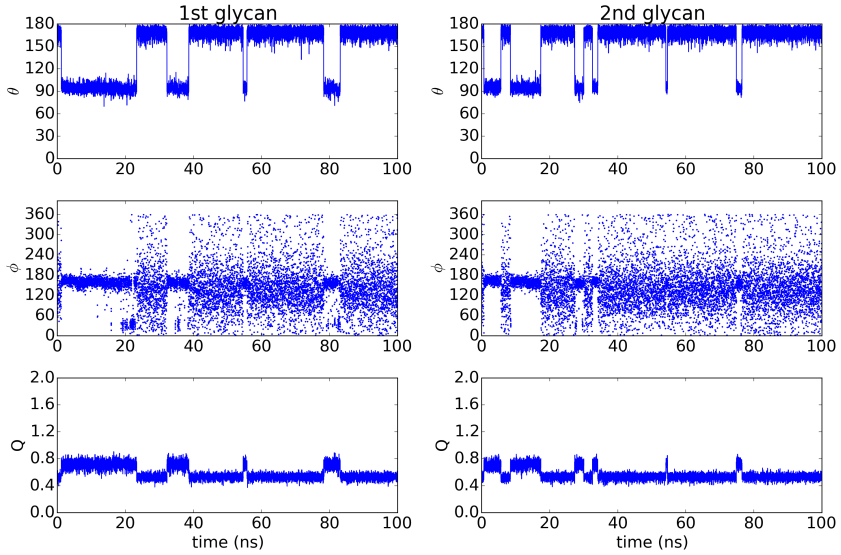


Figure 6.6: ***Cremer and Pople parameters for the IdoA - IdoA monomer pair.*** *From the top - down: θ, ϕ and Q.*

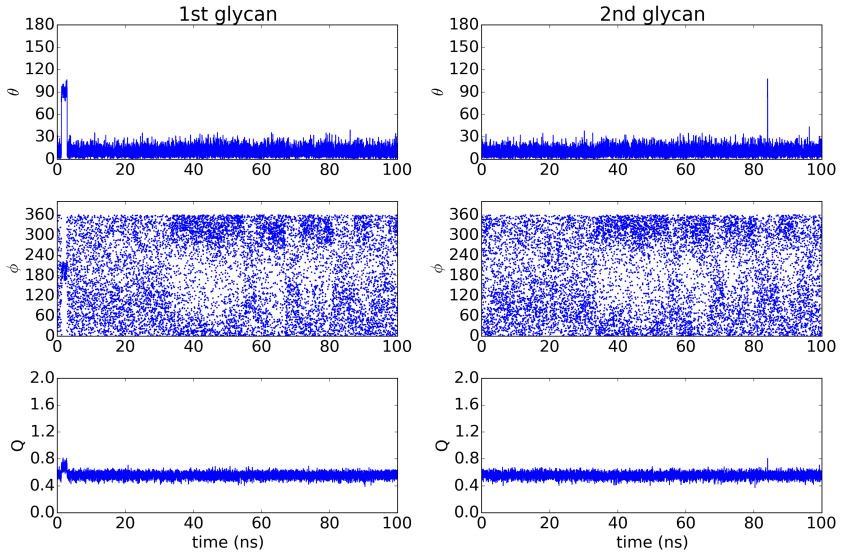


Figure 6.7: ***Cremer and Pople parameters for the GlcNAc - GlcNAc monomer pair.*** *From the top - down: θ, ϕ and Q.*

it is expected that the IdoA2S monosaccharide would have a forward ($^1C_4 \rightarrow^4 C_1$) transition rate of 2 times per μs and backward ($^4C_1 \rightarrow^1 C_4$) transition rate of 176 times per μs [20]. The overall occupancy is predicted at 92% $^1C_4$, 1% $^4C_1$, 3% ^2S_O, 1% $^3S_1$ (skew-boat) and 1% B_{3O} (boat conformation). Both GlcNAc and GlcNS6S monosaccharides have been shown to exhibit a low rate of chair transition

($^4C_1 \rightarrow^1 C_4$) and this was reflected in the simulations performed in this work, as no transitions occur in any of the $100ns$ trajectories.

Overall, flipping between chair conformations is infrequent, as it is not observed on either any of the monomers investigated (as seen in Figures 6.5, 6.6, 6.7) with the exception of one transition in the case of IdoA2S (Figure 6.4).

Another comprehensive way to present Cremer and Pople parameters is by a scatter plot of the ϕ and θ angles. Scatter plots are ideal for observation of the allowed conformation space. For such representations, the sinusoidal projection is frequently used to represent the spherical nature of the parameters. For reference, pucker parameter plots by Sattelle et al. (Supplementary material) [20] for IdoA2S are presented in Figure 6.8, as calculated from a fully converged 5 μs simulation. The annotations on Figure 6.8 indicate the most common conformations for the IdoA2S pucker: Two chair conformations, $^1C_4$ and $^4C_1$ as well as the two skew boat conformations, $^1S_3$ and ^2S_O.

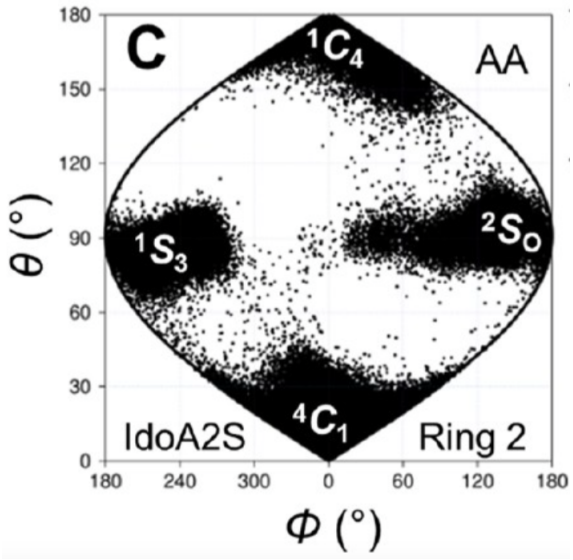
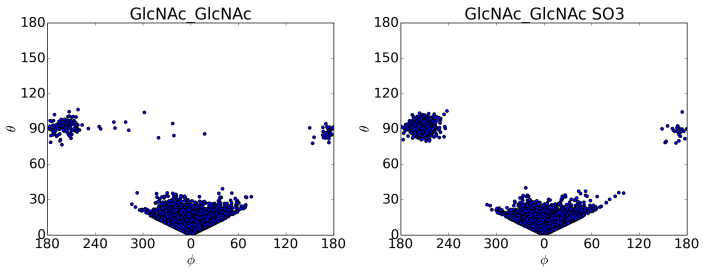


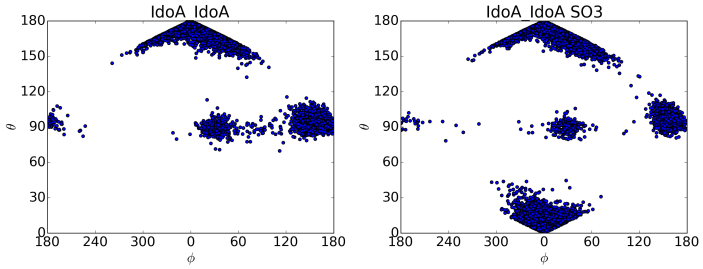
Figure 6.8: ***Sinusoidal projection of puckering space for a fully converged IdoA2S simulation (5 μs) [20]*** *Puckering conformations are annotated and used for reference in the following of the plots of this kind in this chapter.*

Results from monomer simulations in the present work are shown in Figures 6.9a and 6.9b. In order to create the scatter plots, the parameters calculated from both

monomers from each simulation were used, as they are identical.



(a) *GlcNAc monomers, unsulfated (left) and sulfated (right)*



(b) *IdoA monomers, unsulfated (left) and sulfated (right)*

Figure 6.9: ***Sinusoidal projection of puckering space for monomers:*** *Cremer and Pople parameters ϕ and θ calculated for the sulfated and unsulfated (a) GlcNAc and (b) IdoA pair simulations*

Ring pucker conformations are more clearly observed in the scatter plots, with the θ angle defining the $^1C_4$ chair conformation at 150-180 degrees and the $^4C_1$ conformation at 0-30 degrees. Two shorter-lived states were observed for both IdoA and IdoA2S, as shown in Figure 6.9b. The $^1S_3$ state at 90 degrees θ and 150-180 degrees ϕ and a minor one at 90 degrees θ and 0-60 degrees ϕ.

It is typical to also observe an ^2S_O state at 90 degrees θ and 120-180 degrees ϕ, however due to the short length of the simulation and microsecond scale transition-rates reported in the literature for pucker changes, (see Chapter 3) the state does not appear in any of the results. This non-thorough conformation sampling was allowed, as pucker sampling was not considered as fundamental convergence condition for this particular computation. No less than $3\mu s$ of simulation time would have been necessary to achieve this convergence [20].

Tetramer ring pucker

In order to investigate the potential effect of inter-tetramer distance, pucker parameters were calculated separately for all residues and for each of the ABF simulation windows. Ring-flexing behaviour is expected to differ depending on the position of the residue within the glycan, and as such, it is also expected to differ from the single monomer results. This works vice-versa, as the pucker geometry also affects the overall shape of the polymer.

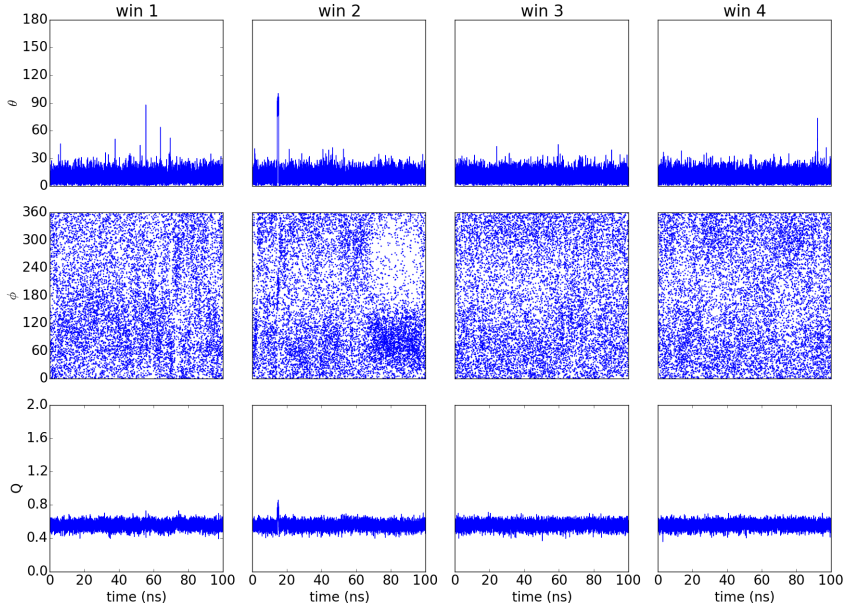
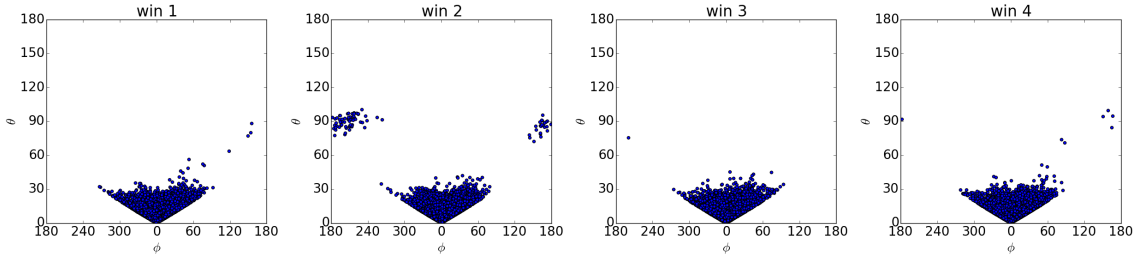


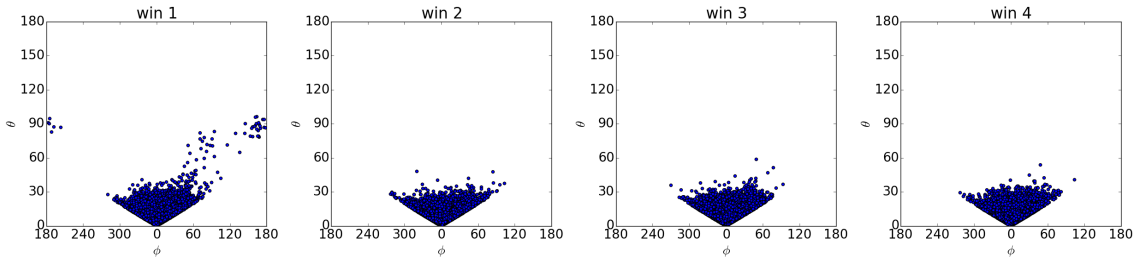
Figure 6.10: ***Cremer and Pople parameters for the terminal GlcNS6S residue of the sulfated tetramer:*** *O2O distance definition.*

In the case of both the terminal and middle GlcNS6S residues present in the sulfated tetramers, pucker parameter averages were constant at $0 - 30°\theta$, which corresponds to the $^4C_1$ chair conformation. No difference was observed in ring pucker parameters between the simulation windows. There was also no observable difference between results obtained from the two distance definitions. This result is expected, given that the calculated chair-chair transition rate is close to 1 μs^-1. For illustration, results for the terminal GlcNS6S residue are shown for one of two residues (the results were identical for both glycans) in the case of O2O distance

definition in Figure 6.10 and as a ϕ vs θ scatter plot in Figure 6.11. Calculated parameters for the non-terminal GlcNS6S showed a similar pattern.



(a) *O2O distance*



(b) *CoM distance*

Figure 6.11: ***Cremer and Pople parameters for the GlcNS6S (non-terminal) of the sulfated tetramer:*** *(a) O2O distance definition and (b) CoM distance definition.*

In the case of the terminal GlcNAc residues present in the unsulfated tetramers, pucker parameter averages were of similar values to those of GlcNS6S. No difference was observed in ring pucker parameters between the simulation windows, with the exception of the first window of only one of the tetramers (O2O distance definition) in which one transition was observed to the $^1C_4$ chair conformation (Figure 6.12). There is otherwise no observable difference as a result of distance definition for the ABF. Parameters for the non-terminal GlcNAc also showed no transition in either window.

In general, for all GlcNAc and GlcNS6S calculated parameters, there was little or no fluctuation during the simulation, mostly observed in the first window of ABF, the one where the tetramers are at their closest point. The lack of chair-chair transitions is expected, as hexosamine flexing is considered a micro-second scale phenomenon which has been shown to increase with polymerisation. Thus, it is important to note

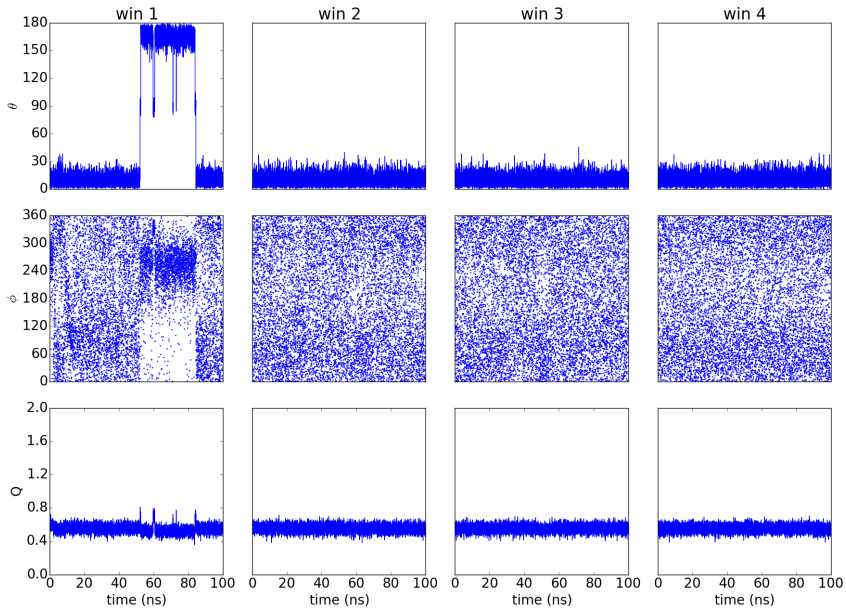
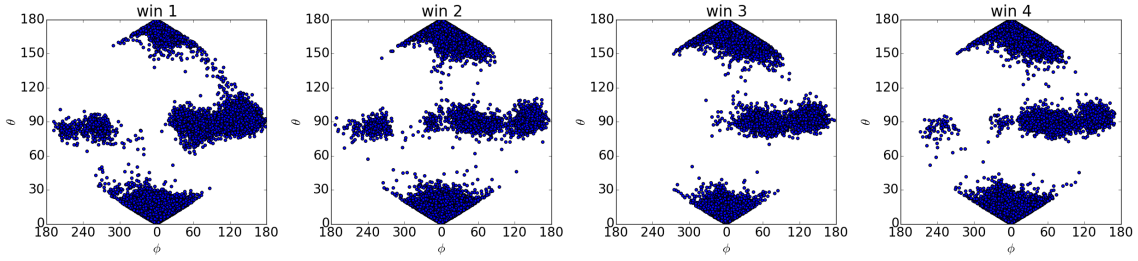


Figure 6.12: ***Cremer and Pople parameters for the terminal GlcNAc residue of the unsulfated tetramer:*** *O2O distance definition.*

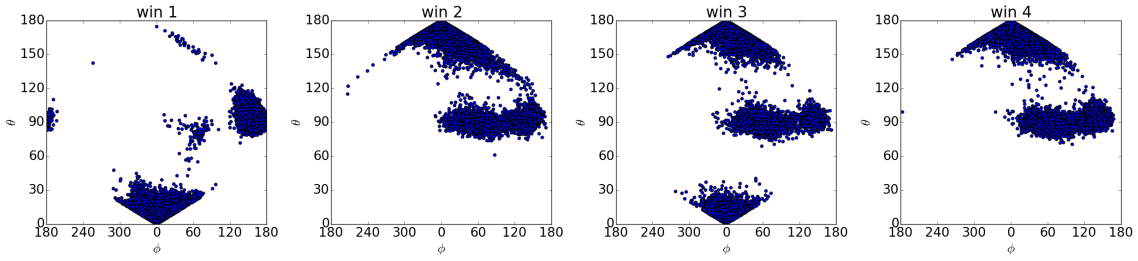
that no convergence, with respect to pucker parameters, has taken place. The one and only transition observed for the first window, indicative of a potentially higher transition rate, in just one trajectory, is most likely linked to polymerisation and unrelated to the biasing method. Overall, non-terminal residues did not undergo any chair-chair transitions, in accordance with all evidence from the literature [20].

The internal (non-terminal) IdoA2S residue is expected to exhibit much more flexibility in comparison to both the free monosaccharide and the terminal positioned residues. This was confirmed in the present work with transitions occurring between chair-chair and also other conformations in some windows and for both distance definitions. For example in windows 3 and 4 in the case of the O2O definition simulation (Figure 6.13a) and one of the residues in the CoM definition (Figure 6.13b).

In the case of the reducing end IdoA2S residues present in the sulfated tetramers, pucker parameters averages showed less variation between windows and a more coherent profile of state populations, for both distance definitions and between the

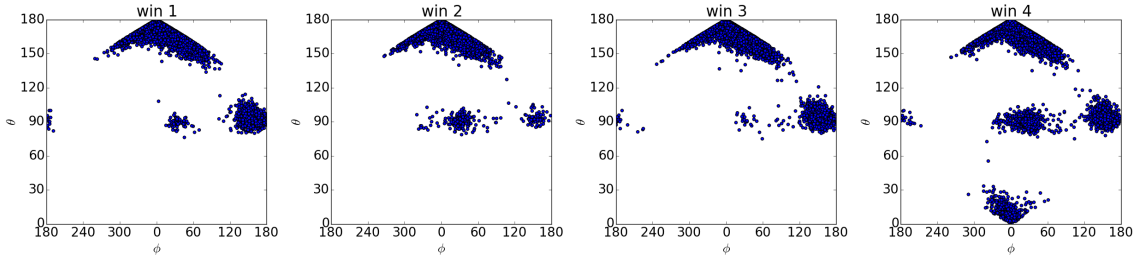


(a) *O2O distance*

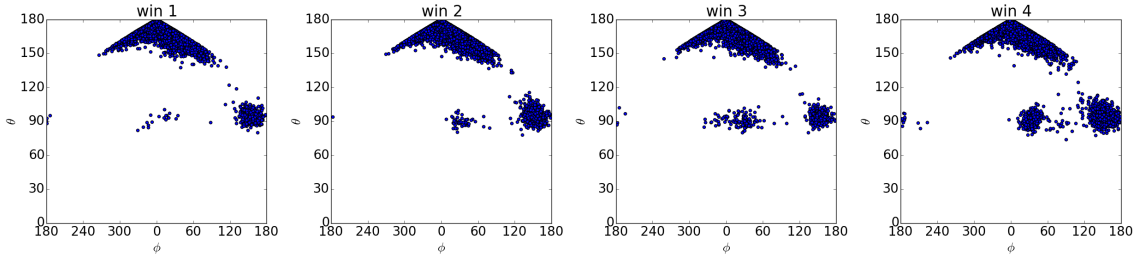


(b) *CoM distance*

Figure 6.13: ***Cremer and Pople parameters for the IdoA2S (non-terminal) of the sulfated tetramer:*** *(a) O2O distance definition and (b) CoM distance definition.*



(a) *O2O distance*



(b) *CoM distance*

Figure 6.14: ***Cremer and Pople parameters for the IdoA2S (reducing end) of the sulfated tetramer:*** *(a) O2O distance definition and (b) CoM distance definition.*

glycan pairs. The main conformation sampled is $^1C_4$ with some transitions to ^2S_O. Only one $^1C_4 \rightarrow^4 C_1$ short-lived transition was observed in one of the residues in window 4 of the O2O definition simulation (Figure 6.14).

In the case of IdoA residues, both internal and terminal ones exhibited high flexibility in most windows and for all simulations. A likely cause of the occurrence of chair - chair flips is polymerisation [129].

Overall, it appears that the ring pucker was not affected by the biasing force, certainly not in any way that produces previously unseen shapes by distorting the pyranose ring. Flipping rates and occupancy of states was difficult to quantify, given the short length of simulation time, however the proportion of conformations observed was consistent with expected results according to the literature.

Glycosidic linkages

Glycosidic linkage parameters were calculated for all tetramer combinations in order to further characterise the shape of the glycans during the simulation and with respect to each simulation window. Results are compared to extensive calculations in the literature, calculated by Tessier [237] [234] in simulations of a similar tetrasaccharides, as shown in Figure 6.15 and validated through experimental data from NMR spectroscopy.

Similarities between the equivalent linkages of IdoA2S - GlcNS6S and vice versa are examined in this section. These linkages correspond to the DE and FE linkages with reference to the schematic in Figure 6.15 using the glycans numbered 2 and 3 for the sulfated and unsulfated tetramers respectively. The mixed sulfation groups were not examined (groups numbered 1).

The main differences between the molecule used for comparison and the tetramers studied in this work are the GlcA molecule in the reducing end, in place of IdoA (or IdoA2S for the sulfated tetramer) as well as the replacement of the di-sulfated hexosamines with mono-sulfated ones (GlcNS in place of GlcNS6S).

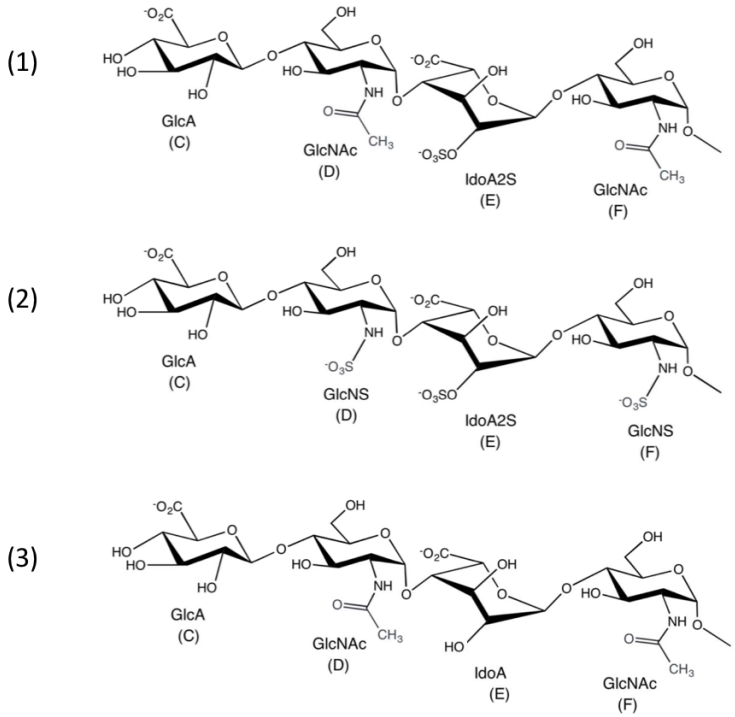
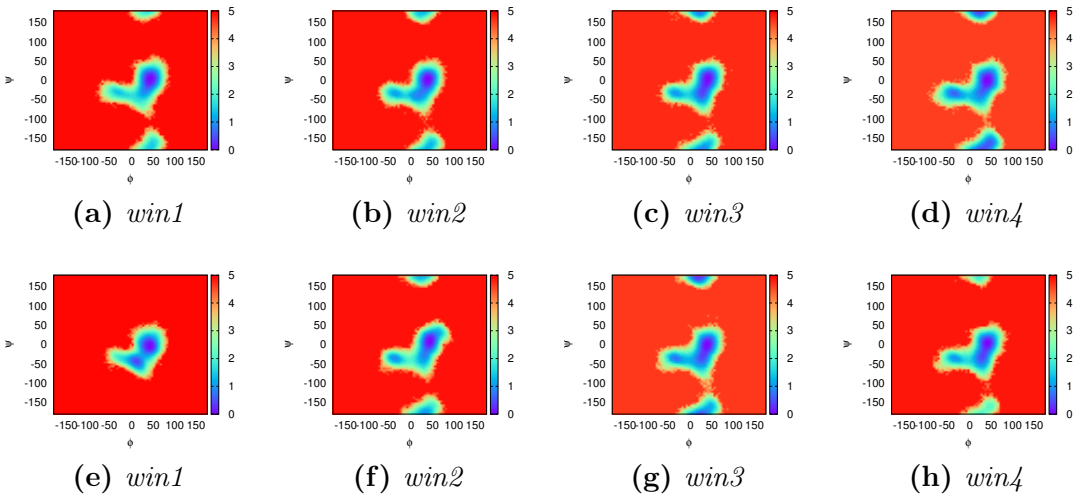


Figure 6.15: ***Tetramers used for comparison from the literature.*** *The core structure of the GAG tetrasaccharides (1-3) in sequence and chemical structure forms. Labels C-F are used to identify the monosaccharide residue components, regardless of sulfation [237] [234].*

Results shown in figures (6.16 - 6.21) from the present computational experiments are the combined heat maps of both identical molecules in each simulation (a combined total of $200ns$).

Torsion angle results for the GlcNS6S - IdoA2S linkage (Figure 6.16) closely resemble the equivalent (GlcNS - IdoA2S) results mentioned above in the literature, with the main rotamer characterised by ϕ angle approximately at 45 degrees and ψ at 0 degrees. Two minor rotamers also appear, at (-45, -15) and (45, +/-180). Similar results are obtained for the internal/terminal GlcNS6S - IdoA2S linkage, although there is no direct comparison to any of the linkages in the work by Tessier [237].

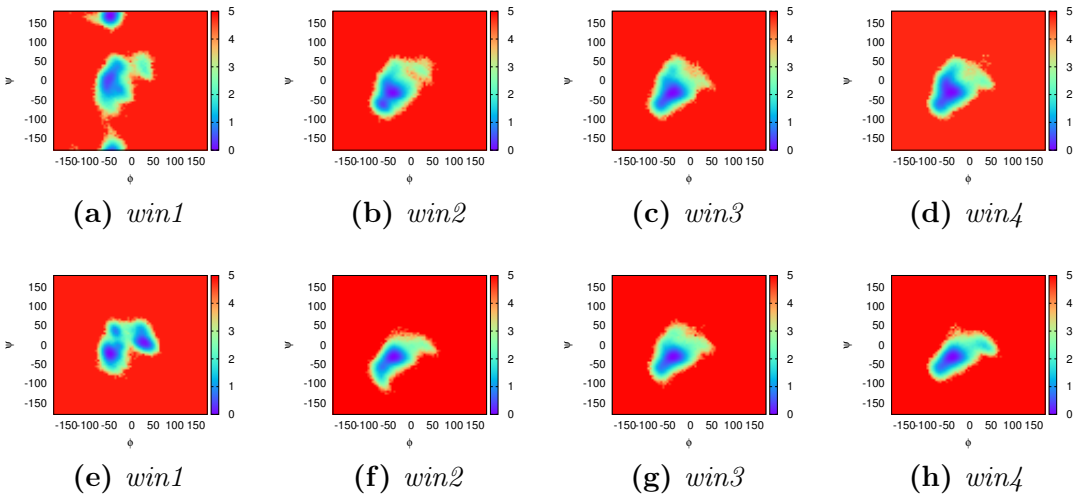
In the case of non-reducing IdoA2S - GlcNS6S linkage, results in Tessier [237]



(a) *win1* **(b)** *win2* **(c)** *win3* **(d)** *win4*

(e) *win1* **(f)** *win2* **(g)** *win3* **(h)** *win4*

Figure 6.16: ***Calculated torsion angles characterising the sulfated tetramer reducing end GlcNS6S - IdoA2S linkage.*** *Comparison between windows. Top row O2O, bottom CoM*

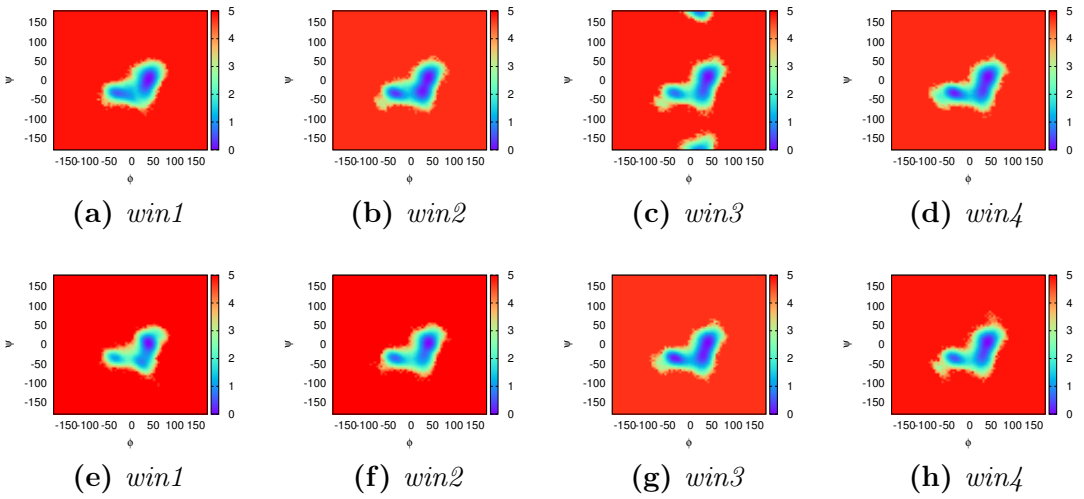


(a) *win1* **(b)** *win2* **(c)** *win3* **(d)** *win4*

(e) *win1* **(f)** *win2* **(g)** *win3* **(h)** *win4*

Figure 6.17: ***Calculated torsion angles characterising the sulfated tetramer non-reducing IdoA2S - GlcNS6S linkage.*** *Comparison between windows. Top row O2O, bottom CoM*

differ significantly according to the IdoA2S pucker conformation (computations were constrained accordingly for enhanced sampling). They are roughly comparable to results obtained in this work, mainly because in the present calculations pucker conformations are mixed in each window, since no constraints were imposed. As seen in Figure 6.17, the shape formed in the heat maps differs between windows.

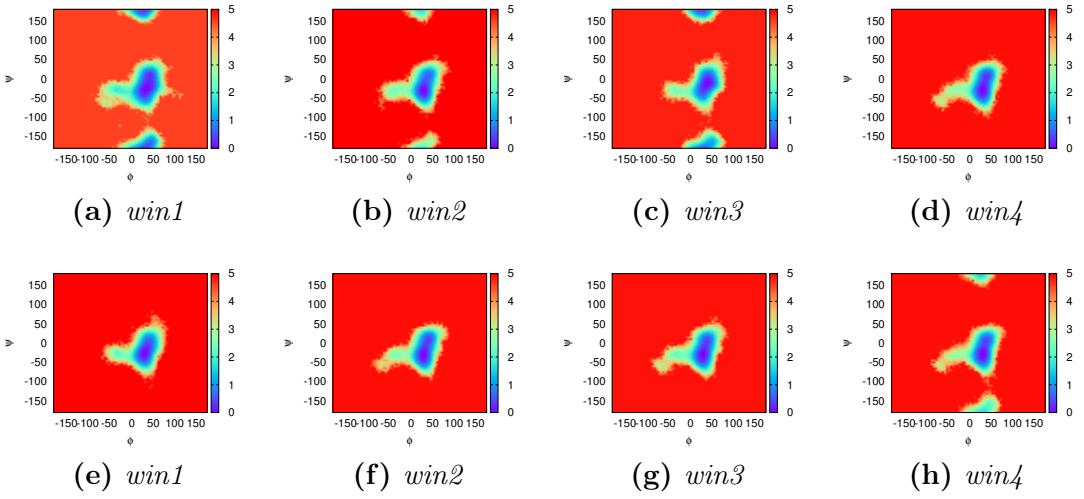


(a) *win1* **(b)** *win2* **(c)** *win3* **(d)** *win4*

(e) *win1* **(f)** *win2* **(g)** *win3* **(h)** *win4*

Figure 6.18: ***Calculated torsion angles characterising the sulfated tetramer terminal GlcNS6S - IdoA2S linkage.*** *Comparison between windows. Top row O2O distance, bottom CoM*

This is a result of variations in the pucker conformation mix for IdoA2S and the fact that simulations haven't converged with respect to the ring pucker as explained in the previous section. It is this linkage in particular that is heavily influenced by the IdoA2S pucker, most notably for the sulfated residue.



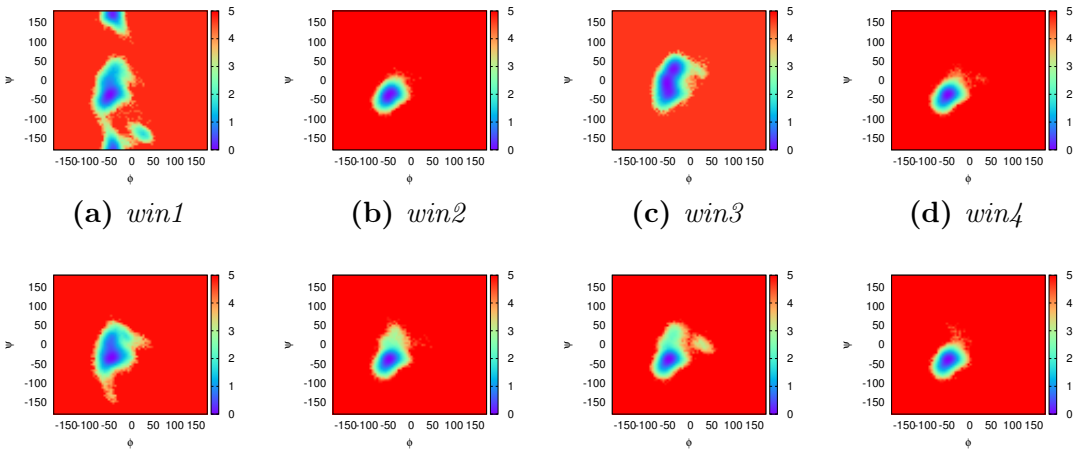
(a) *win1* **(b)** *win2* **(c)** *win3* **(d)** *win4*

(e) *win1* **(f)** *win2* **(g)** *win3* **(h)** *win4*

Figure 6.19: ***Calculated torsion angles characterising the sulfated tetramer reducing end GlcNAc - IdoA linkage.*** *Comparison between windows. Top row O2O distance, bottom CoM*

Results for unsulfated linkages (as shown in Figures 6.19 - 6.21) were identical

to those obtained by Tessier, given the great similarity of the molecules and linkages in question. Neither simulation window nor distance definition appeared to play a significant role in the calculated torsion angles when accounting for ring flexing in the IdoA residue. As observed by Tessier, the ring pucker does affect some linkages, in particular the non-reducing case (Figure 6.20), however not as much as for the sulfated case.

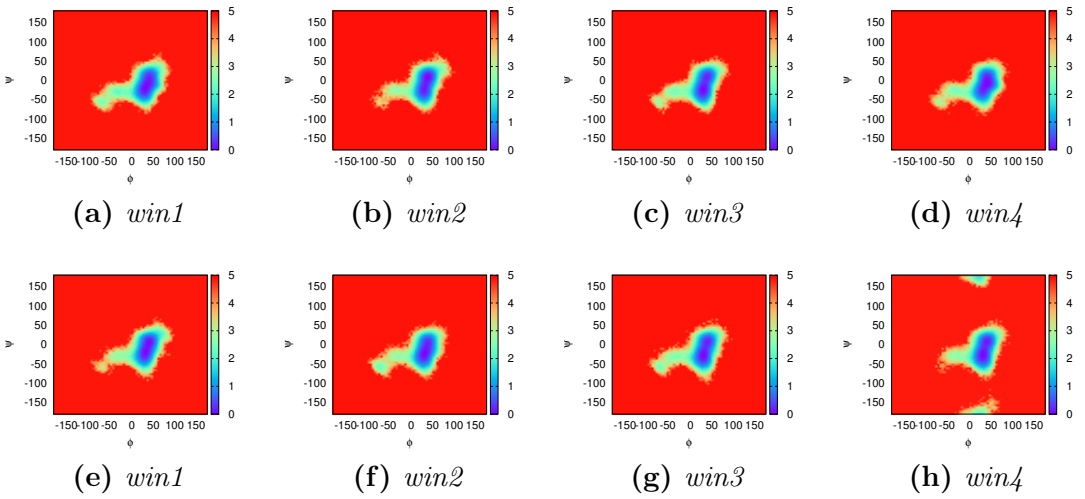


(a) *win1* **(b)** *win2* **(c)** *win3* **(d)** *win4*

(e) *win1* **(f)** *win2* **(g)** *win3* **(h)** *win4*

Figure 6.20: ***Calculated torsion angles characterising the sulfated tetramer non-reducing IdoA - GlcNAc linkage.*** *Comparison between windows. Top row O2O distance, bottom CoM*

For all linkages observed, there was overall no differences attributed to either distance definition method or window. With the exception of the non-reducing GlcNS6S - IdoA2S linkage, most are largely unaffected by IdoA conformation or sulfation. Given that this was also the case for the ring pucker, it appears that the biasing force did not distort the natural conformations of either mono or tetrasaccharides, as observed in other computational studies in the literature, which in turn have been experimentally validated.



(a) *win1* **(b)** *win2* **(c)** *win3* **(d)** *win4*

(e) *win1* **(f)** *win2* **(g)** *win3* **(h)** *win4*

Figure 6.21: ***Calculated torsion angles characterising the sulfated tetramer terminal GlcNAc - IdoA linkage.*** *Comparison between windows. Top row O2O distance, bottom CoM*

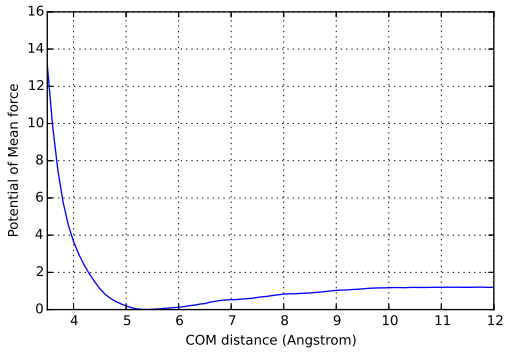
6.4.2 Simulation derived PMF curves

PMF curves were calculated directly from the ABF simulations for all monomer and tetramer systems. The calculation of the mean force and subsequent thermodynamic integration methods that produces these curves has been described in detail in Chapter 4.

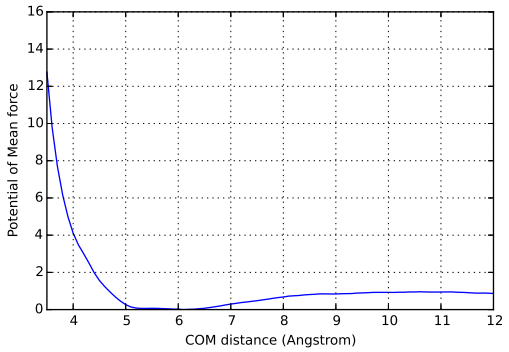
Monomers

ABF simulation derived curves for all monosaccharide pairs are shown in Figure 6.22. Due to the higher energy involved at the closest point as well as the higher number of degrees of freedom the further molecules are separated, the two endpoints of distance calculations are the most difficult to sample. In all simulations, with the exception of GlcNAc - GlcNAc, the maximum (PMF for the closest distance at 3.5 Å) had converged within a tolerance of 1 kcal/mol, that is, the value at $100ns$ was approximately the same as at 50 and $75ns$.

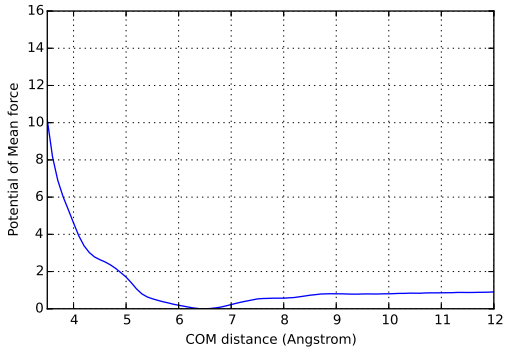
In the case of GlcNAc - GlcNAc, the maximum value did not converge as rapidly. As such, it was selected as the test case for stratified simulation (see the section on



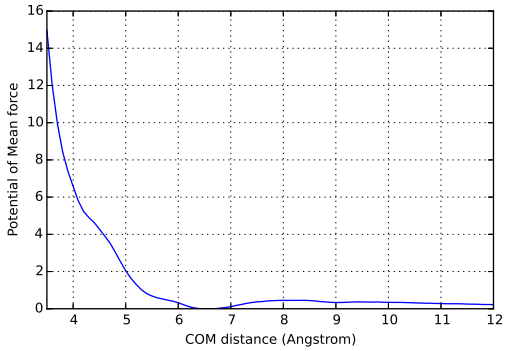
(a) *GlcNAc - GlcNAc*



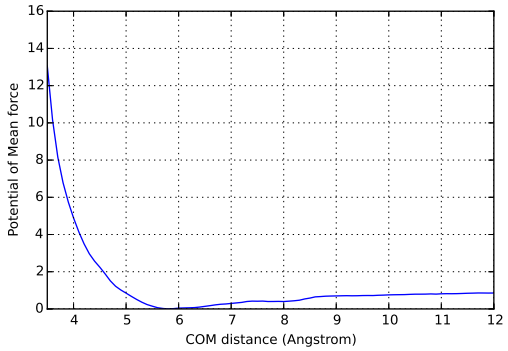
(b) *GlcNS6S - GlcNS6S*



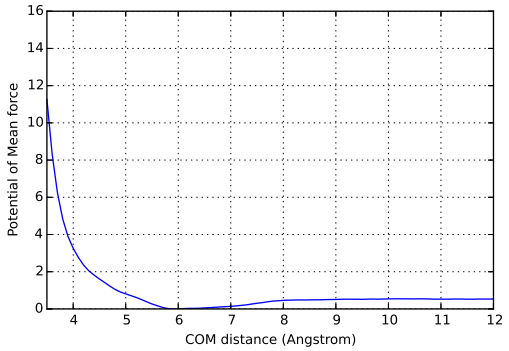
(c) *IdoA - IdoA*



(d) *IdoA2S - IdoA2S*



(e) *GlcNAc - IdoA*



(f) *GlcNS6S - IdoA2S*

Figure 6.22: ***Monomer PMFs as calculated during 100****ns* ***simulations.*** *Left: unsulfated pairs of monomers, right: sulfated pairs*

monomer simulation setup above). Comparison of the obtained PMF curves did not highlight significant differences between the stratified and the original simulation (see Figure 6.23). It was therefore decided that the precision gained was not worth the quadrupling of simulation time.

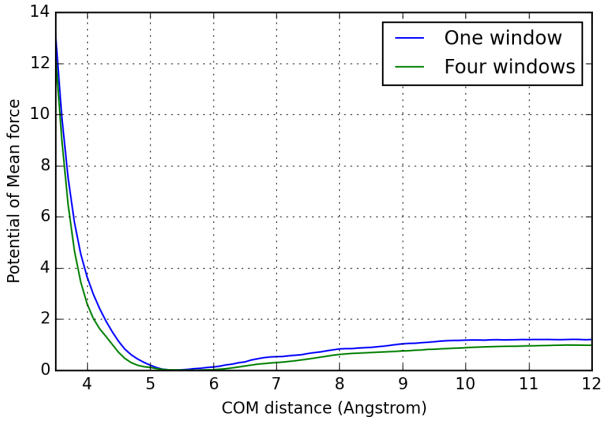


Figure 6.23: ***GlcNAc - GlcNAc simulation resulting from the implementation of 4 windows ABF.*** *Comparison with one window version*

Curves derived for IdoA and IdoA2S are clearly distinct from those of GlcNAc and GlcNS6S, mainly as a result of the minimum potential lying approximately 1 Å further for IdoA and IdoA2S. In the case of sulfation, it appears that it has a larger impact on the iduronic acid in comparison to its effect on the glucosamine. The difference between the highest PMF point (at closest proximity) and lowest separation, which varied for each pair, was between 14 and 10 kcal/mol. Curves for mixed-type of monosaccharides appear to exhibit roughly average values between those of their constituent parts.

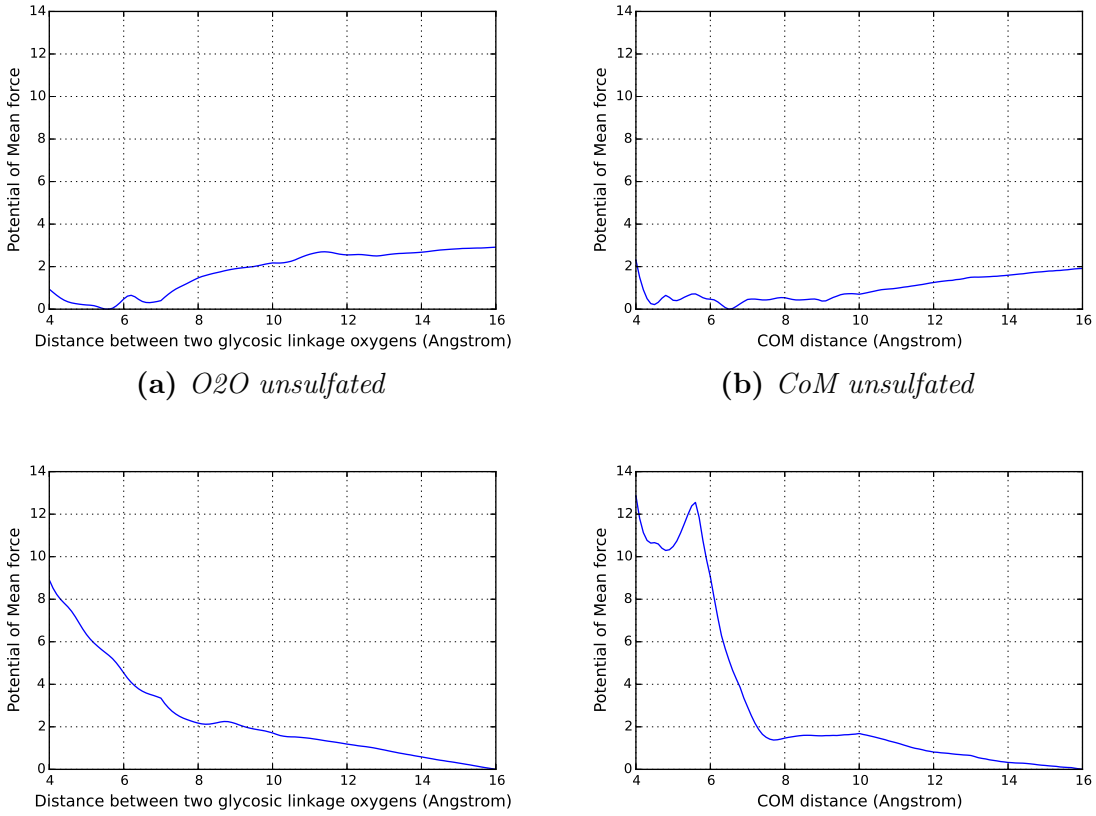
Of all the monomers investigated in this work, the sulfation of IdoA appears to have the largest impact on the overall PMF at close packing with a PMF difference of approximately 4 kcal/mol at 3.5 Å. This result is somewhat unexpected, as the unsulfated IdoA already carries negative charge, and therefore the larger charge difference in sulfated and unsulfated GlcNAc would be expected to result in higher PMF values in the case of sulfation as compared to IdoA. However, as per the discussion above with regards to distance-windows in the case of GlcNAc and as is evidenced in Figure 6.23, more precise simulations lead to slightly lower PMF values for the entirety of the curve.

It is worthwhile stressing at this point that for the PMF curves under consider-

ation, including the ones for tetramers discussed in the next section, the absolute values have no physical meaning. Only the relative potential values have meaning, as well as the shape of each curve.

Tetramers

PMF curves were also calculated directly from the ABF calculations in the case of the tetrassacharide pairs, for both distance definitions (Figure 6.24).



(a) *O2O unsulfated* **(b)** *CoM unsulfated*

(c) *O2O sulfated* **(d)** *CoM sulfated*

Figure 6.24: ***Tetramer PMFs calculated directly from ABF simulations:*** *(a) Unsulfated tetramers using O2O distance metric, (b) unsulfated tetramers using CoM distance, (c) sulfated tetramers using O2O distance and (d) sulfated tetramers using CoM distance.*

The difference between the highest and lower PMF value, which is reversed between short and long distances depending on sulfation, was between 2 and 12 kcal/mol. Comparing between the sulfated and unsulfated tetramers, it becomes apparent that the high energy point in the case of the unsulfated tetramers lies before

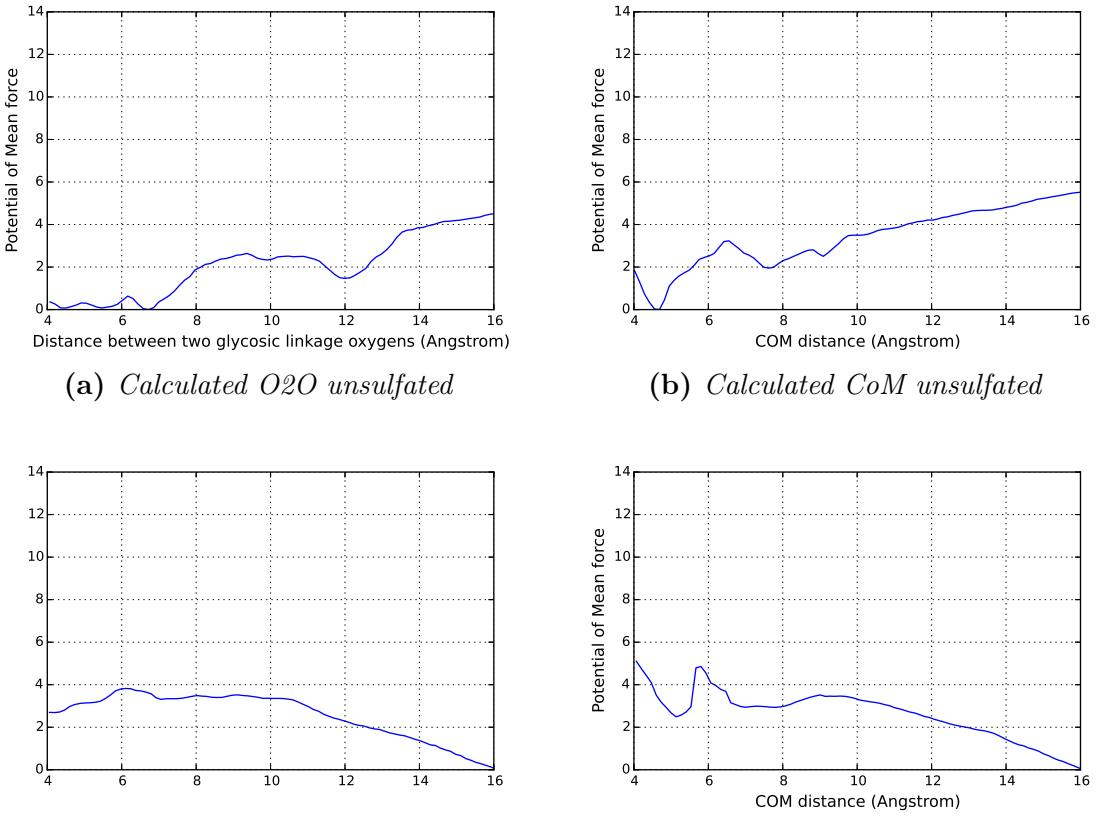
4.0 Å and that the simulation only sampled the curve close to its minimum point. The fact that the value of PMF increases with distance in the case of unsulfated tetramers is most likely a result of inadequate sampling at longer distances, despite the window method applied to reduce this effect. It could also be the effect of the non-bonded interactions cutoff at 12 Å for both van der Waals and electrostatics for which PME was used, although electrostatics effects in particular tend to drop off quickly.

By comparing results between the O2O and CoM definitions, it becomes immediately clear how important they are in determining the outcome. It appears that the CoM definition forces the tetramers to come much closer, hence the larger PMF value up to 7 Å.

PMF curves based on the monomer model

Tetramer PMF curves were also derived based on monomer distances, using the produced tetramer distance trajectory that included distances between every centre of mass, as well as curves fitted to the actual monomer PMF curves as explained in the previous section of this chapter (see Equation 6.3).

Overall, predicted curves were more faithful to the neutral tetrassacharide originally calculated PMF curves, and the shape and relative size of the curve was more faithfully reproduced for the case of O2O distance definition. The absolute error for each PMF calculation based on monomer PMFs is shown in Figure 6.26. Calculation of the non-sulfated tetramer PMFs based on O2O distance has the lowest error, at under 2 kcal/mol for the entire range of the reaction coordinate (Figure 6.26a). It performs best for short-range dynamics and has a root mean squared error (RMSE) of 0.74 Å. Calculations derived from the CoM distance for the unsulfated tetramers (Figure 6.26b) exhibit roughly twice as much absolute error with similar performance on short and long-range dynamics as in the case of O2O. In this case, RMSE is significantly higher at 2.7 Å.



(a) *Calculated O2O unsulfated*

(b) *Calculated CoM unsulfated*

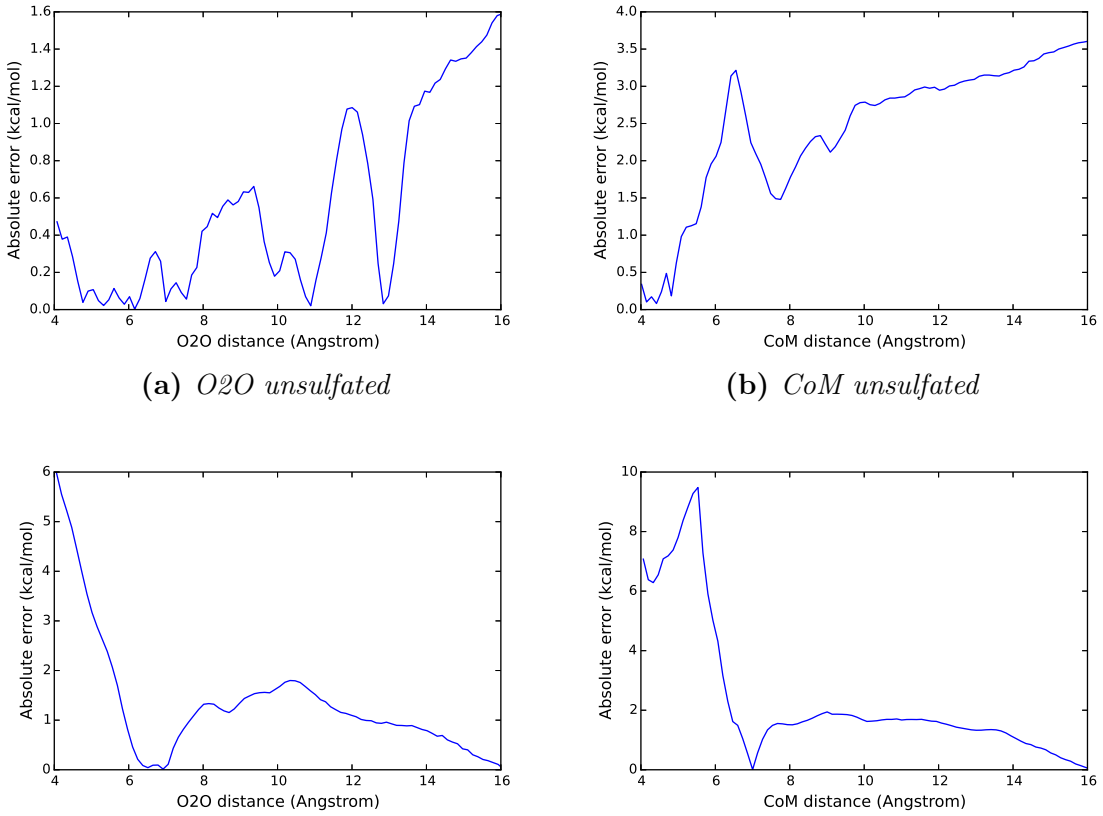
(c) *Calculated O2O sulfated*

(d) *Calculated CoM sulfated*

Figure 6.25: ***Tetramer PMFs as calculated from monomer positions and their respective pairwise PMF curves.*** *(a) Unsulfated tetramers using O2O distance metric, (b) unsulfated tetramers using CoM distance, (c) sulfated tetramers using O2O distance and (d) sulfated tetramers using CoM distance.*

In the case of the sulfated HS tetramers, the monomer model performs significantly better for long-range dynamics, with an absolute error of under 2 kcal/mol between 6 and 16 Å , as opposed to short range. As observed above in the case of the unsulfated tetramers, curves calculated for the O2O method (Figure 6.26c) have a lower overall error (RMSE of 1.8 Å) in comparison to curves derived from the CoM method (Figure 6.26d) where the RMSE was 3.2 Å.

From the monomeric SO3 PMF curves (Figures 6.22b, 6.22d, 6.22f), it becomes obvious that they cannot reproduce the sulfated tetramer calculations, as the repulsion begins at under 5 Å for all the monosaccharides. This is one cause of under-estimation for short-range interactions.



(a) *O2O unsulfated* **(b)** *CoM unsulfated*

(c) *O2O sulfated* **(d)** *CoM sulfated*

Figure 6.26: ***Absolute error of monomer PMF model method.*** *(a) Unsulfated tetramers using O2O distance metric, (b) unsulfated tetramers using CoM distance, (c) sulfated tetramers using O2O distance and (d) sulfated tetramers using CoM distance.*

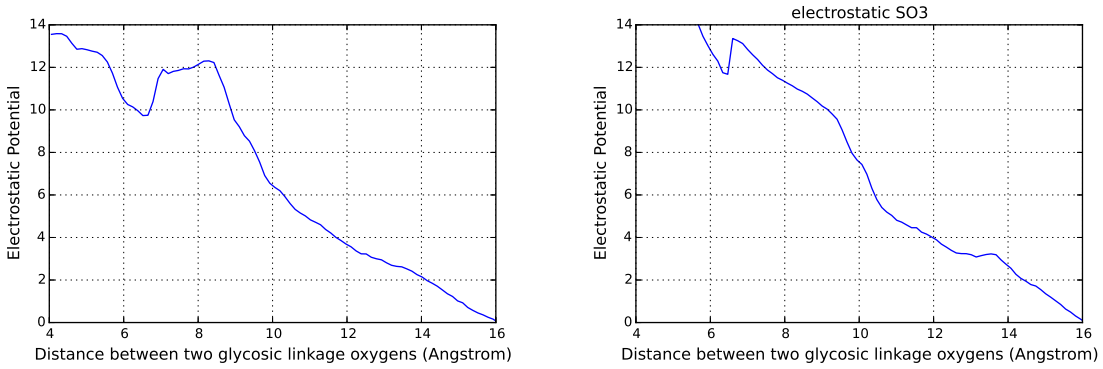
This model further suffers from a configurational as well as an exposure bias. The configurational bias is attributed to the fact that, as part of the tetramers, there are limited configurations that can be adopted by constituent residues. Monomers have the ability to reorient in ways the corresponding tetrasaccharide residues cannot do so. The exposure bias springs from the fact that ion and solvent accessibility to the GAGs is not in any way addressed in this model. For the tetramers, there's a smaller surface area exposed directly to the solvent and ions. Polymerisation alone increases ion and solvent exclusion, which can only further increase with the addition of another GAG in the vicinity.

Another caveat of applying this method was the need to extrapolate the monomer fitted curves to larger distances, rendering long-range results less accurate. However,

production ABF simulation runs for longer distance would have meant even larger simulation boxes with more solvent. In that case, longer simulation times would become necessary and usage of ABF windows would not be optional.

PMF curves based on the electrostatic model

A simple electrostatic model was also used in place of the fitted monomer curves for the purpose of calculating tetrasaccharide PMF curves (see Equation 6.5). Since it is solely based on electrostatic interactions, it was not applied to the case of neutral tetramers, as only the IdoA residues are charged. The electrostatic model tended to amplify the charge component of the molecules and produced larger PMF values in close proximity, using either distance definition method.

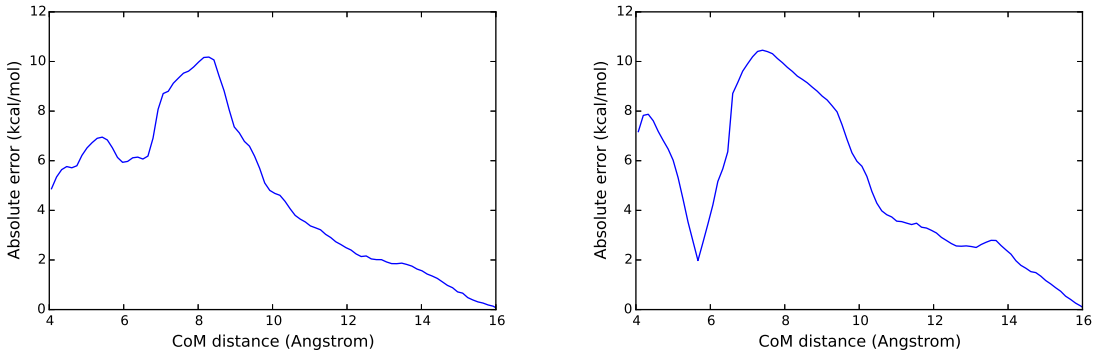


(a) *Calculated O2O* **(b)** *Calculated CoM*

Figure 6.27: ***Tetramer PMFs as calculated from monomer positions and the electrostatic model.*** *(a) sulfated tetramers using O2O distance and (b) sulfated tetramers using CoM distance.*

The absolute error for both distance definitions for the electrostatic model is shown in Figure 6.28. As with the monomer model (sulfated tetramers), short range PMF is overestimated. However the trend continues to mid-range interactions as well, starting to drop off after 8 Å for both reaction coordinates. RMSE was 5.5 Å for the O2O distance definition and 5.8 Å for the CoM distance definition model.

It became apparent that the electrostatic model was too crude in comparison to the monomer-derived one, as it fails to account for attractive forces (van de



(a) *O2O* **(b)** *CoM*

Figure 6.28: ***Absolute error of electrostatic PMF model method.*** *(a) sulfated tetramers using O2O distance and (b) sulfated tetramers using CoM distance.*

Waals) and solvent dampening effects that are implicitly included in monomer PMF calculations. It therefore overestimates the charge contribution to the repulsion experienced by the charged residues of the tetramers.

Of the two PMF predictive models proposed in this work, the monomer model proved to outperform the simple electrostatics model, despite its own limitations. Both models don't explicitly account for the influence of ions or the complexity of GAG geometry and their flexibility. However, the monomer model captures the solvent and ion contribution to a certain extend. O2O distance definition proved to be more robust for use in the predictive models in all cases, in comparison to CoM distance.

6.5 Summary of findings

Overall, conformational analysis results for all HS glycans show good accordance with previous validated *in silico* investigations [20] [237] [234]. However, there was no convergence with regards to ring shape exploration, given the microsecond scale of ring flexing. This was deemed acceptable, given that the focus of this investigation was on GAG interaction rather than GAG conformation characterisation and the working hypothesis of electrostatics and other non-bonded interactions being the

main determinants of the pairwise interaction dynamics.

In the case of tetramers, glycosidic linkage angle parameters were also investigated and found in good accordance with other computational and experimental studies. There were no significant distortion effects of the close proximity of the glycans to either pucker or glycosidic linkage angle conformation space.

The ABF biasing method was applied for the calculation of potential of mean force (PMF) curves with regards to a reaction coordinate defined as a distance metric between a pair of glycans in the case of both monomer and tetramer simulations. Using a window-ABF method was deemed superfluous for monomer simulations but vital in the case of tetramer pairs. Two distance definitions were used in the tetrasaccharides, one based on a glycosidic oxygen distance (O2O distance) and another based on the centre of mass of two residues in the middle of the HS strands (CoM distance). Distance definition had no effect on glycan conformation parameters as described above.

PMF curves were calculated from the ABF simulations for all monomer and tetramer systems. Although sulfation appeared to play a minor role in monomer interactions, its importance was striking in tetramer PMF results, which revealed high energy spikes in short range interactions for the sulfated tetramers. Conversely the unsulfated tetramers showed significantly less short-range resistance to attraction. Results indicate that, as expected, unsulfated HS strands are able to approach each more easily before experiencing strong repulsive forces.

Two potential models of PMF re-construction for tetramer simulations, were constructed: the first one based on a monomer PMF model and the second one based on simple electrostatics. Overall, the predictive accuracy of the monomer model was superior to that of the electrostatic model. In all models the overall accuracy was highest in the case of O2O over that of CoM distance definition.

Chapter 7

Discussion on the two MD models

This chapter provides an overarching discussion on the relationship between the two models presented in chapters 5 and 6 and the way in which they complement each other. This is followed by a discussion on each model's capacity to address the original tasks as they were set out at the end of Chapter 3. The chapter ends with a presentation of the limitations of this work as they apply to both models.

7.1 Glycan in flow and glycan as a charged polymer brush

The glycan in physiological flow model assumes that the full extend of the GAG strands experience flow along their entire length, from the attachment points on the transmembrane proteoglycans through to the very ends, hundreds of sugar rings above. This model, however, does not capture the effects of varying GAG density that is expected to be at its highest close to the EC surface and at its lowers in the end regions of the glycocalyx, given that the thickness of glycocalyx can be of several microns, exceeding the thickness of the endothelial cell itself. This view of the glycocalyx as a layer of decreasing fibre density is supported by high resolution electron microscopy evidence as seen in Figure 7.1.

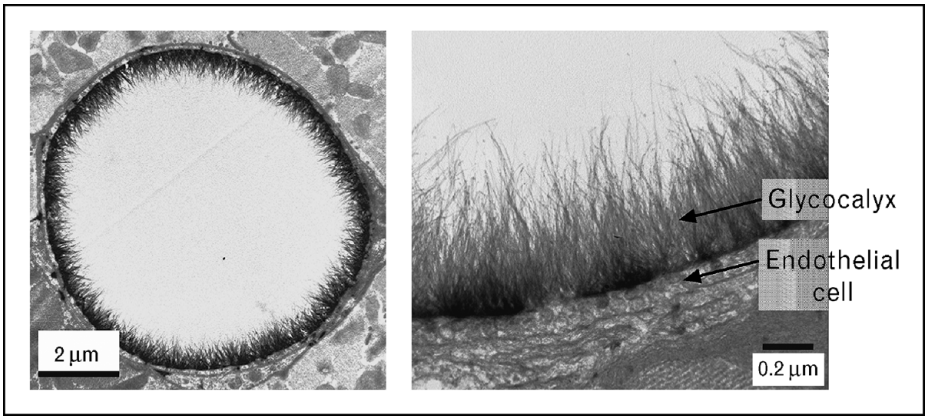


Figure 7.1: ***High resolution electron microscopy image:*** *The endothelial glycocalyx in a coronary capillary [238].*

Aside from the WSS of the blood flow, ECs, particularly in larger vessels, also experience circumferential stress due to the pulsatile nature of the flow. Other forces act on ECs, for example as a result of the passing of red and white blood cells, mostly felt in tight fitting capillaries. Furthermore, the presence of devices such as stents can result in high loads being applied on treated vessels. For such reasons, and assuming that the EGL does not collapse under the weight of such forces and allowing for the its capacity to be sufficiently restored, it becomes apparent that the compressible EGL can exist in states of low and high fibre density.

The working hypothesis in the present work is that the furthest regions of the EGL from the apical cell membrane are in direct exposure to flow, whereas the low density regions close to the cell membrane experience mainly forces of electrostatic repulsion as a result of closer packing, which can be enhanced by flow as well as other forces acting on the vessel wall. A further assumption is that dense regions of the EGL do not directly experience flow, as it becomes largely attenuated close to the cell membrane, in accordance with the dominant view in the literature, by which fluid WSS is zero near the cell membrane.

Given the above hypothesis, the glycan in flow model attempts to elucidate the response of glycans under the direct effect of flow, whereas the polymer brush model attempts to model the dense regions of the EGL and the effects of charge has on

their dynamics, irrespective of the force (be it flow or otherwise) that causes the high density packing.

7.2 Aims of this thesis

The overarching aims and corresponding specific tasks of this work were presented following the end of the literature review, in the final section of Chapter 3. The next two subsections are a discussion on how the tasks were implemented in accordance with the original aims, restated here:

- **Aim 1:** To determine whether and how, physiological flow can effect the properties of the GAG components of the EGL.
- **Aim 2:** To investigate whether and how physiological flow can effect the anchoring glycoprotein and whether shear stress can be transmitted via the glycoprotein to the cytoskeleton.
- **Aim 3:** To assess the impact of sulfation, and therefore charge, on the compressibility of the EGL
- **Aim 4:** To investigate novel ways of calculating the effect of sulfation based on individual building blocks of the EGL.

A discussion on limitations for each model and this work more broadly is given in the final section of this chapter.

7.2.1 Tethered Glycan Under Physiological Flow

In accordance with aims 1 and 2, a fully atomistic glycocalyx in physiological flow model was constructed and MD simulations took place with the aim to accomplish four specific tasks. They are presented here again alongside a discussion on how each of them was addressed in the glycan in flow model:

Task 1: Add physiological flow to existing atomistic glycocalyx model and evaluate the resulting velocity profile.

Flow was imposed on an existing model [157] by the application of an accelerating force to water oxygen molecules. The glycan was tethered by coordinate restraints on the asparagine amino-acid end to simulate the resistance to flow of the glycocalyx fibres. In order to investigate the impact of the accelerating force on the resulting flow velocity, two force regimes were investigated, resulting in two separate trajectories, one of low ($v_{x(low)} = 0.30 \pm 0.14 m/s$) and one of high ($v_{x(high)} = 0.50 \pm 0.15 m/s$) average velocity. The effect of 10% reduction in acceleration translated in 60% reduction on the average flow velocity. Moreover, the velocity profile along the length of the glycan was calculated for 7 Å -wide slabs of the simulation box. For both flow regimes, the velocity was lower in areas of the box with high glycan density and lower towards the tethered end and just above the uppermost reach of the glycan within the water box.

Task 2: Establish the validity of the model by performing fully converged simulations and comparing to gold-standard experimentally derived glycan conformational properties.

The resulting MD trajectories in the diffusion and flow simulations were deemed, at $200 ns$ each to have sufficiently converged, given that each conformational state of inter-residue orientation (quantified as glycosidic torsion angles) was sampled at least twice during the course of the simulations. Calculated conformational properties of the glycan under investigation via the resulting MD trajectories were compared to results from NMR experiments for a similar N-branched oligosaccharide. Experimentally derived properties such as dihedral angles and NOE distances were found to be in good accordance with simulation results in case of both the diffusion and

flow regimes.

Task 3: Investigate the orientation and conformational properties of the glycocalyx under the effect of flow and compare with a diffusion-only model.

The conformational analysis described above further illustrated that, within the time-frame of the simulations carried out, flow does not significantly alter the conformation space explored by the glycan, as indicated by glycosidic torsion angles and proton distance calculations. In only one case, a linkage explored certain additional angle ranges in the flow regime. This could, however, be a result of enhanced sampling the flow provides, rather than actual change in conformation properties. It is possible, therefore, that the flow could provide the additional energy required for the glycan to explore more conformations for a given time-frame as compared to the flow.

Task 4: Investigate the conformational properties of the anchoring glycoprotein (represented by the terminal amino-acid connecting the glycan component of the glycocalyx) and compare with the diffusion-only model as above.

The introduced "resistance" to alignment with flow produced by the tethering of the glycan, gives rise to the difference in angle distribution, as measured from the asparagine end to the branches of the glycan. It was further shown that flow translates to a population shift for torsion angles in the asparagine end amino-acid, in accordance with the flow direction. This result shows that the force applied to the water oxygen molecules does translate into a force acting on the tethered end of the glycan, representing here the anchoring proteoglycan. This finding is in agreement with the dominant consensus according to which the fluid WSS is transmitted via the EGL to the cytoskeleton.

7.2.2 GAG interactions in close proximity

In accordance with aims 3 and 4, a fully atomistic glycocalyx as represented by sulfated and unsulfated HS glycans was constructed and MD simulations took place with the aim to accomplish five specific tasks. They are presented here again alongside a discussion on how each of them was addressed in the charged glycan model.

Task 1: Establish the validity of the above model by performing fully converged simulations and comparing to gold-standard experimentally derived glycan conformational properties.

Conformational analysis results for all HS glycans show good accordance with previous *in silico* investigations by Sattelle et al. [20]. This was deemed sufficient validation, given that the extensive glycan simulations in the work by Sattelle et al. were thoroughly validated through NMR and laser-light scattering experiments and were found to be in good agreement. In the present model, there was no convergence with regards to ring shape exploration, given the microsecond scale of ring flexing. This was deemed acceptable, given that the focus of this investigation was on GAG interaction rather than GAG conformation characterisation.

Task 2: Using a novel enhanced sampling method, calculate the free energy of closely packed monomer components of HS as a function of distance between their centre of mass.

We utilised ABF calculations for the calculation of PMF curves along a reaction coordinate defined as a distance metric (centre of mass) between a pair of HS monomers in the presence and absence of sulfation. The reaction coordinate was investigated in the region between 3.5 and 12 Å. The difference between the highest PMF point (at closest proximity) and lowest separation, which varied for each pair,

was between 14 and 10 kcal/mol. Although sulfation appeared to play a minor role in monomer interactions, it varied depending on the types of monomers involved in each interacting pair, with the iduronate (IdoA) monomer pair PMF exhibiting the largest difference (5 kcal/mol) between the highest and lowest PMF values, as a result of sulfation.

Task 3: Using the same enhanced sampling method, calculate the free energy of a system of closely packed GAGs as a function of density (represented here as distance between two HS strands). Compare between sulfated and unsulfated GAG regions.

We utilised ABF calculations for the calculation of PMF curves along a reaction coordinate defined as a distance metric between a pair of HS tetramers in the presence and absence of sulfation. Two distance definitions were used in the tetrasaccharides, one based on a glycosidic oxygen distance (O2O distance) and another based on the centre of mass of two residues in the middle of the HS strands (CoM distance). The reaction coordinate was investigated in the region between 4 and 16 Å. A 4 window ABF system was utilised in order to adequately and uniformly sample each region of the reaction coordinate space. In the case of unsulfated tetramers, the distance space explored (by both definitions) was potentially not wide enough in the short-range side to include the sharp repulsive part of the tetramer interaction. Conversely, in the case of sulfated tetramers, the distance space explored was not wide enough on the long-range side to include the PMF plateau at longer distances. Both of these observations highlight the need for longer and wider-range ABF simulation in order to capture the short and long range dynamics of the systems. Overall the effect of sulfation on tetramer interactions was striking in comparison to monomer PMFs.

Task 4: Investigate the effect of close proximity on the conformational properties of HS and its constituent monomers.

Calculated glycosidic linkage angle and puckering parameters for both monomers and tetramers were in good accordance with experimental results as described in Objective 1. There were no significant effects of the close proximity of the glycans to either pucker or glycosidic linkage angle conformation space. Distance definition for tetramer ABF simulations had no effect on glycan conformation and puckering overall shape. Given these results, the model can be confidently shown to produce realistic HS conformations as no distortion effects were observed as a result of the enhanced sampling method.

Task 5: Construct and evaluate a model for the calculation of free energy of sulfated and unsulfated strands of HS based on the types of monomers it is composed of.

Two potential models of PMF re-construction for tetramer simulations were constructed: The first model was based on linear interpolation (and where needed, extrapolation) using the monomer PMF curves pairwise for each tetrasaccharide residue. The second model was based on simple electrostatics between each residue of the tetrasaccharides. Overall, the predictive accuracy of the monomer model was superior to that of the electrostatic model. The electrostatic model was not appropriate for unsulfated tetramers. In the case of sulfated tetramers, short-range interactions were underestimated by the monomer model but overestimated by the electrostatic model. In all models the overall accuracy was highest in the case of O2O over that of CoM distance definition. The monomer PMF re-construction model would benefit from longer monomer simulations exploring a wider range of the reaction coordinate space, thus limiting the need for extrapolation.

7.3 Limitations

7.3.1 Tethered Glycan Under Physiological Flow

This simple branched glycan model does not involve the complexity of multiple strands of long, negatively charged glycosaminoglycans and therefore explores no aspect of their interaction in the electrolyte solution. Although the flow regime is the most realistic, in terms of magnitude, achieved in MD simulations of the glycocalyx to-date, the solute velocities produced by the water oxygen acceleration method remain high with regards to the magnitude expected near the vessel wall. Furthermore, the simulation box was not large enough to study the difference in flow velocity directly above (and below, due to periodic boundary conditions) the glycan.

There are further challenges with regards to the flow implementation method. Firstly, it is hard to calibrate, and requires a trial and error approach with results that are not re-usable in different (for example larger) systems. Secondly, this technique has severe limitations for scaling up, besides the well known issues of using large water boxes with explicit solvent that already increase simulation time dramatically even in the absence of flow.

Finally, the strict restraints imposed on the asparagine end limit the orientation space (with regards to flow and other wise) the glycan is able to explore. The energy barrier imposing the constraints was high enough to absolutely freeze the affected atoms. In a more realistic scenario the atoms would be allowed to vibrate to some degree.

7.3.2 GAG interactions in close proximity

The monomer simulations would be significantly improved by covering a wider range of distances to allow for the utilisation of resulting PMF curves in PMF prediction for HS chains. A window system such as the one used for the tetramer simulations

would be required to cover a wider range of the reaction coordinate. Furthermore, no cases of mixed (sulfated and unsulfated) pairs were investigated even though the EGL has is a mixture of both regions.

In the case of tetramer simulations, as already mentioned the range of the reaction coordinate for both definitions was not sufficient to capture the dynamics of the system either in close range (unsulfated tetramers) or short range (sufated tetramers). Furthermore, the use of one reaction coordinate rather than multiple limits the types of systems that can be investigated to particularly small ones, as the CoM and O2O distances are rough distance approximations when dealing with large HS strands.

With regards to the monomer predictive model, the ionic strength of the solution has not been taken into account, neither for the monomer model nor for the electrostatic model. The number of ions in each system of monomers as well as tetramers is different and as a result would affect the electrostatic interactions, in particular for short-range dynamics. The contribution of ions also depends on their ability to approach the glycans (exposure bias). Both the monomer and electrostatic model lack a parameter describing ion and solvent accessibility, as a result of polymerisation as well as the closeness of the glycans. This could be achieved with the inclusion of an accessible surface area measure. Furthermore, monomer configurations that do not feature in the polymerised state should not be taken into account when fitting the PMF model. In the case of the naive electrostatic model, it could further benefit from the inclusion of simple van der Waals terms as well as a term describing solvent dampening effects, both of which are implicitly included in the monomer model.

Chapter 8

Conclusions and future direction of MD EGL simulations

In this final chapter, we summarise our findings and present the conclusions for each case study we performed as presented in chapters 5 and 6. We additionally present recommendations for future work involving each of the utilised methods. Finally, we discuss the future directions of atomistic simulations whose aim is to improve the understanding of the endothelial glycocalyx layer.

8.1 Glycan in flow simulations

The computational experiments described in Chapter 5 were based on a model of a flexible glycocalyx molecule reorienting itself as a result of the flow, exerting a pulling force on the rigid transmembrane protein. This atomistic model of the endothelial glycocalyx is the first, to our knowledge, to introduce a flow velocity close to physiological levels. This is contrary to previous models of this scale that implement unrealistic velocities in the tens of m/s with the effect of inducing bending and other conformational changes on the molecules under study [173] [174] [159].

This simple branched glycan model does not involve the complexity of multiple strands of long, negatively charged glycosaminoglycans and therefore explores no

aspect of their interaction in the electrolyte solution. Despite the model's limitations, covered extensively in the previous chapter, the findings of this investigation can be used in order to assess the type and magnitude of the effects of flow we can expect on such flexible biomolecules.

8.1.1 Conclusions of glycan in flow model

- **Flow implementation** The flow implementation used in this work, where a force is added to water oxygen atoms at every timestep, was successful in producing two regimes of bulk water flow at 0.5 and 0.3 m/s, which is the closest to realistic physiological velocity levels in comparison to all other glycocalyx in flow studies to date.

- **Effect of flow on the glycan** We found that flow had no significant effect on the glycan itself, which was free to explore any conformational space, as the tethering was applied to the amino-acid end. The conformations explored by the oligosaccharide as well as their relative populations were shown to be in accordance with experimental studies for the case of both flow and diffusion conditions.

- **Effect of flow on the amino-acid** The introduction of flow, however, did alter the conformational space explored by the tethered amino-acid end of the glycan, by favouring states that allow for maximum alignment with the flow. The influence of tethering was also evident in the orientation of the glycan as a whole, which shifted towards states that corresponded to higher alignment with the flow, given the restraints imposed, as it was tethered in a way that did not allow for it to become parallel to the flow.

- **Implications of the effect of flow on the amino-acid** The change in the conformation properties of the aglycon, could indicate that, in vivo, the long strands of polysaccharides exert a *bending* force that changes the conformation

of the transmembrane proteoglycans, essentially playing the part of the initial mechanical signal transmitted through to the cytoskeleton in the process of mechanotransduction.

8.1.2 Future direction for flow simulations

Many aspects of the glycan-in-flow glycocalyx model can be improved and expanded:

- **Scaling up** Simulations with lower flow and multiple tethered glycans as well as larger (with respect to the z-axis) boxes to allow for a bigger layer of water above the glycan where shear effects are not observed. Such a system would require larger simulation times in order to accurately capture the velocity profile and measure shear-stress inside and outside the glycocalyx layer. It would also require a different flow implementation method, as discussed further in this section.

- **Improved tethering of aglycon** A more realistic investigation of the effects of flow on the proteoglycan could be achieved firstly by exploring different tethering options, such as fewer restrained atoms to allow for more degrees of freedom and lower restraint forces that would permit the tethered atoms to vibrate. Measuring the vibration properties under diffusion and flow would give an estimation of the exact energy transferred from the flow to the asparagine end.

- **Using GAGs and proteoglycans in flow simulations** The model could be further improved with the usage of actual GAGs (such as the tetramers used in Chapter 6), as well as the inclusion of the tethered protein. It is possible to include a coarse-grained representation of the large protein in order to speed up the calculations. Longer simulation times would allow for the implementation of better NOE approximations to include molecular tumbling effects on long-range NOEs [232]. It is worth noting that tumbling times would be reduced

on proteoglycans due to the size of the protein and the polymerised GAGs. These NOE estimation effects would likely be less noticeable in experimental NOEs of a proteoglycan system.

- **Implement flow on previously developed coarse grained model** A very promising route to take would be the adaptation of the Sattelle et al. [20] [158] coarse grained model of GAGs and proteoglycans that has been specifically developed and parameterised to capture effects seen in only the atomic scale, such as pyranose ring flexing. Such a model coupled with a better flow implementation would allow for great speedup on the calculations and reliable results for large systems.

- **Novel flow implementation method** It is our opinion that the flow implementation used in this work, by taking advantage of NAMD's scripting plugin, is inefficient and difficult to calibrate for new systems. One way to overcome this obstacle in implementing flow is to hardcode the implementation by customising the computational package's base code. This could be NAMD for example the Amber computational package (here, we used NAMD mostly for its scripting plugin). The flow can either be added in as a force in the same manner done within the script, or directly at the thermostat level where a bias can be added on the x-velocity component of water oxygens.

- **Pressure-driven flow** A pressure-driven flow can be implemented using a "gated" design, where a slab of water molecules is accelerated at the inlet of the simulation box. This design can take long to reach steady state, which is why it could be turned on subsequently to using an all-water acceleration method. Implementing an actual pressure gradient in MD would not work as it would not allow for periodicity in the direction of the gradient.

8.2 GAG simulations

The sulfated/unsulfated GAG model explores aspects of the glycocalyx structure not included in the flow study, namely the charged nature of several glycocalyx components as well as the interaction between them. The focus remained in the atomic scale and employed a fine grained level of detail, using state the of the art force field for GAGs [234] as well as the novel sampling method ABF [220]. This model is not concerned with the nature of the force driving the close packing, which can be a result of flow, naturally occurring density in static conditions or interaction with large cells that deform the EGL.

8.2.1 Conclusions on the GAG simulations

- **Validation of novel approach to EGL modelling** Conformational characteristics were reproduced with good validity in comparison to both experimental and *in silico* results. This adds to our confidence that the usage of ABF for studying glycans in close proximity does not produce anomalies in residue ring shapes or the internal motion of the flexible polysaccharides.

- **Effect of sulfation on the dynamics of GAG pairs** The potential of mean force (PMF) for each system with regards to a reaction coordinate (representing distance) confirmed our initial hypothesis, that two systems that vary only in the degree of sulfation will exhibit significantly different dynamics in close proximity.

- **Reaction coordinate definition** ABF simulations were implemented by definition of different reaction coordinates on the same system. PMF curves produced according to the two reaction coordinates (distance definitions) were similar but distinctive and the effects were shown to matter both in direct PMF calculations and in the case of the constructed predictive models.

- **PMF predictive models** The effect of distance definitions was further illustrated by the predictive accuracy of the monomer model, which was able to more accurately reproduce PMF curves calculated using the glycosidic oxygen distance definition. The electrostatics based model overestimated the PMF of both systems and demonstrated that, in the current, naive versions of both models, the monomer model is superior at reconstructing PMF curves for larger polysaccharides.

8.2.2 Direction of GAG simulations

- **Improvements on the reaction coordinate** Observations derived from the GAG ABF simulations give many indications for the future direction of this project. Firstly, in any future attempt to model a similar system, the reaction coordinate for unsulfated tetrasaccharide simulations should expand to cover even shorter distances, under 4 Å, assuming that either the O2O or CoM definitions of distance are used. In terms of reaction coordinates, there are several choices of univariate definitions and a more complex model would take into account two reaction coordinates, of which one would remain a distance metric and the other would characterise the overall geometric shape of the glycan (for example, a measure of end-to-end distance to characterise the internal motion of one of the glycans).

- **Improvements on the predictive model** The ionic strength of the solution has not been taken into account, neither for the monomer model nor for the electrostatic model. The number of ions in each system of monomers as well as tetramers is different and as a result would affect the electrostatic interactions, in particular for short-range dynamics. The contribution of ions also depends on their ability to approach the glycans (exposure bias). Both the monomer and electrostatic model lack a parameter describing ion and solvent accessibility, as a result of polymerisation as well as the closeness of the glycans. This

could be achieved with the inclusion of an accessible surface area measure. Furthermore, monomer configurations that do not feature in the polymerised state should not be taken into account when fitting the PMF model. In the case of the naive electrostatic model, it could further benefit from the inclusion of simple van der Waals terms as well as a term describing solvent dampening effects, both of which are implicitly included in the monomer model.

- **Expanding to other GAGs** Once the current model is further refined, the possibilities for expanding to different GAG systems are vast. Studies on the nature of hyaluran (HA), dermatan and chondroitin sulfate etc, as well as heterogeneous paired simulations can also be explored. The benefit of parameterising and fitting a simple monomer model to predict the dynamics of large systems is that there is no need for ABF simulations for systems of multiple large polymers. A further point of interest is the application of this model to large polymers, for example hexasaccharides in order to investigate the scaling properties of PMF calculation in GAGs.

8.3 Future work on EGL models

There are limits to the complexity of systems that a fully atomistic method such as molecular dynamics can be stretched to. In this work, we have investigated two major aspects of the endothelial glycocalyx morphology, namely its close coupling to the blood flow as well as its electrostatic nature and representation as a biological polymer brush. We have chosen this detailed atomistic approach because the field of glycomics and molecular simulations have in recent years evolved sufficiently to allow for this kind of investigation on GAGs.

Flow simulations involving GAGs can be optimised using the PMF curves obtained by HS in the present work, and refined by further simulations of other GAGs to define optimum packing distance (which can translate to tethering distance) for

systems of sulfated and unsulfated glycocalyx components. Both models can be repeated with more sophisticated water models (such as 4 or 5-point charge models) in order to estimate correction terms for both PMF as well as flow results.

It is, however, improbable that a representative, fully atomistic glycocalyx system, inclusive of a lipid bilayer as well as several proteoglycans and a variety of attached GAGs can be investigated in MD for an adequate simulation time in order to yield meaningful and converged results in the immediate future. Instead, we believe that only by modelling different aspects of the EGL and parameterising appropriate coarse grained models can this bottom-up approach be realised. This is indeed the work already pursued for GAGs by glycomics experts [20] [239] [158] and we believe that coupling such coarse grained models with flow as well as GAG interactions studies such as the ones of this work, the complex nature of the glycocalyx and plasma system in its full detail can begin to emerge.

Bibliography

[1] Maria Siebes and Yiannis Ventikos. The role of biofluid mechanics in the assessment of clinical and pathological observations: sixth International Bio-Fluid Mechanics Symposium and Workshop, March 28-30, 2008 Pasadena, California. *Annals of Biomedical Engineering*, 38(3):1216–24, mar 2010.

[2] Salim Yusuf, David Wood, Johanna Ralston, and K Srinath Reddy. The World Heart Federation's vision for worldwide cardiovascular disease prevention. *Lancet (London, England)*, 386(9991):399–402, jul 2015.

[3] I Halliday, M Atherton, C M Care, M W Collins, D Evans, P C Evans, D R Hose, A W Khir, C S König, R Krams, P V Lawford, S V Lishchuk, G Pontrelli, V Ridger, T J Spencer, Y Ventikos, D C Walker, and P N Watton. Multi-scale interaction of particulate flow and the artery wall. *Medical Engineering & Physics*, 33(7):840–8, sep 2011.

[4] S Laurent, P Boutouyrie, R Asmar, I Gautier, B Laloux, L Guize, P Ducimetiere, and A Benetos. Aortic stiffness is an independent predictor of all-cause and cardiovascular mortality in hypertensive patients. *Hypertension*, 37(5):1236–41, may 2001.

[5] P Gaede, P Vedel, N Larsen, G V H Jensen, H-H Parving, and O Pedersen. Multifactorial intervention and cardiovascular disease in patients with type 2 diabetes. *The New England journal of medicine*, 348(5):383–93, jan 2003.

[6] P Jankowski, G Bilo, and K Kawecka-Jaszcz. The pulsatile component of blood pressure: its role in the pathogenesis of atherosclerosis. *Blood pressure*, 16(4):238–45, jan 2007.

[7] C F Dewey. Effects of fluid flow on living vascular cells. *Journal of Biomechanical Engineering*, 106(1):31–5, feb 1984.

[8] PF Davies, A Remuzzi, and EJ Gordon. Turbulent fluid shear stress induces vascular endothelial cell turnover in vitro. *Proceedings of the*, 83(April):2114–2117, 1986.

[9] P V Lawford, Y Ventikos, A W Khir, M Atherton, D Evans, D R Hose, C M Care, P N Watton, I Halliday, D C Walker, A P Hollis, and M W Collins. Modelling the interaction of haemodynamics and the artery wall: current status and future prospects. *Biomedicine & Pharmacotherapy*, 62(8):530–5, oct 2008.

[10] T Ohashi, K Tsuyuki, and M Sato. Real-Time Observations of Actin Filament Dynamics in Cultured Endothelial Cells Exposed to Fluid Flow. *Journal of Japanese Society of Biorheology*, 17(1):24–28, 2003.

[11] M D Savery, J X Jiang, P W Park, and E R Damiano. The endothelial glycocalyx in syndecan-1 deficient mice. *Microvascular research*, 87:83–91, may 2013.

[12] B F Becker, D Chappell, D Bruegger, T Annecke, and M Jacob. Therapeutic strategies targeting the endothelial glycocalyx: acute deficits, but great potential. *Cardiovascular Research*, 87(2):300–10, jul 2010.

[13] S Weinbaum, J M Tarbell, and E R Damiano. The structure and function of the endothelial glycocalyx layer. *Annual Review of Biomedical Engineering*, 9(1):121–167, 2007.

[14] Carlos H. Miranda, Marcos de Carvalho Borges, André Schmidt, José Antônio Marin-Neto, and Antônio Pazin-Filho. Evaluation of the endothelial glycocalyx damage in patients with acute coronary syndrome. *Atherosclerosis*, feb 2016.

[15] D Chappell, M Jacob, B F Becker, K Hofmann-Kiefer, P Conzen, and M Rehm. Expedition glycocalyx. A newly discovered "Great Barrier Reef". *Der Anaesthesist*, 57(10):959–969, 2008.

[16] D. Chappell, M. Jacob, K. Hofmann-Kiefer, D. Bruegger, M. Rehm, P. Conzen, U. Welsch, and B.F Becker. Hydrocortisone preserves the vascular barrier by protecting the endothelial glycocalyx. *Anesthesiology*, 107(5):776–784, 2007.

[17] P Jabbour, C Koebbe, E Veznedaroglu, R P Benitez, and R Rosenwasser. Stent-assisted coil placement for unruptured cerebral aneurysms. *Neurosurgical focus*, 17(5):E10, nov 2004.

[18] D Fiorella, F C Albuquerque, V R Deshmukh, H H Woo, P A Rasmussen, T J Masaryk, and C G McDougall. Endovascular reconstruction with the Neuroform stent as monotherapy for the treatment of uncoilable intradural pseudoaneurysms. *Neurosurgery*, 59(2):291–300, aug 2006.

[19] D Fiorella, F C Albuquerque, H Woo, P A Rasmussen, T J Masaryk, and C G McDougall. Neuroform in-stent stenosis: incidence, natural history, and treatment strategies. *Neurosurgery*, 59(1):34–42, jul 2006.

[20] B M Sattelle, J Shakeri, and A Almond. Does microsecond sugar ring flexing encode 3D-shape and bioactivity in the heparanome? *Biomacromolecules*, 14(4):1149–59, apr 2013.

[21] J E Wagenseil and RP Mecham. Vascular extracellular matrix and arterial mechanics. *Physiological reviews*, 89(3):957–89, jul 2009.

[22] A J Lusis. Atherosclerosis. *Nature*, 407(September):233–241, 2000.

[23] M Tamminen, G Mottino, J H Qiao, J L Breslow, and J S Frank. Ultrastructure of early lipid accumulation in ApoE-deficient mice. *Arteriosclerosis, thrombosis, and vascular biology*, 19(4):847–53, apr 1999.

[24] C Hahn and M A Schwartz. Mechanotransduction in vascular physiology and atherogenesis. *Nature reviews. Molecular cell biology*, 10(1):53–62, jan 2009.

[25] K S Cunningham and A I Gotlieb. The role of shear stress in the pathogenesis of atherosclerosis. *Laboratory investigation; a journal of technical methods and pathology*, 85(1):9–23, jan 2005.

[26] J Suvatne, A I Barakat, and M E O'Donnell. Flow-induced expression of endothelial Na-K-Cl cotransport: dependence on K+ and Cl{-} channels. *Am J Physiol Cell Physiol*, 280(1):C216–227, jan 2001.

[27] Y-S J Li, J H Haga, and S Chien. Molecular basis of the effects of shear stress on vascular endothelial cells. *Journal of biomechanics*, 38(10):1949–71, oct 2005.

[28] R Ross. The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature*, 362(6423):801–9, apr 1993.

[29] P-W Fok. Growth of necrotic cores in atherosclerotic plaque. *Mathematical medicine and biology : a journal of the IMA*, 29(4):301–27, dec 2012.

[30] N J Sniadecki. A tiny touch: activation of cell signaling pathways with magnetic nanoparticles. *Endocrinology*, 151(2):451–7, feb 2010.

[31] G B Monshausen and S Gilroy. Feeling green: mechanosensing in plants. *Trends in cell biology*, 19(5):228–35, may 2009.

[32] A van der Flier and A Sonnenberg. Function and interactions of integrins. *Cell and tissue research*, 305(3):285–98, sep 2001.

[33] H Lodish, A Berk, S L Zipursky, P Matsudaira, D Baltimore, and J Darnell. Molecular Cell Biology, 2000.

[34] T Ohashi and M Sato. Remodeling of vascular endothelial cells exposed to fluid shear stress: experimental and numerical approach. *Fluid Dynamics Research*, 37(1-2):40–59, jul 2005.

[35] Narbeh Melikian, Michael D Seddon, Barbara Casadei, Philip J Chowienczyk, and Ajay M Shah. Neuronal nitric oxide synthase and human vascular regulation. *Trends in cardiovascular medicine*, 19(8):256–62, nov 2009.

[36] S Lehoux, Y Castier, and A Tedgui. Molecular mechanisms of the vascular responses to haemodynamic forces. *Journal of internal medicine*, 259(4):381–92, apr 2006.

[37] Deshun Lu and Ghassan S Kassab. Role of shear stress and stretch in vascular mechanobiology. *Journal of the Royal Society, Interface / the Royal Society*, 8(63):1379–85, oct 2011.

[38] P M Vanhoutte, C M Boulanger, and J V Mombouli. Endothelium-derived relaxing factors and converting enzyme inhibition. *The American journal of cardiology*, 76(15):3E–12E, nov 1995.

[39] Francis J Haddy, Paul M Vanhoutte, and Michel Feletou. Role of potassium in regulating blood flow and blood pressure. *American journal of physiology. Regulatory, integrative and comparative physiology*, 290(3):R546–52, mar 2006.

[40] J Ando and K Yamamoto. Effects of shear stress and stretch on endothelial function. *Antioxidants & redox signaling*, 15(5):1389–403, sep 2011.

[41] Peter F Davies. Hemodynamic shear stress and the endothelium in cardiovascular pathophysiology. *Nature clinical practice. Cardiovascular medicine*, 6(1):16–26, jan 2009.

[42] P F Davies. Flow-mediated endothelial mechanotransduction. *Physiological reviews*, 75(3):519–60, jul 1995.

[43] P M Schiffers, D Henrion, C M Boulanger, E Colucci-Guyon, F Langa-Vuves, H van Essen, G E Fazzi, B I Lévy, and J G De Mey. Altered flow-induced arterial remodeling in vimentin-deficient mice. *Arteriosclerosis, thrombosis, and vascular biology*, 20(3):611–6, mar 2000.

[44] I R Hutcheson and T M Griffith. Mechanotransduction through the endothelial cytoskeleton: mediation of flow- but not agonist-induced EDRF release. *British journal of pharmacology*, 118(3):720–6, jun 1996.

[45] A M Malek, A L Greene, and S Izumo. Regulation of endothelin 1 gene by fluid shear stress is transcriptionally mediated and independent of protein kinase C and cAMP. *Proceedings of the National Academy of Sciences of the United States of America*, 90(13):5999–6003, jul 1993.

[46] T Ishida, T E Peterson, N L Kovach, and B C Berk. MAP kinase activation by flow in endothelial cells. Role of beta 1 integrins and tyrosine kinases. *Circulation research*, 79(2):310–6, aug 1996.

[47] J. M. Muller, W. M. Chilian, and M. J. Davis. Integrin Signaling Transduces Shear Stress Dependent Vasodilation of Coronary Arterioles. *Circulation Research*, 80(3):320–326, mar 1997.

[48] K D Chen, Y S Li, M Kim, S Li, S Yuan, S Chien, and J Y Shyy. Mechanotransduction in response to shear stress. Roles of receptor tyrosine kinases, integrins, and Shc. *The Journal of biological chemistry*, 274(26):18393–400, jun 1999.

[49] I S Bhullar, Y S Li, H Miao, E Zandi, M Kim, J Y Shyy, and S Chien. Fluid shear stress activation of IkappaB kinase is integrin-dependent. *The Journal of biological chemistry*, 273(46):30544–9, nov 1998.

[50] Akira Katsumi, A Wayne Orr, Eleni Tzima, and Martin Alexander Schwartz. Integrins in mechanotransduction. *The Journal of biological chemistry*, 279(13):12001–4, mar 2004.

[51] Eleni Tzima, Mohamed Irani-Tehrani, William B Kiosses, Elizabetta Dejana, David A Schultz, Britta Engelhardt, Gaoyuan Cao, Horace DeLisser, and Martin Alexander Schwartz. A mechanosensory complex that mediates the endothelial cell response to fluid shear stress. *Nature*, 437(7057):426–31, sep 2005.

[52] S Gudi, J P Nolan, and J A Frangos. Modulation of GTPase activity of G proteins by fluid shear stress and phospholipid composition. *Proceedings of the National Academy of Sciences of the United States of America*, 95(5):2515–9, mar 1998.

[53] Charles R White and John A Frangos. The shear stress of it all: the cell membrane and mechanochemical transduction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 362(1484):1459–67, aug 2007.

[54] S Bergaya, P Meneton, M Bloch-Faure, E Mathieu, F Alhenc-Gelas, B I Levy, and C M Boulanger. Decreased Flow-Dependent Dilation in Carotid Arteries of Tissue Kallikrein-Knockout Mice. *Circulation Research*, 88(6):593–599, mar 2001.

[55] Rosario Maroto, Albert Raso, Thomas G Wood, Alex Kurosky, Boris Martinac, and Owen P Hamill. TRPC1 forms the stretch-activated cation channel in vertebrate cells. *Nature cell biology*, 7(2):179–85, feb 2005.

[56] Yun Fang, Gernot Schram, Victor G Romanenko, Congzhu Shi, Lisa Conti, Carol A Vandenberg, Peter F Davies, Stanley Nattel, and Irena Levitan. Functional expression of Kir2.x in human aortic endothelial cells: the dominant role of Kir2.2. *American journal of physiology. Cell physiology*, 289(5):C1134–44, nov 2005.

[57] J Ando and K Yamamoto. Flow detection and calcium signalling in vascular endothelial cells. *Cardiovascular research*, 99(2):260–8, jul 2013.

[58] Thomas Weimbs. Polycystic kidney disease and renal injury repair: common pathways, fluid flow, and the function of polycystin-1. *American journal of physiology. Renal physiology*, 293(5):F1423–32, nov 2007.

[59] K Kim, I Drummond, O Ibraghimov-Beskrovnaya, K Klinger, and M A Arnaout. Polycystin 1 is required for the structural integrity of blood vessels. *Proceedings of the National Academy of Sciences of the United States of America*, 97(4):1731–6, feb 2000.

[60] Carlo Iomini, Karla Tejada, Wenjun Mo, Heikki Vaananen, and Gianni Piperno. Primary cilia of human endothelial cells disassemble under laminar shear stress. *The Journal of cell biology*, 164(6):811–7, mar 2004.

[61] P J Butler, G Norwich, S Weinbaum, and S Chien. Shear stress induces a time- and position-dependent increase in endothelial cell membrane fluidity. *American journal of physiology. Cell physiology*, 280(4):C962–9, apr 2001.

[62] M Gouverneur, B van den Berg, M Nieuwdorp, E Stroes, and H Vink. Vasculoprotective properties of the endothelial glycocalyx : effects of fluid shear stress. *Journal of Internal Medicine*, pages 393–400, 2006.

[63] M A Schwartz and R K Assoian. Integrins and cell proliferation: regulation of cyclin-dependent kinases via cytoplasmic signaling pathways. *Journal of cell science*, 114(Pt 14):2553–60, jul 2001.

[64] Bingmei M Fu and John M Tarbell. Mechano-sensing and transduction by endothelial surface glycocalyx: composition, structure, and function. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine*, 5(3):381–90, 2013.

[65] Rob J Dekker, Reinier A Boon, Mariska G Rondaij, Astrid Kragt, Oscar L Volger, Yvonne W Elderkamp, Joost C M Meijers, Jan Voorberg, Hans Pannekoek, and Anton J G Horrevoets. KLF2 provokes a gene expression pattern that establishes functional quiescent differentiation of the endothelium. *Blood*, 107(11):4354–63, jun 2006.

[66] A R Brooks, P I Lelkes, and G M Rubanyi. Gene expression profiling of vascular endothelial cells exposed to fluid mechanical forces: relevance for focal susceptibility to atherosclerosis. *Endothelium : journal of endothelial cell research*, 11(1):45–57, 2004.

[67] J-J Chiu and S Chien. Effects of disturbed flow on vascular endothelium: pathophysiological basis and clinical perspectives. *Physiological reviews*, 91(1):327–87, jan 2011.

[68] John M Tarbell, Zhong-Dong Shi, Jessilyn Dunn, and Hanjoong Jo. Fluid Mechanics, Arterial Disease, and Gene Expression. *Annual review of fluid mechanics*, 46:591–614, jan 2014.

[69] N Adhikari, N Charles, U Lehmann, and J L Hall. Transcription factor and kinase-mediated signaling in atherosclerosis and vascular injury. *Current atherosclerosis reports*, 8(3):252–60, may 2006.

[70] Roland Kaunas, Shunichi Usami, and Shu Chien. Regulation of stretch-induced JNK activation by stress fiber orientation. *Cellular signalling*, 18(11):1924–31, nov 2006.

[71] D G Harrison, J Widder, I Grumbach, W Chen, M Weber, and C Searles. Endothelial mechanotransduction, nitric oxide and vascular inflammation. *Journal of internal medicine*, 259(4):351–63, apr 2006.

[72] Frederick R Maxfield and Ira Tabas. Role of cholesterol and lipid organization in disease. *Nature*, 438(7068):612–21, dec 2005.

[73] A Wayne Orr, Rebecca Stockton, Michael B Simmers, John M Sanders, Ian J Sarembock, Brett R Blackman, and Martin Alexander Schwartz. Matrix-specific p21-activated kinase activation regulates vascular permeability in atherogenesis. *The Journal of cell biology*, 176(5):719–27, feb 2007.

[74] Roland Stocker and John F Keaney. Role of oxidative modifications in atherosclerosis. *Physiological reviews*, 84(4):1381–478, oct 2004.

[75] P Barter. The role of HDL-cholesterol in preventing atherosclerotic disease. *European Heart Journal Supplements*, 7(Suppl F):F4–F8, jul 2005.

[76] S. M. Razavian, V. Atger, P. Giral, M. Cambillau, M. Del-Pino, A. C. Simon, N. Moatti, and J. Levenson. Influence of HDL subfractions on erythrocyte aggregation in hypercholesterolemic men. PCVMETRA Group. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 14(3):361–366, mar 1994.

[77] Michael J Joyner and Daniel J Green. Exercise protects the cardiovascular system: effects beyond traditional risk factors. *The Journal of physiology*, 587(Pt 23):5551–8, dec 2009.

[78] T J Bowker, P N Watton, P E Summers, J V Byrne, and Y Ventikos. Rest versus exercise hemodynamics for middle cerebral artery aneurysms: a computational study. *AJNR. American journal of neuroradiology*, 31(2):317–23, feb 2010.

[79] Jaume Padilla, Grant H Simmons, Shawn B Bender, Arturo A Arce-Esquivel, Jeffrey J Whyte, and M Harold Laughlin. Vascular effects of exercise: endothelial adaptations beyond active muscle beds. *Physiology (Bethesda, Md.)*, 26(3):132–45, jun 2011.

[80] Shali Chen, Suranjana Mukherjee, Chandan Chakraborty, and Subrata Chakrabarti. High glucose-induced, endothelin-dependent fibronectin synthesis is mediated via NF-kappa B and AP-1. *American journal of physiology. Cell physiology*, 284(2):C263–72, feb 2003.

[81] Harkiran Kaur, Shali Chen, Xiping Xin, Jane Chiu, Zia A Khan, and Subrata Chakrabarti. Diabetes-induced extracellular matrix protein expression is mediated by transcription coactivator p300. *Diabetes*, 55(11):3104–11, nov 2006.

[82] Jeffrey W Stephens, Manish P Khanolkar, and Stephen C Bain. The biological relevance and measurement of plasma markers of oxidative stress in diabetes and cardiovascular disease. *Atherosclerosis*, 202(2):321–9, feb 2009.

[83] A. R. Pries, T. W. Secomb, T. Gessner, M. B. Sperandio, J. F. Gross, and P. Gaehtgens. Resistance to blood flow in microvessels in vivo. *Circ. Res.*, 75(5):904–915, nov 1994.

[84] A. R. Pries, T. W. Secomb, H. Jacobs, M. Sperandio, K. Osterloh, and P. Gaehtgens. Microvascular blood flow resistance: role of endothelial surface layer. *American journal of Physiology. Heart and Circulatory Physiology*, 273(5):H2272–2279, nov 1997.

[85] C. Desjardins and B. R. Duling. Heparinase treatment suggests a role for the endothelial cell glycocalyx in regulation of capillary hematocrit. *Am J Physiol Heart Circ Physiol*, 258(3):H647–654, 1990.

[86] J C Lewis, R G Taylor, N D Jones, R W St Clair, and J F Cornhill. Endothelial surface characteristics in pigeon coronary artery atherosclerosis. I. Cellular alterations during the initial stages of dietary cholesterol challenge. *Laboratory Investigation; A Journal of Technical Methods and Pathology*, 46(2):123–38, feb 1982.

[87] H Vink and B R Duling. Identification of distinct luminal domains for macromolecules, erythrocytes, and leukocytes within mammalian capillaries. *Circulation research*, 79(3):581–9, sep 1996.

[88] J M Tarbell and M Y Pahakis. Mechanotransduction and the glycocalyx. *Journal of Internal Medicine*, 259(4):339–350, 2006.

[89] Ye Zeng, Eno E Ebong, Bingmei M Fu, and John M Tarbell. The structural stability of the endothelial glycocalyx after enzymatic removal of glycosaminoglycans. *PloS one*, 7(8):e43168, jan 2012.

[90] E. R. Damiano, D. S. Long, and M. L. Smith. Estimation of viscosity profiles using velocimetry data from parallel flows of linearly viscous fluids: application to microvascular haemodynamics. *Journal of Fluid Mechanics*, 512:1–19, jul 2004.

[91] M D Savery and E R Damiano. The endothelial glycocalyx is hydrodynamically relevant in arterioles throughout the cardiac cycle. *Biophysical Journal*, 95(3):1439–47, 2008.

[92] S. Weinbaum, X. Zhang, Y. Han, H. Vink, and S. C. Cowin. Mechanotransduction and flow across the endothelial glycocalyx. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 100(13):7988–7995, 2003.

[93] J M Squire, M Chew, G Nneji, C Neal, J Barry, and C Michel. Quasi-periodic substructure in the microvessel endothelial glycocalyx: a possible explanation for molecular filtering? *Journal of structural biology*, 136(3):239–55, dec 2001.

[94] K P Arkill, C Knupp, C C Michel, C R Neal, K Qvortrup, J Rostgaard, and J M Squire. Similar endothelial glycocalyx structures in microvessels from a range of mammalian tissues: evidence for a common filtering mechanism? *Biophysical journal*, 101(5):1046–56, sep 2011.

[95] Ke Bai and Wen Wang. Shear stress-induced redistribution of the glycocalyx on endothelial cells in vitro. *Biomechanics and modeling in mechanobiology*, 13(2):303–11, apr 2014.

[96] Ye Zeng and John M Tarbell. The adaptive remodeling of endothelial glycocalyx in response to fluid shear stress. *PloS one*, 9(1):e86249, jan 2014.

[97] H Vink, B R Duling, and J A E Spaan. Mechanical properties of the endothelial surface layer. In *Faseb Journal*, volume 13. Federation American Society Exp Biology, 1999.

[98] Y. Han, S. Weinbaum, J. A. E. Spaan, and H. Vink. Large-deformation analysis of the elastic recoil of fibre layers in a Brinkman medium with application to the endothelial glycocalyx. *Journal of Fluid Mechanics*, 554(-1):217, apr 2006.

[99] Ke Bai and Wen Wang. Spatio-temporal development of the endothelial glycocalyx layer and its mechanical property in vitro. *Journal of the Royal Society, Interface / the Royal Society*, pages rsif.2011.0901–, mar 2012.

[100] A R Pries, T W Secomb, and P Gaehtgens. The endothelial surface layer. *Review Literature And Arts of The Americas*, pages 653–666, 2000.

[101] T W Secomb, R Hsu, and A R Pries. Effect of the endothelial surface layer on transmission of fluid shear stress to endothelial cells. *Biorheology*, 38(2-3):143–50, jan 2001.

[102] H. Vink and B. R. Duling. Capillary endothelial surface layer selectively reduces plasma solute distribution volume. *Am J Physiol Heart Circ Physiol*, 278(1):H285–289, 2000.

[103] H. Vink, T. M. Stace, and E. R. Damiano. High resolution 3D intravital fluorescence microscopy reveals partial exclusion of anionic tracers within a 1 micron thick capillary endothelial cell glycocalyx. *Faseb Journal*, 17(4), jun 2003.

[104] E R Damiano and T M Stace. An Electrochemical Model of the Transport of Charged Molecules through the Capillary Glycocalyx. *Biophysical Journal*, 80(4):1169–1670, 2001.

[105] E. R. Damiano. A Mechano-Electrochemical Model of Radial Deformation of the Capillary Glycocalyx. *Biophysical Journal*, 82(3):1153–1175, mar 2002.

[106] J R Levick. Revision of the Starling principle: new views of tissue fluid balance. *The Journal of Physiology*, 557(Pt 3):704, jun 2004.

[107] R H Adamson, J F Lenz, X Zhang, G N Adamson, S Weinbaum, and F E Curry. Oncotic pressures opposing filtration across non-fenestrated rat microvessels. *The Journal of physiology*, 557(Pt 3):889–907, jun 2004.

[108] X Hu and S Weinbaum. A new view of Starling's hypothesis at the microstructural level. *Microvascular research*, 58(3):281–304, nov 1999.

[109] C Charles Michel. Fluid Exchange in the Microcirculation. *The Journal of Physiology*, 557(Pt 3):701–702, 2004.

[110] a W Mulivor and H H Lipowsky. Role of glycocalyx in leukocyte-endothelial cell adhesion. *American Journal of Physiology. Heart and Circulatory Physiology*, 283(4):H1282–91, oct 2002.

[111] Daniel Chappell, Bernhard Heindl, Matthias Jacob, Thorsten Annecke, Congcong Chen, Markus Rehm, Peter Conzen, and Bernhard F Becker. Sevoflurane reduces leukocyte and platelet adhesion after ischemia-reperfusion by protecting the endothelial glycocalyx. *Anesthesiology*, 115(3):483–91, sep 2011.

[112] J-L Balligand, O Feron, and C Dessy. eNOS activation by physical forces: from short-term regulation of contraction to chronic remodeling of cardiovascular tissues. *Physiological reviews*, 89(2):481–534, apr 2009.

[113] R H Adamson and G Clough. Plasma proteins modify the endothelial cell glycocalyx of frog mesenteric microvessels. *The Journal of Physiology*, 445:473–86, jan 1992.

[114] J. A. Florian, J. R Kosky, K. Ainslie, Z. Pang, R. O Dull, and J. M Tarbell. Heparan sulfate proteoglycan is a mechanosensor on endothelial cells. *Circulation Research*, 93(10):e136–42, nov 2003.

[115] M Hecker, A Mülsch, E Bassenge, and R Busse. Vasoconstriction and increased flow: two principal mechanisms of shear stress-dependent endothelial autacoid release. *The American Journal of Physiology*, 265(3 Pt 2):H828–33, sep 1993.

[116] S. Mochizuki, H. Vink, O. Hiramatsu, T. Kajita, F. Shigeto, J. A E Spaan, and F. Kajiya. Role of hyaluronic acid glycosaminoglycans in shear-induced endothelium-derived nitric oxide release. *American Journal of Physiology. Heart and Circulatory Physiology*, 285(2):H722–6, aug 2003.

[117] Wanyi Yen, Bin Cai, Jinlin Yang, Lin Zhang, Min Zeng, John M Tarbell, and Bingmei M Fu. Endothelial surface glycocalyx can regulate flow-induced nitric oxide production in microvessels in vivo. *PloS one*, 10(1):e0117133, jan 2015.

[118] Soo-Hyun Kim, Jeremy Turnbull, and Scott Guimond. Extracellular matrix and cell signalling: the dynamic cooperation of integrin, proteoglycan and growth factor receptor. *The Journal of endocrinology*, 209(2):139–51, may 2011.

[119] T Ariga, T Miyatake, and R K Yu. Role of proteoglycans and glycosaminoglycans in the pathogenesis of Alzheimer's disease and related disorders: amyloidogenesis and therapeutic strategies–a review. *Journal of neuroscience research*, 88(11):2303–15, aug 2010.

[120] C Bertozzi and D Rabuka. Structural Basis of Glycan Diversity. In A Varki, RD Cummings, and JD Esko, editors, *Essentials of Glycobiology*, chapter 2. Cold Spring Harbor Laboratory Press, New York, 2nd edition, 2009.

[121] John T Gallagher and Andrew Walker. Molecular distinctions between heparan sulphate and heparin. *Biochemical Journal*, 230:665–674, 1985.

[122] George A. Jeffrey and John H. Yates. Stereographic representation of the cremer-pople ring-puckering parameters for pyranoid rings. *Carbohydrate Research*, 74(1):319–322, sep 1979.

[123] Herbert L. Strauss and Herbert M. Pickett. Conformational structure, energy, and inversion rates of cyclohexane and some related oxanes. *Journal of the American Chemical Society*, 92(25):7281–7290, dec 1970.

[124] D. Cremer and J. A. Pople. General definition of ring puckering coordinates. *Journal of the American Chemical Society*, 97(6):1354–1358, mar 1975.

[125] Jeffrey D Esko, Koji Kimata, and Ulf Lindahl. Proteoglycans and Sulfated Glycosaminoglycans. In A Varki, RD Cummings, and JD Esko, editors, *Essentials of Glycobiology*, chapter 16. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2nd edition, 2009.

[126] H G Bazin, I Capila, and R J Linhardt. Conformational study of synthetic Δ4-uronate monosaccharides and glycosaminoglycan-derived disaccharides. *Carbohydrate Research*, 309(2):135–144, may 1998.

[127] Lan Jin, Milos Hricovíni, Jon A Deakin, Malcolm Lyon, and Dusan Uhrín. Residual dipolar coupling investigation of a heparin tetrasaccharide confirms the limited effect of flexibility of the iduronic acid on the molecular shape of heparin. *Glycobiology*, 19(11):1185–96, nov 2009.

[128] D Mikhailov, R Linhardt, and K Mayo. NMR solution conformation of heparin-derived hexasaccharide. *Biochemical Journal*, 328:51–61, nov 1997.

[129] D R Ferro, A Provasoli, M Ragazzi, B Casu, G Torri, V Bossennec, B Perly, P Sinaÿ, M Petitou, and J Choay. Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences. *Carbohydrate research*, 195(2):157–67, jan 1990.

[130] V Blanchard, F Chevalier, A Imberty, B R Leeflang, Basappa, K Sugahara, and Johannis P Kamerling. Conformational studies on five octasaccharides isolated from chondroitin sulfate using NMR spectroscopy and molecular modeling. *Biochemistry*, 46(5):1167–75, feb 2007.

[131] B M Sattelle, J Shakeri, I S Roberts, and A Almond. A 3D-structural model of unsulfated chondroitin from high-field NMR: 4-sulfation has little effect on backbone conformation. *Carbohydrate research*, 345(2):291–302, jan 2010.

[132] B Mulloy, M J Forster, C Jones, A F Drake, E A Johnson, and D B Davies. The effect of variation of substitution on the solution conformation of heparin: a spectroscopic and molecular modelling study. *Carbohydrate research*, 255:1–26, mar 1994.

[133] J T Gallagher and A Walker. Molecular distinctions between heparan sulphate and heparin. Analysis of sulphation patterns indicates that heparan sulphate and heparin are separate families of N-sulphated polysaccharides. *The Biochemical journal*, 230(3):665–74, sep 1985.

[134] Daniel R Potter and Edward R Damiano. The hydrodynamically relevant endothelial cell glycocalyx observed in vivo is absent in vitro. *Circulation research*, 102(7):770–6, apr 2008.

[135] T Arisaka, M Mitsumata, M Kawasumi, T Tohjima, S Hirose, and Y Yoshida. Effects of shear stress on glycosaminoglycan synthesis in vascular endothelial cells. *Annals of the New York Academy of Sciences*, 748:543–54, jan 1995.

[136] Mirella Gouverneur, Jos A E Spaan, Hans Pannekoek, Ruud D Fontijn, and Hans Vink. Fluid shear stress stimulates incorporation of hyaluronan into endothelial cell glycocalyx. *American Journal of Physiology. Heart and Circulatory Physiology*, 290(1):H458–2, jan 2006.

[137] Bernard M van den Berg, Jos A E Spaan, Titia M Rolf, and Hans Vink. Atherogenic region and diet diminish glycocalyx dimension and increase intima-to-media ratios at murine carotid artery bifurcation. *American journal of physiology. Heart and circulatory physiology*, 290(2):H915–20, feb 2006.

[138] Kristina M. Giantsos-Adams, Andrew Jia-An Koo, Sukhyun Song, Jiro Sakai, Jagadish Sankaran, Jennifer H. Shin, Guillermo Garcia-Cardena, and C. Forbes Dewey. Heparan Sulfate Regrowth Profiles Under Laminar Shear Flow Following Enzymatic Degradation. *Cellular and Molecular Bioengineering*, feb 2013.

[139] Johan Kreuger, Dorothe Spillmann, Jin-ping Li, and Ulf Lindahl. Interactions between heparan sulfate and proteins: the concept of specificity. *The Journal of cell biology*, 174(3):323–7, jul 2006.

[140] M Bernfield, M Götte, P W Park, O Reizes, M L Fitzgerald, J Lincecum, and M Zako. Functions of cell surface heparan sulfate proteoglycans. *Annual review of biochemistry*, 68:729–77, jan 1999.

[141] Cristal I Gama, Sarah E Tully, Naoki Sotogaku, Peter M Clark, Manish Rawat, Nagarajan Vaidehi, William A Goddard, Akinori Nishi, and Linda C Hsieh-Wilson. Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nature chemical biology*, 2(9):467–73, sep 2006.

[142] Constant A. A. van Boeckel and Maurice Petitou. The Unique Antithrombin III Binding Domain of Heparin: A Lead to New Synthetic Antithrombotics. *Angewandte Chemie International Edition in English*, 32(12):1671–1690, dec 1993.

[143] Hongyan Kang, Yubo Fan, Anqiang Sun, and Xiaoyan Deng. Compositional or charge density modification of the endothelial glycocalyx accelerates flow-dependent concentration polarization of low-density lipoproteins. *Experimental biology and medicine (Maywood, N.J.)*, page ebm.2011.011072v1, jun 2011.

[144] J A Bevan and G Siegel. Blood vessel wall matrix flow sensor: evidence and speculation. *Blood vessels*, 28(6):552–6, jan 1991.

[145] G Siegel, A Walter, A Kauschmann, M Malmsten, and E Buddecke. Anionic biopolymers as blood flow sensors. *Biosensors & Bioelectronics*, 11(3):281–94, jan 1996.

[146] Masako Sugihara-Seki, Takeshi Akinaga, and Tomoaki Itano. Flow across microvessel walls through the endothelial surface glycocalyx and the interendothelial cleft, apr 2008.

[147] Bingmei M Fu, Bin Chen, and Wenhao Chen. An electrodiffusion model for effects of surface glycocalyx layer on microvessel permeability. *American journal of physiology. Heart and circulatory physiology*, 284(4):H1240–50, apr 2003.

[148] N J Agrawal and R Radhakrishnan. The Role of Glycocalyx in Nanocarrier-Cell Adhesion Investigated Using a Thermodynamic Model and Monte Carlo Simulations. *The journal of physical chemistry. C, Nanomaterials and interfaces*, 111(43):15848–15856, nov 2007.

[149] P P Sumets, J E Cater, D S Long, and R J Clarke. A boundary-integral representation for biphasic mixture theory, with application to the post-capillary glycocalyx. *Proceedings. Mathematical, physical, and engineering sciences / the Royal Society*, 471(2179):20140955, jul 2015.

[150] John M Tarbell and Zhong-Dong Shi. Effect of the glycocalyx layer on transmission of interstitial flow shear stress to embedded cells. *Biomechanics and modeling in mechanobiology*, 12(1):111–21, jan 2013.

[151] Mahsa Dabagh, Payman Jalali, Peter J Butler, and John M Tarbell. Shear-induced force transmission in a multicomponent, multicell model of the endothelium. *Journal of the Royal Society, Interface / the Royal Society*, 11(98):20140431, sep 2014.

[152] Jacob Klein. Chemistry. Repair or replacement–a joint perspective. *Science (New York, N.Y.)*, 323(5910):47–8, jan 2009.

[153] Angus Hucknall, Andrew J Simnick, Ryan T Hill, Ashutosh Chilkoti, Andres Garcia, Matthew S Johannes, Robert L Clark, Stefan Zauscher, and Buddy D Ratner. Versatile synthesis and micropatterning of nonfouling polymer brushes on the wafer scale. *Biointerphases*, 4(2):FA50–7, jun 2009.

[154] K Binder, T Kreer, and A Milchev. Polymer brushes under flow and in other out-of-equilibrium conditions. *Soft Matter*, 7(16):7159, aug 2011.

[155] Luca Lanotte, Giovanna Tomaiuolo, Chaouqi Misbah, Lionel Bureau, and Stefano Guido. Red blood cell dynamics in polymer brush-coated microcapillaries: A model of endothelial glycocalyx in vitro. *Biomicrofluidics*, 8(1):014104, jan 2014.

[156] Mingge Deng, Xuejin Li, Haojun Liang, Bruce Caswell, and George Em Karniadakis. Simulation and modelling of slip flow over surfaces grafted with polymer brushes and glycocalyx fibres. *Journal of fluid mechanics*, 711, nov 2012.

[157] M. Eriksson, T. K Lindhorst, and B. Hartke. Differential effects of oligosaccharides on the hydration of simple cations. *The Journal of Chemical Physics*, 128(10):105105, mar 2008.

[158] B M Sattelle, J Shakeri, M J Cliff, and A Almond. Proteoglycans and Their Heterogeneous Glycosaminoglycans at the Atomic Scale. *Biomacromolecules*, 16(3):951–961, 2015.

[159] E R Cruz-Chu, A Malafeev, T Pajarskas, I V Pivkin, and P Koumoutsakos. Structure and response to flow of the glycocalyx layer. *Biophysical Journal*, 106(1):232–43, jan 2014.

[160] M. Cieplak, J. Koplik, and J. Banavar. Boundary Conditions at a Fluid-Solid Interface. *Physical Review Letters*, 86(5):803–806, jan 2001.

[161] J Shiomi and S Maruyama. Water transport inside a single-walled carbon nanotube driven by a temperature gradient. *Nanotechnology*, 20(5):055708, feb 2009.

[162] J. M. D. MacElroy. Nonequilibrium molecular dynamics simulation of diffusion and flow in thin microporous membranes. *The Journal of Chemical Physics*, 101(6):5274, sep 1994.

[163] R. Cracknell, D. Nicholson, and N. Quirke. Direct Molecular Dynamics Simulation of Flow Down a Chemical Potential Gradient in a Slit-Shaped Micropore. *Physical Review Letters*, 74(13):2463–2466, mar 1995.

[164] J. Koplik, J. R Banavar, and J. F Willemsen. Letters 28. *Physical Review*, 60(13):1282–1286, 1988.

[165] J. Koplik, J. R. Banavar, and J. F. Willemsen. Molecular dynamics of fluid flow at solid surfaces. *Physics of Fluids A: Fluid Dynamics*, 1(5):781, 1989.

[166] V. P. Sokhan, D. Nicholson, and N. Quirke. Fluid flow in nanopores: Accurate boundary conditions for carbon nanotubes. *The Journal of Chemical Physics*, 117(18):8531, nov 2002.

[167] V. P. Sokhan, D. Nicholson, and N. Quirke. Fluid flow in nanopores: An examination of hydrodynamic boundary conditions. *The Journal of Chemical Physics*, 115(8):3878, aug 2001.

[168] EM Kotsalis, J.H. Walther, and P. Koumoutsakos. Multiphase water flow inside carbon nanotubes. *International Journal of Multiphase Flow*, 30(7-8):995–1010, aug 2004.

[169] G Nagayama. Effects of interface wettability on microscale flow by molecular dynamics simulation. *International Journal of Heat and Mass Transfer*, 47(3):501–513, jan 2004.

[170] Quy Dong To, Thanh Tung Pham, Guy Lauriat, and Céline Léonard. Molecular Dynamics Simulations of Pressure-Driven Flows and Comparison with Acceleration-Driven Flows. *Advances in Mechanical Engineering*, 2012:1–10, 2012.

[171] I. Hanasaki, J. H Walther, S. Kawano, and P. Koumoutsakos. Coarse-grained molecular dynamics simulations of shear-induced instabilities of lipid bilayer membranes in water. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 82(5 Pt 1):051602, nov 2010.

[172] T. Tanzeglock, M. Soos, G. Stephanopoulos, and M. Morbidelli. Induction of mammalian cell death by simple shear and extensional flows. *Biotechnology and Bioengineering*, 104(2):360–70, oct 2009.

[173] Zhongzhou Chen, Jizhong Lou, Cheng Zhu, and Klaus Schulten. Flow-induced structural transition in the beta-switch region of glycoprotein Ib. *Biophysical Journal*, 95(3):1303–13, aug 2008.

[174] Xueqing Zou, Yanxin Liu, Zhongzhou Chen, Gloria Ines Cárdenas-Jirón, and Klaus Schulten. Flow-induced beta-hairpin folding of the glycoprotein Ibalpha beta-switch. *Biophysical Journal*, 99(4):1182–91, aug 2010.

[175] D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press, 2nd edition, 2002.

[176] D. J. Gee and M. R. King, editors. *Multiscale Modeling of Particle Interactions: Applications in Biology and Nanotechnology*. Wiley, New Jersey, 2010.

[177] R. J Woods and M. B Tessier. Computational glycoscience: characterizing the spatial and temporal properties of glycans and glycan-protein complexes. *Current Opinion in Structural Biology*, 20(5):575–83, oct 2010.

[178] Atoms In Motion - http://atomsinmotion.com/, 2014.

[179] J. E. Jones. On the Determination of Molecular Fields. II. From the Equation of State of a Gas. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 106(738):463–477, oct 1924.

[180] J E Lennard-Jones. Cohesion. *Proceedings of the Physical Society*, 43(5):461–482, sep 1931.

[181] H J C Berendsen, J P M Postma, and W F van Gunsteren. Interaction models for water in relation to protein hydration. *Intermolecular Forces*, pages 331–342, 1981.

[182] W L Jorgensen, J Chandrasekhar, J D Madura, R W Impey, and M L Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926, jul 1983.

[183] W L Jorgensen. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water. *J. Am. Chem. Soc.;(United States)*, 103(2):335–340, 1981.

[184] M. W. Mahoney and W. L. Jorgensen. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *The Journal of Chemical Physics*, 112(20):8910, may 2000.

[185] D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, jan 2004.

[186] R. Fletcher. Function minimization by conjugate gradients. *The Computer Journal*, 7(2):149–154, feb 1964.

[187] B R Brooks, R E Bruccoleri, B D Olafson, D J States, S Swaminathan, and M Karplus. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, 4(2):187–217, 1983.

[188] Loup Verlet. Computer" experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical Review*, 159(1), 1967.

[189] M P Allen and D J Tildesley. *Computer Simulation of Liquids*. Clarendon Press, 1989.

[190] Paul Peter Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik*, 369(3):253–287, 1921.

[191] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089, jun 1993.

[192] Matteo Frigo and Steven G Johnson. The Fastest Fourier Transform in the West (MIT-LCS-TR-728). *Materials Research*, page 20, 1997.

[193] J C Phillips, R Braun, W Wang, J Gumbart, E Tajkhorshid, E Villa, C Chipot, R D Skeel, L Kalé, and K Schulten. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–802, dec 2005.

[194] B R Brooks, C L Brooks, A D Mackerell, L Nilsson, R J Petrella, B Roux, Y Won, G Archontis, C Bartels, S Boresch, A Caflisch, L Caves, Q Cui, A R Dinner, M Feig, S Fischer, J Gao, M Hodoscek, W Im, K Kuczera, T Lazaridis, J Ma, V Ovchinnikov, E Paci, R W Pastor, C B Post, J Z Pu, M Schaefer, B Tidor, R M Venable, H L Woodcock, X Wu, W Yang, D M York, and M Karplus. CHARMM: the biomolecular simulation program. *Journal of Computational Chemistry*, 30(10):1545–614, jul 2009.

[195] J. Wang, P. Cieplak, and P. A. Kollman. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry*, 21(12):1049, sep 2000.

[196] K. N Kirschner, A. B Yongye, S. M Tschampel, J. Gonza, C. R Daniels, B L. Foley, and R. J Woods. GLYCAM06 : A Generalizable Biomolecular Force Field. Carbohydrates. *Journal of Computational Chemistry*, 2007.

[197] R. D Lins and P.H Hünenberger. A new GROMOS force field for hexopyranose-based carbohydrates. *Journal of Computational Chemistry*, 26(13):1400–12, oct 2005.

[198] J.C. Phillips, Gengbin Zheng, S. Kumar, and L.V. Kale. NAMD: Biomolecular Simulation on Thousands of Processors. In *Supercomputing, ACM/IEEE 2002 Conference*, page 36, Baltimore, 2002. IEEE.

[199] Ulrich Essmann, Lalith Perera, Max L. Berkowitz, Tom Darden, Hsing Lee, and Lee G. Pedersen. A smooth particle mesh Ewald method. *The Journal of Chemical Physics*, 103(19):8577, nov 1995.

[200] Nick Schafer, Radu Serban, and Dan Negrut. An Investigation on New Numerical Methods for Molecular Dynamics Simulation. In *Volume 5: 6th International Conference on Multibody Systems, Nonlinear Dynamics, and Control, Parts A, B, and C*, pages 1467–1471. ASME, 2007.

[201] C. Schafmeister, D. E. Smyth, V. Romanovski, B. Ross, S. Brozell, and W. Zhang. Private communication, 2011.

[202] T. Cheatham, M. Crowley, V. Tsui, J. Pitera, D. Case, H. Gohlke, S. Tanner, E. Absgarten, D. Roe, P. Frybarger, and R. Walker. Private Communication, 2011.

[203] William Humphrey, Andrew Dalke, and Klaus Schulten. {VMD} – {V}isual {M}olecular {D}ynamics. *Journal of Molecular Graphics*, 14:33–38, 1996.

[204] Woods Group. Carbohydrate Builder (http://www.glycam.com). In *GLYCAM Web Complex Carbohydrate Research Center, University of Georgia, Athens, GA.*, 2005.

[205] B. Widom. *Statistical mechanics: a concise introduction for chemists*. Cambridge University Press, 2002.

[206] Sameer Varma and Susan B. Rempe. Coordination numbers of alkali metal ions in aqueous solutions. *Biophysical Chemistry*, 124(3):192–199, 2006.

[207] G N Ramachandran, C Ramakrishnan, and V Sasisekharan. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*, 7:95–9, jul 1963.

[208] A. M Salisburg, A. L Deline, K. W Lexa, G. C Shields, and K.N Kirschner. Ramachandran-Type Plots for Glycosidic Linkages : Examples from Molecular Dynamic Simulations Using the Glycam06 Force Field. *Journal of Computational Chemistry*, pages 1–12, 2008.

[209] S. C Lovell, I. W Davis, W B Arendall, P. I W de Bakker, J M Word, M G Prisant, J S Richardson, and D C Richardson. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins*, 50(3):437–50, feb 2003.

[210] G N Ramachandran and V Sasisekharan. Conformation of polypeptides and proteins. *Advances in Protein Chemistry*, 23:283–438, jan 1968.

[211] J R Brisson and J P Carver. Solution conformation of .alpha.-D-(1-3)- and .alpha.-D-(1-6)-linked oligomannosides using proton nuclear magnetic resonance. *Biochemistry*, 22(6):1362–1368, mar 1983.

[212] L. Poppe, G. S. Brown, J.S. Philo, P. V. Nikrad, and B. H. Shah. Conformation of sLe x Tetrasaccharide, Free in Solution and Bound to E-, P-, and L-Selectin. *Journal of the American Chemical Society*, 119(7):1727–1736, feb 1997.

[213] X. Hu, I. Carmichael, and A. S Serianni. N-acetyl side-chains in saccharides: NMR J-coupling equations sensitive to CH-NH and NH-CO bond conformations in 2-acetamido-2-deoxy-aldohexopyranosyl rings. *The Journal of Organic Chemistry*, 75(15):4899–910, aug 2010.

[214] Mari L DeMarco and Robert J Woods. Structural glycobiology: a game of snakes and ladders. *Glycobiology*, 18(6):426–40, jun 2008.

[215] B. Lachele Foley, Matthew B. Tessier, and Robert J. Woods. Carbohydrate force fields. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(4):652–697, jul 2012.

[216] John G. Kirkwood. Statistical mechanics of fluid mixtures. *Journal Of Chemical Physics*, 3(May 2015):300–313, 1935.

[217] Eric Darve and Andrew Pohorille. Calculating free energies using average force. *Journal of Chemical Physics*, 115(20):9169–9183, 2001.

[218] Cristian Micheletti, Alessandro Laio, and Michele Parrinello. Reconstructing the density of states by history-dependent metadynamics. *Physical Review Letters*, 92(17):170601–1, 2004.

[219] C Bartels and M Karplus. Multidimensional adaptive umbrella sampling: Applications to main chain and side chain peptide conformations. *Journal of Computational Chemistry*, 18(12):1450–1462, 1997.

[220] Eric Darve, David Rodríguez-Gómez, and Andrew Pohorille. Adaptive biasing force method for scalar and vector free energy calculations. *The Journal of chemical physics*, 128(14):144120, apr 2008.

[221] Martin Karplus and J Andrew McCammon. Molecular dynamics simulations of biomolecules. *Nature Structural Biology*, 9(9):646–52, sep 2002.

[222] Jeremy E Turnbull and Robert A Field. Emerging glycomics technologies. *Nature Chemical Biology*, 3(2):74–7, feb 2007.

[223] H J C Berendsen, J P M Postma, W F van Gunsteren, A DiNola, and J R Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684, 1984.

[224] Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman J.C Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23(3):327–341, mar 1977.

[225] D A Case, T A Darden, T E Cheatham, C L Simmerling, J Wang, R E Duke, R Luo, R C Walker, W Zhang, K M Merz, B Roberts, S Hayik, A Roitberg, G Seabra, J Swails, A W Goetz, I Kolossváry, K F Wong, F Paesani, J Vanicek, R M Wolf, J Liu, X Wu, S R Ferrer, C Sagui, V Babin, T Luchko, S Gusarov, A Kovalenko, and P A Kollman. AMBER 12 and AmberTools 13, 2012.

[226] L.D. Landau and E.M. Lifshitz. *Statistical Physics, Third Edition, Part 1: Volume 5*. Butterworth-Heinemann; 3 edition, 1975.

[227] R J Woods, A Pathiaseril, M R Wormald, C J Edge, and R A Dwek. The high degree of internal flexibility observed for an oligomannose oligosaccharide does not alter the overall topology of the molecule. *European journal of biochemistry / FEBS*, 258(2):372–86, dec 1998.

[228] R J Woods, R A Dwek, C J Edge, and B Fraser-Reid. Molecular Mechanical and Molecular Dynamic Simulations of Glycoproteins and Oligosaccharides. 1. GLYCAM_93 Parameter Development. *The Journal of Physical Chemistry*, 99(11):3832–3846, mar 1995.

[229] D Beglov and B Roux. Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations. *The Journal of Chemical Physics*, 100(12):9050, jun 1994.

[230] Roland Bammer, Thomas a Hope, Murat Aksoy, and Marcus T Alley. Time-resolved 3D quantitative flow MRI of the major intracranial vessels: initial experience and comparative evaluation at 1.5T and 3.0T in combination with parallel imaging. *Magnetic resonance in medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 57(1):127–40, jan 2007.

[231] Lars Nybo and Bodil Nielsen. Middle cerebral artery blood velocity is reduced with hyperthermia during prolonged exercise in humans. *The Journal of Physiology*, 534(1):279–286, jul 2001.

[232] G. Chalmers, J.N. Glushka, B.L. Foley, R.J. Woods, and J.H. Prestegard. Direct NOE Simulation from Long MD Trajectories. *Journal of Magnetic Resonance*, 265:1–9, jan 2016.

[233] Andrei-J Petrescu, Adina-L Milac, Stefana M Petrescu, Raymond A Dwek, and Mark R Wormald. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology*, 14(2):103–14, mar 2004.

[234] Arunima Singh, Matthew B Tessier, Kari Pederson, Xiaocong Wang, Andre P Venot, Geert-Jan Boons, James H Prestegard, and Robert J Woods. Extension and Validation of the GLYCAM Force Field Parameters for modeling Glycosaminoglycans. *Canadian Journal of Chemistry*, pages cjc–2015–0606, feb 2016.

[235] Glenn J. Martyna, Douglas J. Tobias, and Michael L. Klein. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics*, 101(5):4177, 1994.

[236] Scott E. Feller, Yuhong Zhang, Richard W. Pastor, and Bernard R. Brooks. Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics*, 103(11):4613, sep 1995.

[237] M. B Tessier. *The Development of Computational Carbohydrate Grafting and the GLYCAM Force Field to Understand how Glycan Structure Alters Biorecognition.* PhD thesis, University of Georgia, 2013.

[238] Max Nieuwdorp, Marijn C Meuwese, Hans Vink, Joost B L Hoekstra, John J P Kastelein, and Erik S G Stroes. The endothelial glycocalyx: a potential barrier between health and vascular disease. *Current Opinion in Lipidology*, 16(5):507–11, 2005.

[239] B M Sattelle and A Almond. Shaping up for structural glycomics: a predictive protocol for oligosaccharide conformational analysis applied to N-linked glycans. *Carbohydrate research*, 383:34–42, jan 2014.