UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Dissertation

submitted to the
Combined Faculty of Mathematics, Engineering and Natural Sciences
of the Ruperto Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Saber Boushehri
Born in Iran
May 2024

Effect of O-glycans on the Structure and Viscosity of Intrinsically Disordered and Glycosylated Protein Lubricin

Referees:	Prof. Dr. Frauke Gräter
	Prof. Dr. Lars Pastewka
	Prof. Dr. Rob Russell
	PD. Dr. Heike Böhm

This work is licensed under a Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.

Abstract

Synovial joints exhibit remarkable lubrication mechanisms, essential for the smooth articulation of bones. This intricate process is orchestrated by key biomolecular components, such as hyaluronic acid, lipids, aggrecan, and glycoproteins. Lubricin, characterized by its intrinsically disordered region and highly glycosylated structure, is part of these glycoproteins. Lubricin significantly contributes to reducing viscosity and ensuring the sustained efficacy of synovial joints. The protein comprises two globular end domains and an intrinsically disordered central domain. The end domains bind to cartilage surfaces, allowing the remaining protein to extend into the synovial fluid. The central domain, known as the mucin-like domain, is abundant in O-glycans and mainly responsible for lubricin functionality. The O-glycans constitute approximately 30-35% of the entire lubricin structure. O-glycans play a pivotal role in enabling lubricin to fulfill its highly specialized function. Despite the recognized importance of O-glycans, a detailed understanding of their influence on the structure and viscosity of lubricin remains elusive.

Using extensive molecular dynamics (MD) simulations, in this thesis we unveil the influence of O-glycans on the rheological and viscosity properties of lubricin. In the first part of this thesis, our focus was on understanding how O-glycans influence the structure of lubricin. To achieve this, glycosylated fragments of lubricin were generated and modeled, and a suitable force field for intrinsically disordered glycoproteins was prepared. Five different segments of lubricin, each with a length of 80 amino acids, were considered. Utilizing Monte Carlo sampling, six different glycosylated fragments were introduced for each segment. Equilibrium molecular dynamics simulations were then conducted to explore the role of glycans in the conformational properties of lubricin. The findings reveal that the presence of O-glycans induces a more extended conformation in fragments of the disordered region of lubricin, resulting in a stiffer structure, and enhancing the exposure of lubricin to solvent molecules. These changes in the lubricin structure are attributed to the electrostatic and steric interactions imposed by the bulky side chains of O-glycans.

The second phase of our study focused on unraveling the influence of glycans on the viscoelastic behavior of liquid systems containing lubricin fragments. Specifically, we aimed to understand how the solution viscosity may have been affected by the presence of this O-glycosylated protein. To achieve this, we employed the Green-Kubo method, i.e. we derived the zero-shear viscosity from the fluctuations of the pressure tensor in equilibrium molecular dynamics simulations. In addition we carried shear-driven non-equilibrium molecular dynamics simulations to determine the viscosity under shear. Our simulations reveal that, unlike pure water, systems containing lubricin display a pronounced shear-thinning response. Furthermore, our findings demonstrate that glycosylation and the mass density of lubricin chains play a crucial role in regulating the system viscosity and its response to shear. Increasing mass density leads to higher viscosity, but the presence of O-glycans results in a reduction in solution viscosity and weakens shear thinning at high shear rates, compared to non-glycosylated systems with the same density. The electrostatic and steric interactions of O-glycans prevent the conglomeration and structuring of lubricin fragments, thereby altering the viscoelastic properties of lubricin.

The results from our computational study provide a mechanistic understanding of previous experimental observations of lubricin, offering a more rational comprehension of its function in the synovial fluid.

Zusammenfassung

Synovialgelenke verfügen über bemerkenswerte Schmiermechanismen, die für die reibungslose Bewegung der Knochen unerlässlich sind. Dieser komplizierte Prozess wird von wichtigen biomolekularen Komponenten wie Hyaluronsäure, Lipiden, Aggrecan und Glykoproteinen gesteuert. Lubricin, das sich durch seine intrinsisch ungeordnete Region und seine hochglykosylierte Struktur auszeichnet, gehört zu diesen Glykoproteinen. Lubricin trägt wesentlich dazu bei, die Viskosität zu verringern und die dauerhafte Funktionsfähigkeit der Synovialgelenke zu gewährleisten. Das Protein besteht aus zwei globulären Enddomänen und einer intrinsisch ungeordneten zentralen Domäne. Die Enddomänen binden an Knorpeloberflächen, während das restliche Protein in die Synovialflüssigkeit eindringt. Die zentrale Domäne, die als mucinähnliche Domäne bekannt ist, ist reich an O-Glykanen und hauptsächlich für die Funktionalität des Lubricins verantwortlich. Die O-Glykane machen etwa 30-35% der gesamten Lubricinstruktur aus. O-Glykane spielen eine zentrale Rolle dabei, dass Lubricin seine hochspezialisierte Funktion erfüllen kann. Trotz der bekannten Bedeutung der O-Glykane fehlt ein detailliertes Verständnis ihres Einflusses auf die Struktur und Viskosität von Lubricin.

Mithilfe umfangreicher Molekulardynamiksimulationen (MD) enthüllen wir in dieser Arbeit den Einfluss von O-Glykanen auf die rheologischen und viskosen Eigenschaften von Lubricin. Im ersten Teil dieser Arbeit konzentrierten wir uns darauf zu verstehen, wie O-Glykane die Struktur von Lubricin beeinflussen. Zu diesem Zweck wurden glykosylierte Fragmente von Lubricin erzeugt und modelliert, und es wurde ein geeignetes Kraftfeld für intrinsisch ungeordnete Glykoproteine erstellt. Fünf verschiedene Segmente von Lubricin, jedes mit einer Länge von 80 Aminosäuren, wurden berücksichtigt. Unter Verwendung von Monte-Carlo-Sampling wurden für jedes Segment sechs verschiedene glykosylierte Fragmente erzeugt. Anschließend wurden Gleichgewichts-Molekulardynamiksimulationen durchgeführt, um die Rolle der Glykane für die Konformationseigenschaften von Lubricin zu untersuchen. Die Ergebnisse zeigen, dass das Präsenz von O-Glykanen eine ausgedehntere Konfpormation in Fragmenten der ungeordneten Region des Lubricins bewirkt, was zu einer steiferen Struktur führt und das Lubricin den Lösungsmittelmolekülen erhöht aussetzt. Diese Veränderungen in der Lubricinstruktur werden auf die elektrostatischen und sterischen Wechselwirkungen zurückgeführt, die durch die sperrigen Seitenketten der O-Glykane verursacht werden.

In der zweiten Phase unserer Studie ging es darum, den Einfluss von Glykanen auf das viskoelastische Verhalten von flüssigen Systemen, die Lubricin-Fragmente enthalten, zu entschlüsseln. Insbesondere wollten wir verstehen, wie die Lösungsviskosität durch die Anwesenheit dieses O-glykosylierten Proteins beeinflusst wird. Zu diesem Zweck verwendeten wir die Green-Kubo-Methode, um die Null-Scherviskosität aus den Fluktuationen des Drucktensors in Gleichgewichts-Molekulardynamiksimulationen abzuleiten. Außerdem führten wir scherungsgetriebene Nicht-Gleichgewichts-Molekulardynamiksimulationen durch, um die Viskosität unter Scherung zu bestimmen. Unsere Simulationen zeigen, dass Systeme, die Lubricin enthalten, im Gegensatz zu reinem Wasser eine ausgeprägte scherverdünnende Reaktion zeigen. Darüber hinaus zeigen unsere Ergebnisse, dass die Glykosylierung und die Massendichte der Lubricinketten eine entscheidende Rolle bei der Regulierung der Viskosität des Systems und seiner Reaktion auf Scherung spielen. Eine Erhöhung der Massendichte führt zu einer höheren Viskosität, aber das Vorhandensein von O-Glykanen führt zu einer Verringerung der Lösungsviskosität und schwächt die Scherverdünnung bei hohen Schergeschwindigkeiten im Vergleich zu nicht-glykosylierten Systemen mit der gleichen Massendichte. Die elektrostatischen und sterischen Wechselwirkungen von O-Glykanen verhindern die Zusammenballung und Strukturierung von Lubricinfragmenten und verändern dadurch die viskoelastischen Eigenschaften von Lubricin.

Die Ergebnisse unserer Berechnungsstudie ermöglichen ein mechanistisches Verständnis früherer experimenteller Beobachtungen von Lubricin und bieten ein rationaleres Verständnis seiner Funktion in der Synovialflüssigkeit.

Acknowledgment

I express my deepest gratitude to all those who contributed to the completion of this Ph.D. thesis.

First and foremost, I would like to extend my heartfelt thanks to my supervisor, Prof. Dr. Frauke Gräter, for her unwavering support, guidance, and the opportunity to work in an excellent environment. Her advice, insightful suggestions, and constant encouragement have collectively been instrumental in advancing this Ph.D.

I extend sincere appreciation to my advisor, Dr. Camilo Aponte-Santamaría, for his patient and constant support in my research, guiding the progress of my work through countless helpful discussions, essential to the completion of the thesis.

Special thanks to Prof. Dr. Lars Pastewka, Prof. Dr. Peter Gumbsch, Dr. Hannes Holey, and Dr. Mohamed Tarek Elewa Hassan for their insightful discussions and collaboration.

Gratitude is extended to the Heidelberg Institute for Theoretical Studies, University of Heidelberg, Karlsruhe Institute of Technology (KIT), German Research Foundation (DFG) through the RTG/GRK-2450 graduate school program, the Klaus Tschira Foundation, and the state of Baden-Württemberg through bwHPC for providing a conducive academic environment and access to essential resources for meaningful research. Special appreciation goes to Prof. Dr. Marcus Elstner for his pivotal role in developing the RTG/GRK-2450 graduate school, shaping the course of this academic journey.

Special thanks to the dedicated members of the MBM group, both past and present, including Leif, Marcus, Andrea, Matthias, Nick, Benedikt, Kai, Svenja, Daniel, Debora, Eric, Isabel, Fabian, Florian, Erik, Jannik, Johanna, Aysecan, Nuriza, Nickolas, and Christina. Our discussions and exchange of ideas significantly enriched the content of this thesis. This journey was made more enjoyable with your invaluable contributions and help. Thank you for your collaboration and shared enthusiasm during the research phase.

I express gratitude to all my friends and colleagues who provided moral support, encouragement, and moments of respite during the demanding phases of this academic pursuit.

Last but not least, heartfelt appreciation goes to my family for their unwavering support, understanding, and encouragement. Particularly, my wife Rayehe, for her love, resilience, and countless sacrifices.

Contents

List of Figures

List of Tables

1

Introduction

Viscoelastic fluids are materials which exhibit complex rheological properties. As their name suggests, they are characterized by a combination of viscous and elastic behaviors. These materials exhibit frequency-dependent energy storage and dissipation. When the fluid is subjected to stress, its non-Newtonian properties become evident, manifesting in a viscosity change during external shear flow. For instance, under high strain rates, the microscopic structure of viscoelastic fluids undergoes deformation, leading to nonlinear rheological responses such as shear thinning or shear thickening [1–3]. Viscoelastic materials and viscous fluids play a crucial role in diverse applications, ranging from biological structures to industrial materials. Examples of their application include oil, molecular inks, Nafion inks, cellulose hydrogels, and mucus [4–7].

The synovial fluid is a crucial viscoelastic element within the human body. It is present in joints like the knee, hip, and shoulder. It plays a key role in enabling smooth movements during different activities. The complex rheological characteristics of the synovial fluid arise from the dynamic interplay among its components, which include hyaluronic acid, albumin, phospholipids, and lubricin [8–11]. Previous studies have demonstrated that lubricin, a vital mucinous glycoprotein and intrinsically disordered protein (IDP), plays a crucial role in improving lubrication within synovial joints by efficiently reducing friction [12–14].

In the upcoming sections, we will provide biological insights into synovial joints, focusing on essential components like lubricin and hyaluronic acid (HA). Additionally, we will scrutinize the illustration of some biological properties of proteins, including intrinsically disordered and glycosylated ones.

1.1 Synovial joint

A biological joint is where two or more bones come together in the body. Human bodies have different types of joints, with the synovial or diarthrodial ones (shown in Figure 1.1), like the knee, hip, and shoulder, being the most common type [15, 16]. These joints enable the bones to move freely and facilitate their motion [17]. They can achieve a very low friction coefficient, ranging from 0.0005 to 0.04 (as a comparison that of Teflon is 0.05 to 0.2 [18]) and endure pressures of around 200 atm [19,20]. Healthy joints exhibit effective lubrication and resistance to wear, withstanding over 100 million of shearing and loading cycles. These remarkable lubricating properties and wear resistance are needed for facilitating smooth movement of articulating bones on each other [21], and degradation of these properties can result in osteoarthritis [22, 23].

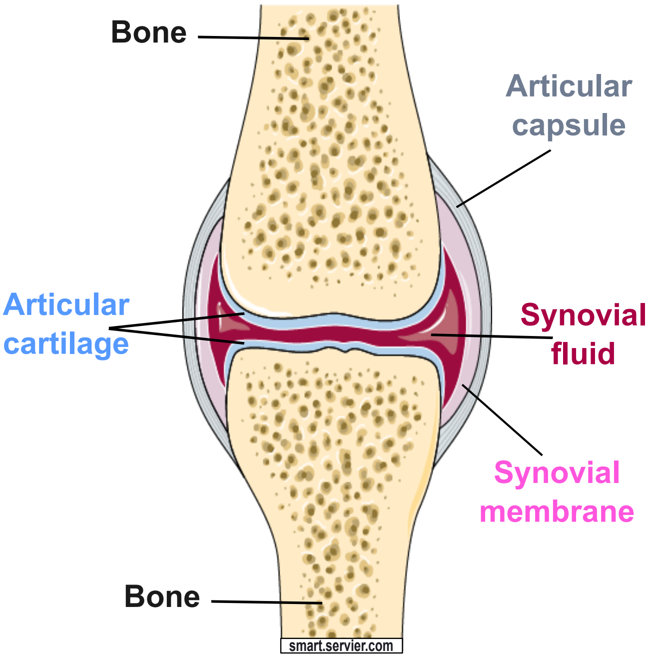

Figure 1.1. A schematic representation of a synovial joint. The interplay between articular cartilage and synovial fluid is crucial, providing exceptional lubrication essential for the smooth articulation of bones.

The synovial joint is a complex structure made up of an articular capsule, two bony surfaces covered with hyaline cartilage, and synovial fluid (Figure 1.1). The articular capsule is a fibrous connective tissue structure and shapes the boundaries of the joint space. The articular cartilage, a thin, elastic, and fibrous connective tissue, covers the entire articulating surface of each bone. The cavity between articular cartilages is filled with a viscous, non-

Newtonian synovial fluid [17,24]. For example, in the human knee joint, the cartilage area is around 120 cm^2 and typically contains 1-2 ml of synovial fluid [25–27]. There are six different types of synovial joints in the human body, distinguished by the types of movements they permit, namely planar, hinge, pivot, condyloid, saddle, or ball-and-socket [28,29].

Various models have been proposed to explain the mechanisms facilitating the attainment of low friction in this context. These include hydrodynamic and elastohydrodynamic lubrication (especially applicable at high rates of motion and large separations between cartilages) [30,31], weeping [32,33], biphasic [34], and boundary lubrication (notably at direct cartilage-cartilage contact) [35–39].

The achievement of effective lubrication in joint mechanisms is intricately linked to the collaborative interactions between articular cartilage and synovial fluid, as well as their key components such as Proteoglycan4 (PRG4), also known as lubricin, and hyaluronic acid [40–42]. In the upcoming sections, we will probe deeper into the intricate details of these biological components, exploring their roles and contributions to the overall lubrication mechanism within joints.

1.1.1 Articulate cartilage

Articular cartilage is a biological and highly specialized connective tissue found in synovial joints. This hyaline cartilage is about 2 to 4 mm thick and has four zones: superficial, middle, deep, and calcified zone. It lacks nerves and blood vessels and consists of a few chondrocytes embedded within an extracellular matrix which is composed of water, collagen II, proteoglycans, noncollagenous proteins, and glycoproteins (Figure 1.2, left) [43,44]. Articular cartilage, which is vital for keeping joints working well and preventing damage to bones, not only withstands high loads but also provides a remarkably low coefficient of friction. This combination enables smooth movements of bones even under high pressure [45,46].

The superficial zone, the outermost and thinnest layer of articular cartilage, makes up about 10-20% of the total cartilage thickness. This zone experiences fluid flow and high pressures because of its contact with the synovial fluid and the opposing cartilage (Figure 1.2, right). To ensure smooth movement at the joint surface, effective lubrication relies on charged macromolecules like lubricin and hyaluronic acid, situated on the superficial zone's surface [47,48].

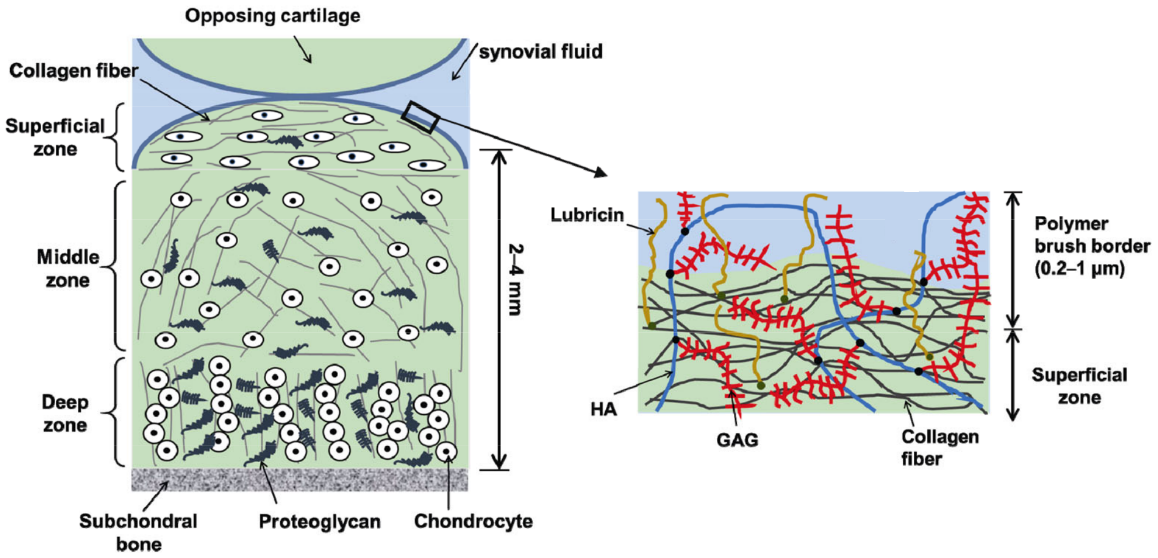

Figure 1.2. Schematic representation of articular cartilage structure and its surface. The left side illustrates the whole structure of articular cartilage and its zones, while the right side focuses on the superficial zone and its major components. Figure adapted from ref. [46]

1.1.2 Synovial fluid

Synovial fluid (SF) is a highly viscous, heterogeneous, and non-Newtonian (strongly shear thinning) fluid that fills the spaces in the joint cavities (Figure 1.1). In synovial joints, the principal role of SF is to provide both lubrication and nutrition. The typical volume of synovial fluid in a regular human knee joint is approximately 1-2 mL. Its contribution, along with cartilage, ensures synovial joints move smoothly and stay healthy [49–51].

The SF displays intricate rheological and tribological properties under shear. Its viscosity ranges from 10–100 Pa·s at zero-shear viscosities to approximately 0.01 Pa·s at higher shear rates [52, 53]. The viscoelastic characteristics of SF arise from components such as hyaluronic acid and PRG4 (lubricin), along with various other elements like albumin, γ-globulin, and lipids. Table 1.1 shows the composition of SF and each constituent's concentrations [50].

1.2 Hyaluronic acid

Hyaluronic acid (HA), also known as hyaluronan, is a linear anionic biopolysaccharide and glycosaminoglycan (GAG) chain. It was discovered in 1934 by scientists Karl Meyer and John Palmer, who extracted it from bovine eyes. HA is composed of approximately 2000-25000 repeating disaccharide units, specifically D-glucuronic acid (GlcA) and N-acetyl-D-glucosamine (GlcNAc), linked by $\beta(1\rightarrow3)$ and $\beta(1\rightarrow4)$ glycoside bonds, respectively.

Table 1.1. Key chemical composition and concentrations of main constituents in normal synovial fluid [50, 54, 55].

Constituent	Notation	Concentration (mg/mL)
Glycosaminoglycan	HA	1–4
Glycoprotein	PRG4	0.05–0.35
Serum protein	Albumin	8–12
	γ-Globulin	2–7
Ions	Na^+, Cl^-	7.28
Lipids		0.17–0.28
Small molecules	Urea, Glucose	0.756
Water		

Figure 1.3 illustrates the structure of HA and its units. This unbranched molecule carries a negative charge and is found in various parts of the body, including the skin, eyes, heart valves, and joints. Its molecular weight can reach up to 10^7 Da. HA exhibits strong hydrophilic properties due to the presence of COO- and OH- groups. In synovial joints, the concentration of HA is about 3 mg/mL. Previous studies have shown that the viscoelastic properties of synovial joints are strongly dependent on HA [56–61].

The rheological and structural properties of HA have been extensively investigated. Luigi Ambrosio et al. [62] and Jihoon Kim et al. [63] demonstrated that HA can exhibit either Newtonian or non-Newtonian behavior at varying concentrations. Their findings revealed that an increase in the mass density of HA can shift the solution's response from Newtonian to shear thinning. Furthermore, the effects of ion concentration [64], system pH [65], crosslinking [66], and molecular weights [67] on HA behaviors have also been investigated.

1.3 Lubricin

Studies have shown that synovial joints need another vital ingredient to provide excellent lubrication and work smoothly, which is called lubricin. In boundary conditions (when two cartilages are under high load and very close to each other), lubricin plays a key role in keeping our joints healthy by forming a protective and slippery layer on the surfaces of cartilage. Lubricin, the superficial zone protein, reduces friction on the surface of cartilage [12–14]. It was first discovered in 1985 by David Swann [68].

Lubricin is encoded by the proteoglycan 4 gene and is expressed by synoviocytes and superficial zone chondrocytes. It exists as a mucin-like glycoprotein and it is an intrinsically

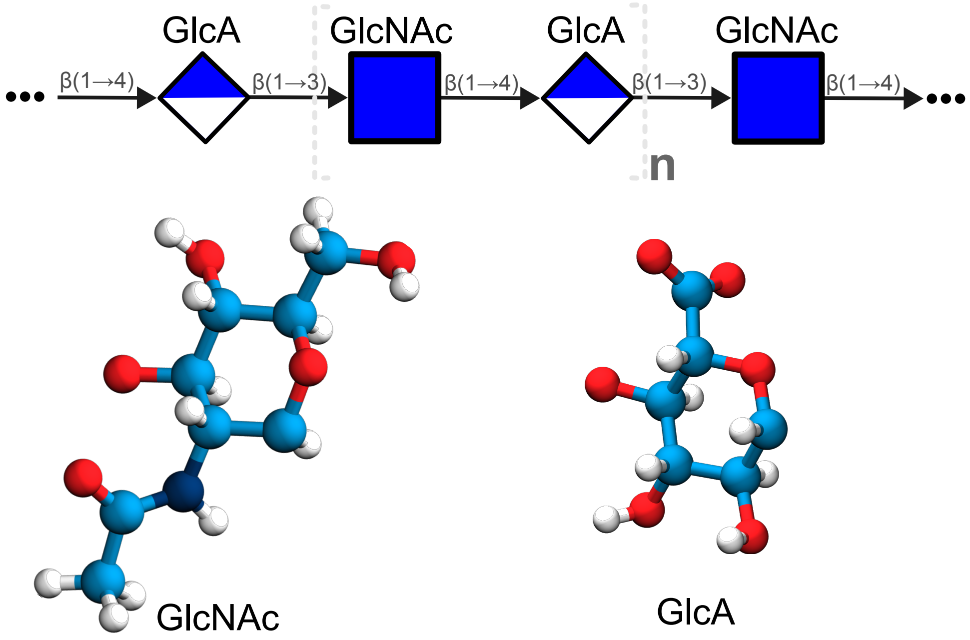

Figure 1.3. Schematic and atomistic representation of hyaluronic acid (HA) and its chemical structure. HA is composed of GlcA and GlcNAc sugars and it can consist of up to n=25,000 units [57].

disordered protein, lacking a specific structure, with 1404 amino acids and a length of approximately 200 nm ± 50 nm. Structurally, lubricin consists of three primary components: two end domains designated as the N- and C-terminus (SMB- and PEX-like domains), and a central domain recognized as the mucin-like domain (Figure 1.4) [69–71].

The terminal domains of lubricin consist of two somatomedin B (SMB)-like domains at the N-terminus and a hemopexin (PEX)-like domain at the C-terminus, depicted in blue and green in the figure 1.4, respectively. These domains showcase a distinct and ordered structure, as illustrated in the 3D representation in figure 1.6. The terminal, globular, and adhesive domains of lubricin, characterized by their non-glycosylation, positive charge, and hydrophobic nature, enable effective attachment to the surface of cartilage proteins [72–75]. This property allows the rest of the protein to extend into the synovial joint cavity, contributing to lubrication and joint functionality. The binding to oligomeric matrix protein (COMP), hyaluronic acid, collagen II, and fibronectin anchors lubricin to the surfaces of cartilage [76–79].

Unlike the ordered structure of the end domains, the mucin-like domain lacks a specific structure and is highly O-linked glycosylated. This region comprises approximately 800 amino acids and extends to around 200 nm. It is rich in KEPAPTT motifs, proline, threonine, and serine residues. Core I and core II types of glycans bind to threonine and serine residues, comprising 30-35% of the whole structure. These O-glycans impart a neg-

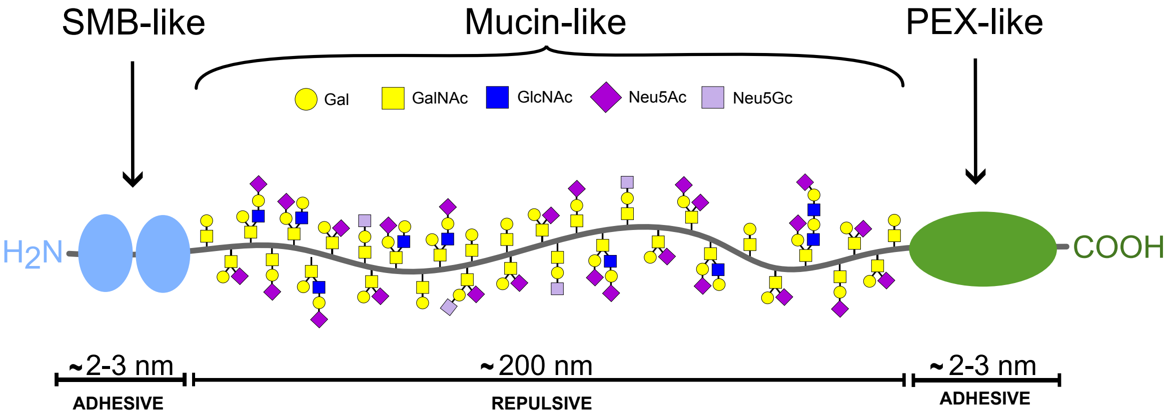

Figure 1.4. Structural overview of lubricin. Comprising two globular end domains (N- and C-terminus) and a central mucin-like domain, lubricin plays a vital role in joint lubrication. The sticky end domains anchor the protein to cartilage, while the glycosylated central domain reduces friction and provides essential boundary lubrication in joints [8, 12].

ative charge, rendering the region highly hydrophilic. The mucin domain is responsible for the highly lubricating abilities of lubricin [80, 81]. For lubricin to be effective at boundary lubrication, both its ability to bind to cartilage and the presence of O-linked glycans are required [82].

1.3.1 O-glycans of lubricin

Previous studies have shown that lubricin undergoes post-translational modifications with two types of O-linked glycans, namely CoreI and CoreII. These glycans are located in the central region, specifically the mucin-like domain, and bind to serine (Ser) and threonine (Thr) residues, making up about 150 kDa of the entire structure. These glycans play an important role in providing boundary lubrication and reducing friction within joints. [68, 83–85].

In 2015, E. Svala and colleagues [86] conducted research on synovial fluid from horses and discovered that lubricin consists of eleven different types of CoreI and CoreII O-glycans (shown in Figure 1.5A). They found that approximately 90% of these glycans belong to the CoreI type, with the remaining 10% being CoreII. These glycans are composed of sugars such as Gal (D-Galactose), GalNAc (N-Acetyl-D-Galactosamine), GlcNAc (N-Acetyl-glucosamine), Neu5Ac (N-Acetyl-Neuraminic Acid), and Neu5Gc (N-Glycolyl-Neuraminic acid). Notably, Neu5Ac and Neu5Gc carry a negative charge. Figure 1.5A provides an overview of the overall structure of these glycans, while Figure 1.5B illustrates the detailed atomic arrangement of each sugar unit.

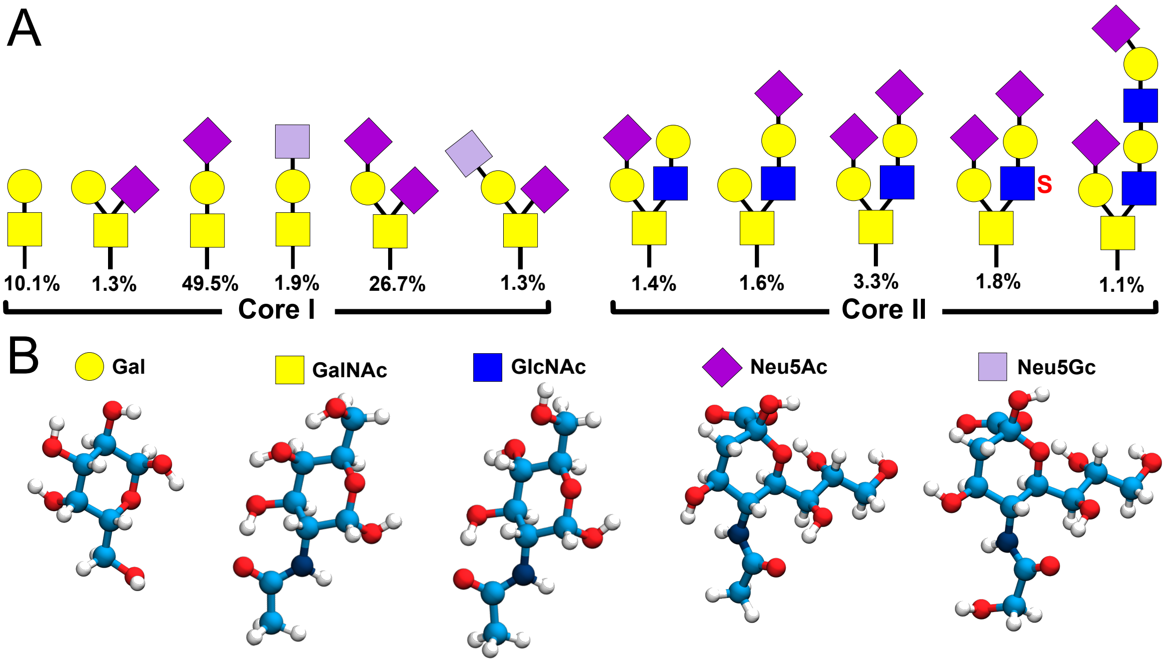

Figure 1.5. Variety of lubricin O-glycans. (A) The 11 O-glycan types in lubricin are categorized into CoreI and CoreII, composed of five distinct sugar groups. (B) Atomistic representation of individual sugar units.

1.4 Intrinsically disordered proteins (IDPs)

Lubricin consists of three parts: the two end domains are globular (Figure 1.6, in blue and green), while the central domain lacks a specific structure and is a disordered region (Figure 1.6 in pink). Intrinsically disordered proteins (IDPs) lack a stable secondary or three-dimensional structure.

This lack of a confined 3D structure can create a larger interaction surface and can be more suitable for post-translational modifications, allowing the localization of their structure to introduce additional features. IDPs are usually rich in proline and charged residues but have a low proportion of hydrophobic residues [88–90]. In the human body, around 51% of all human proteins are IDPs or have an IDP region. They play a variety of important functions in the human body across eight functional classes, including molecular assembly, cellular signaling, transcription, molecular recognition, and more [91–93].

There are several experimental techniques to elucidate the structural and dynamic properties of IDPs, such as Nuclear Magnetic Resonance (NMR) Spectroscopy, Circular Dichroism (CD) Spectroscopy, Fluorescence Spectroscopy, and Small-Angle X-ray Scattering (SAXS) [94]. Another method to study IDPs is molecular dynamics simulations. Given the limitations of atomic resolution in experimental methods, molecular dynamics simula-

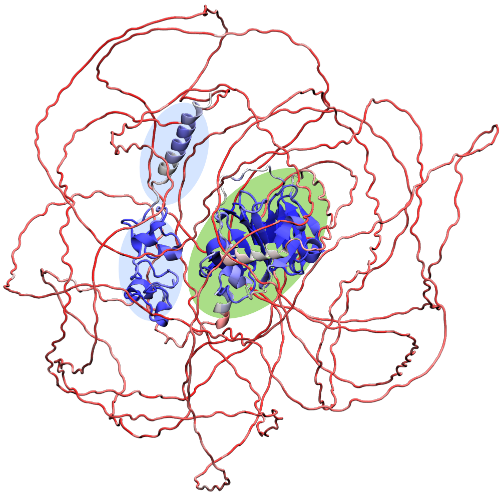

Figure 1.6. A three-dimensional view of lubricin's structure. Lubricin is composed of three structural components. It has two globular end domains (N and C terminus) highlighted in blue and green. The central mucin-like domain of lubricin has an intrinsically disordered structure depicted in pink. This visualization is rendered by using a PDB-code from AlphaFold [87].

tion emerges as a valuable tool to complement ensemble-averaged results of experimental techniques, facilitating a deeper understanding of the structural and dynamic properties of IDPs at the atomic level [95–97]. In the next chapter, I will delve further into this technique and provide detailed insights into utilizing MD to generate reliable results for IDPs.

1.5 Aims of the thesis

Despite the abundance of data on the synovial joint, the connection between the physicochemical properties of the molecular components and the macroscopic low-friction and excellent lubrication response remains inadequately understood. A pivotal player in this scenario is lubricin. This intrinsically disordered glycoprotein plays a crucial role in achieving excellent lubrication in synovial joints. However, there is a lack of an atomistic-level understanding of how lubricin functions. In particular, the influence of glycans on the rheological and viscoelastic behavior of lubricin remained unknown. The primary objective of this thesis is to address this question. To achieve this goal, molecular dynamics simulations and associated computational tools are employed. Through the study and analysis of individual and multi-chain lubricin fragments, this research investigates how these fragments respond to varying levels of glycosylation and external shear stresses. The connection to the macroscopic viscoelastic response is established through the calculation of the viscosity and the structural properties that explain it.

1.6 Organization of the thesis

The thesis is divided into three main chapters. Chapter 2 introduces the methods and materials used in the study, focusing on the theoretical background of molecular dynamics simulations. It details how we prepared a suitable force field for studying glycoproteins, specifically intrinsically disordered ones. Additionally, it explores methodologies for viscosity calculations, encompassing the Green-Kubo method for zero shear viscosity determination and the box deformation method used to evaluate viscosities under external shear stress.

The initial results chapter (chapter 3) explores the impact of O-glycans on the structure of lubricin fragments. This chapter details our modeling approach for achieving atomic-level resolution of lubricin fragments and presents the outcomes of equilibrium molecular dynamics simulations for various glycosylated fragments of lubricin. Here we aim to understand the impact of O-glycans on the structural properties of lubricin.

Chapter 4 sheds light on the viscoelastic response of lubricin fragments to the different levels of glycosylation and external shear stresses. In this chapter, by computing the zero shear viscosity and viscosities, we aim to show how O-glycan can affect the viscosity induced by lubricin.

In Chapter 5, the thesis closes with final concluding remarks and an outlook of future work.

2

Theory and methods

The advent of computers in the 20th century marked a juncture in the landscape of scientific research. It altered the way researchers conduct their studies and introduced a new field to science called computational physics. Computer simulations can provide a direct connection between the microscopic details of a system (such as the type, mass, and interactions of atoms) and its macroscopic properties (pressure, volume, temperature, and viscosity). In addition, in fields such as materials science, chemistry, and molecular biology, computer simulations assist researchers in exploring conditions that are challenging or risky to achieve in experiments. Computer simulation proves to be a valuable tool, particularly when experiments encounter limitations, such as resolution issues or challenges involving extreme temperatures or pressures [98–100].

In the field of molecular simulations, several computational methods are employed, including Molecular Dynamics (MD), Monte Carlo (MC), and Brownian Dynamics (BD) modeling techniques. In this study, MD simulations have been utilized to elucidate the structure, function, and interactions of proteins. This method predicts the movement of atoms and molecules, as well as their interactions, based on physical principles derived from Newtonian physics [101–103]. For historical context, the first MD simulation took place in 1957, conducted by Alder and Wainwright [104]. Around two decades later, in 1977, McCammon et al. [105] achieved another milestone by conducting the first MD simulation of a protein. Over the years, with advancements in computer power and improvements in simulation methods and force field, MD simulations have evolved from simple statistical models to powerful tools capable of predicting properties of biological (like proteins, lipid membranes, DNA, and RNA) and industrial materials.

In the upcoming section, I will delve into the principles and algorithms governing MD simulations. Following that, we will explain the general theories and methodologies employed in this study, such as equilibrium and non-equilibrium MD simulations. Lastly, I will elaborate on calculating zero shear viscosity and viscosity at various shear rates derived from MD simulations.

2.1 Molecular dynamics simulations

In quantum mechanics, the dynamics of a molecular system is described by the Schrödinger equation:

$$\hat{H}|\psi(t)\rangle = -i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle. \tag{2.1}$$

In this equation, $\hat{H}$ describes the Hamiltonian, which is the sum of the kinetic and potential energy operators ($\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})$). Here, ψ stands for the wave function, and $\hbar$ represents the reduced Planck constant. However, solving the Schrödinger equation for systems with many particles, such as molecules containing numerous atoms (like a proteins or DNA), is computationally expensive and often impractical.

The Born-Oppenheimer approximation [106] allows us to separate the movement of atomic nuclei from the motion of electrons. This separation allows us to treat electronic and nuclear motions independently. This approximation simplifies the equations, breaking them into two parts: one for the changing position of nuclei and another for the stable states of electrons. The Ehrenfest theorem [107] can then be applied to describe the motion of atomic nuclei in a classical manner. Essentially, we move from a quantum mechanics description of atoms to a classical mechanics depiction of their motion. Studying systems in classical mechanics is simpler, demands fewer computational resources, and is well-suited for exploring biological systems such as proteins and DNA.

Transitioning to classical mechanics, the Newton's equations of motion describes the dynamics of particles. The MD simulation numerically solves Newton's equation:

$$\begin{aligned} m_i\ddot{\boldsymbol{r}}_i &= \boldsymbol{F}_i, \qquad\qquad (i = 1,...,N) \\ m_i\frac{\partial^2\boldsymbol{r}_i}{\partial t^2} &= -\sum_{i\neq j}\frac{\partial V_{ij}}{\partial r_i}, \end{aligned} \tag{2.2}$$

to study the motion and behavior of atoms and molecules over time. Here, m, r, F, and V represent mass, position, force, and potential energy, respectively. Solving these equations for a system of N particles in three dimensions results in 6N first-order differential equations.

By solving these equations at each time step, the position and velocitiy of the particles can be determined.

In addition to the foundational Born-Oppenheimer approximation and the classical treatment of atomic dynamics, MD simulation relies on another critical assumption: the utilization of empirical force fields for describing interactions and the energy of the system. In the next section, we will explain the force field in details.

2.1.1 Force fields and interaction potentials

In molecular dynamics simulations, we make certain assumptions, and the third assumption necessitates a framework for discussing the system's energy. This framework is called a force field. A force field uses simple arithmetic functions to describe how atoms interact It includes potential energies (V_i), which are functions of the atomic positions r_i, describing all the important connections between atoms.

$$\boldsymbol{F}_i = -\nabla V_i = \nabla V(r_1, r_2, ..., r_N) \tag{2.3}$$

The potential energy (V_i) of a system is typically divided into bonded and non-bonded terms. Bonded terms involve interactions between atoms that are connected by chemical bonds. This includes bond stretching (V_b), angle bending (V_a), dihedrals (V_{dih}), and improper dihedral interactions ($V_{imp.dih}$). Non-bonded terms deal with interactions between atoms that are not directly linked by bonds. These interactions are modeled through a Lennard-Jones potential (V_{LJ}) for van der Waals interactions and a Coulombic potential ($V_{coulomb}$) for electrostatic interactions. These interactions are illustrated in Figure 2.1. A typical expression for a force field function may defined as [108]:

$$V = \left[V_{bond} + V_{angle} + V_{dihedral} + V_{imp.dihedral}\right]_{bonded} + \left[V_{LJ} + V_{coulomb}\right]_{nonbonded} \tag{2.4}$$

The bond energy in molecular dynamics force fields comes from the interaction between two atoms, labeled i and j, when they form a covalent bond. This covalent bond is like a connection between the atoms, and is described as a harmonic potential energy function,

$$V_{bond} = \sum_{ij} \frac{1}{2} k_{ij}^{b} (r_{ij} - b_{ij})^2, \tag{2.5}$$

where k_{ij}^b, $r_{ij} = |r_j - r_i|$, and b_{ij} represent the bond strength parameter, the actual distance between the two bonded atoms, and the equilibrium bond length, represently.

In molecular dynamics force fields, the interaction among sets of three particles (labeled i, j, and k) is described as an angular restriction. We use a simple formula, called a harmonic

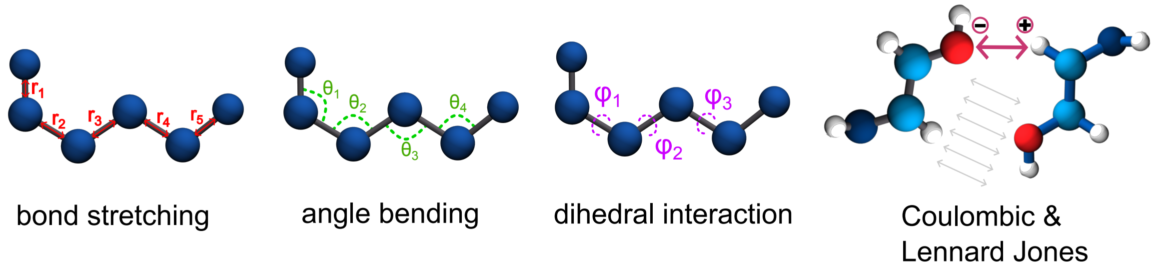

Figure 2.1. Bonded and non-bonded potential types in a classical force fields. A classical force field comprises two types of potentials: bonded (such as bond, angle, and dihedral) and non-bonded (including LJ and Coulomb) interactions.

potential energy function, to calculate the energy associated with bending this angle. The formula is represented as:

$$V_{angle} = \sum_{ijk} \frac{1}{2} k^{\theta}_{ijk} (\theta_{ijk} - \theta^{0}_{ijk})^2, \tag{2.6}$$

where k^{θ}_{ijk} is the harmonic force constant, θ_{ijk} is the angle formed by the three bonded atoms, and θ^{0}_{ijk} is the equilibrium angle.

Dihedral interactions, also known as torsional or proper dihedral terms, represent the potential energy associated with the rotation of a series of four connected atoms (ijkl) around a central bond. The dihedral angle represents the twist between two planes formed by pairs of atoms (i, j, k) and (j, k, l). The potential energy function for a dihedral angle is commonly represented as:

$$V_{dihedral} = \sum_{ijkl} \frac{1}{2} k^{\phi}_{ijkl} (1 + cos(n\phi_{ijkl} - \phi^{0}_{ijkl})), \tag{2.7}$$

where k^{ϕ}_{ijkl} is the force constant, ϕ_{ijkl} is the dihedral angle, and ϕ^{0}_{ijkl} is the equilibrium dihedral angle.

Improper dihedrals are crucial in molecular dynamics simulations, maintaining planarity in groups like aromatic rings and preventing distortions. The improper dihedral potential energy is typically modeled using a harmonic potential energy function,

$$V_{imp.dihedral} = \sum_{ijkl} \frac{1}{2} k^{\xi}_{ijkl} (\xi_{ijkl} - \xi^{0}_{ijkl})^2, \tag{2.8}$$

, considering the force constant (k^{ξ}_{ijkl}), the improper dihedral angle formed by the four atoms (ξ_{ijkl}), and the equilibrium improper dihedral angle (ξ^{0}_{ijkl}). Unlike proper dihedrals, improper dihedrals serve the purpose of preserving planarity in specific molecular groups and regulating the orientation of substituents around a central atom without involving rotations about a bond.

The Lennard-Jones (LJ) potential, also known as the 12-6 potential, is a mathematical model to describe the van der Waals interactions between atoms or molecules. The potential energy between two atoms, i and j, as a function of their separation distance r_{ij} is given by:

$$V_{LJ} = \sum_{ij} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right]. \tag{2.9}$$

Here, ε_{ij}, the depth of the potential energy well, represents the strength of the interaction, and σ_{ij} is the distance at which the intermolecular potential between atom i and j becomes zero. The LJ potential consists of two terms. The attraction term, $\left(\frac{\sigma_{ij}}{r_{ij}}\right)^6$, represents van der Waals attraction at long-ranged interactions, and the repulsion term, $\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}$, represents Pauli repulsion at short distances.

Coulomb interaction, also known as electrostatic interaction, refers to the potential energy arising from the interaction between charged particles. The Coulomb potential energy between two point charges, separated by a distance r_{ij}, is given by Coulomb's law:

$$V_{coulomb} = \sum_{ij} \frac{1}{4\pi\varepsilon_0 \varepsilon} \frac{q_i q_j}{r_{ij}^2}. \tag{2.10}$$

Here in this equation, ε_0 and ε are the vacuum permittivity and the relative permittivity of the medium. q_i and q_j are the magnitudes of the charges of particles i and j.

The parameters of bonded and non-bonded potentials (parametres in eq. 2.5 to 2.10) are determined through quantum mechanical calculations or adjusted by fitting them to experimental data from various sources like neutron, X-ray, and electron diffraction, NMR, etc [109, 110].

Several force fields, including amber [111], CHARMM [112], OPLS-AA [113], and GROMOS [114], optimized for specific types of molecules, are available. In our research, we employed two force fields to cover all materials in our system: amber99sb-star-ildnp [115, 116] for proteins and GLYCAM06 [117] for carbohydrates.

2.1.2 Integration of equations of motion

Once the interaction potentials (V) in the force field are defined, the classical Newtonian equation 2.2 allows us to track the positions, velocities, and accelerations of particles over time. Given that this equation operates in three dimensions and there are N particles in the system, it results in a set of 6N interconnected first-order differential equations. Analytically solving these 6N equations is impractical. Therefore, numerical methods are essential for integrating these equations over time. Numerical integration methods in MD simulations rely on Taylor expansions to approximate the continuous functions of particle positions and velocities at future time steps, based on their current values. This discretization of continuous motion enables the numerical integration of equations of motion. Various integration algorithms, including Euler, Verlet, Beeman, Predictor-Corrector, and the Runge-Kutta algorithm, exist. In our study, we employed the leapfrog algorithm [118, 119].

The leapfrog algorithm [120] is widely used as a numerical integration method in computational physics, especially in the context of MD simulations. This is based on the second-order Taylor expansion. The leapfrog algorithm updates particle positions and velocities using the following equations:

$$\begin{aligned} \boldsymbol{v}_i\left(t+\frac{\Delta t}{2}\right) &= \boldsymbol{v}_i\left(t-\frac{\Delta t}{2}\right)+\frac{\boldsymbol{F}_i(t)}{m_i}\Delta t, \\ \boldsymbol{r}_i\left(t+\Delta t\right) &= \boldsymbol{r}_i(t)+\boldsymbol{v}_i\left(t+\frac{\Delta t}{2}\right)\Delta t. \end{aligned} \tag{2.11}$$

Despite the Leapfrog method being well-known for its energy conservation and stability, it does have certain limitations. In this model, position and velocity cannot be calculated at the same time, and its performance is highly sensitive to the chosen time step (Δt).

2.1.3 Simulation details and ensembles

The time step Δt plays a crucial role in determining both the speed and precision of MD simulations. It needs to strike a balance between being large enough to enhance simulation speed and being small enough to accurately represent high-frequency vibrations in covalent bonds, especially those involving hydrogen atoms and heavy atoms. Employing algorithms like LINCS [121], SHAKE [122], and Settle algorithm [123] to constrain bond lengths or angles can effectively eliminate high-frequency vibrations, allowing for the use of larger integration time steps ($\Delta t \geq 2$ fs).

Short and long range non-bonded interactions

Calculating non-bonded interactions in the force field (equation 2.9 and 2.10) is the most demanding part of the integration step, involving a double sum over all atoms. A system with N particles has $N(N-1)/2$ pairwise interactions, leading to a quadratic increase in computational cost. To enhance efficiency, short-range Lennard-Jones interactions are considered within a 1.0 - 1.2 nm cutoff distance. The Verlet list [124], by introducing a second cutoff distance for short-range interactions, effectively reduces computational costs. Nearest neighbor information is updated every 20 fs. For long-range electrostatic interactions, a cutoff scheme is avoided due to potential artifacts. Instead, the particle-mesh Ewald (PME) method [125, 126] is employed, separating the Coulomb term into short and long-range components. PME significantly improves scalability compared to direct calculation, thereby enhancing computational efficiency for both short and long-range interactions in the simulation. For a system with N particles, Cutoff methods and PME can significantly reduce the computational cost to $\mathcal{O}(N)$ and $\mathcal{O}(N \log N)$, respectively [127].

Periodic boundary conditions

In MD simulations, we simulate a system within a finite box known as the simulation cell. However, due to the limited size of the box and the finite number of particles, an artificial effect occurs near the edges of the box. This effect can impact the behavior of particles and impede the simulation of macroscopic systems. To address this limitation and move closer to modeling macroscopic systems, periodic boundary conditions (PBC) can be used. Instead of having particles interact with the physical boundaries of the box, PBC allows particles exiting one side of the box to re-enter on the opposite side.

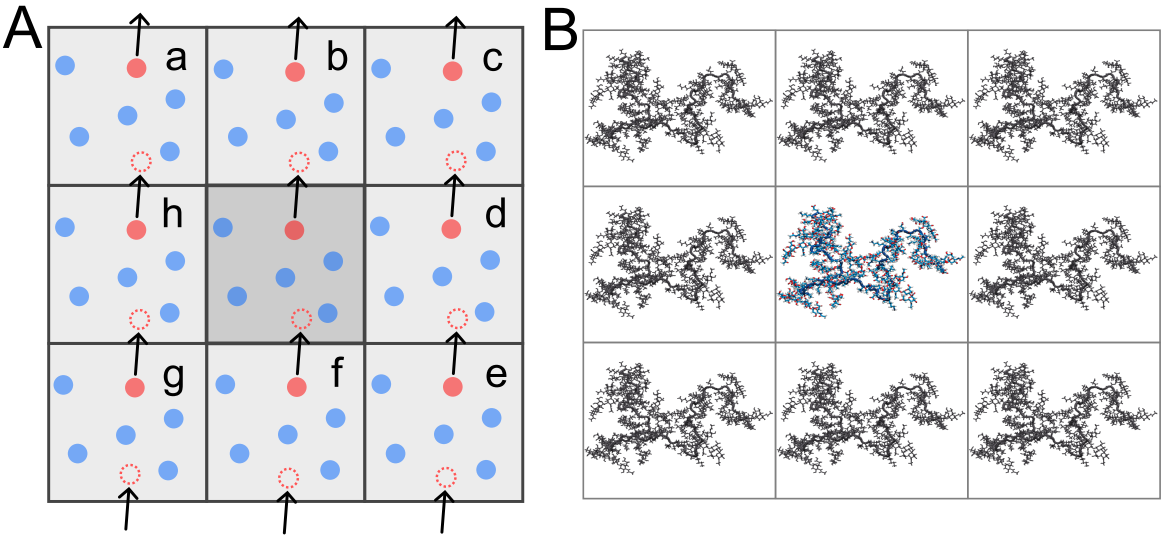

Figure 2.2. Periodic boundary conditions in 2D. A. This figure visually depicts the application of periodic boundary conditions, allowing particles to enter or exit from all edges of the simulation box. PBC enables the study of systems as if they were part of an infinite space. **B.** The figure shows a real MD box, with the protein in the center, surrounded by its periodic images.

To implement periodic boundary conditions (PBC), we consider an infinite set of replicated boxes, referred to as imaged boxes, surrounding the main simulation box. Each imaged box is identical to the main box, containing particles that move and behave in the same way.

When a particle crosses the border of the main box, a corresponding particle from the imaged box enters the main box. As a result, the total number of particles, density, and charge remain constant throughout the simulation. Figure 2.2A provides a visual representation of this concept in 2D, with imaged boxes labeled a, b, c, d, e, f, g, and h. Figure 2.2B shows a protein in the center, surrounded by its periodic images.

Temperature and pressure control

The utilization of Newton's equations of motion, incorporating the leapfrog algorithm within a periodic box, results in a microcanonical ensemble (NVE). This ensemble maintains a constant number of particles (N), box volume (V), and system energy (E). In laboratory conditions where experiments typically operate at a fixed temperature (T) and pressure (P), it becomes essential to control temperature and pressure through the simulation process, giving rise to canonical (NVT) and isothermal-isobaric (NPT) ensembles. This coupling ensures a more accurate representation of the system's behavior by maintaining a constant temperature and pressure, aligning with experimental settings.

Temperature coupling involves regulating the system's temperature, commonly achieved using a thermostat algorithm. This tool adjusts the kinetic energy of particles to maintain the desired temperature. In contrast, pressure coupling aims to regulate the pressure within the system, often accomplished using barostats. Barostats adjust the volume of the simulated system and modify the actual positions of particles to maintain the pressure at the desired value. In Gromacs [108], various methods are available for temperature and pressure coupling. Thermostat options include Berendsen [128], Nose-Hoover [129, 130], and V-rescale [131], while barostat options encompass Berendsen [128], Parrinello-Rahman [132], and c-rescale [133].

In our research, we employ the V-rescale and Parrinello-Rahman protocols to precisely regulate both temperature and pressure according to our reference values. The V-rescale method adjusts the kinetic energies of particles, ensuring temperature stability by rescaling velocities during each simulation step. On the other hand, the Parrinello-Rahman protocol manipulates the simulation box dimensions to maintain the desired pressure conditions. Together, these protocols contribute to the accurate and controlled simulation environment essential for our study.

2.1.4 Lee-Edwards boundary conditions

Typically, classical periodic boundary conditions (PBC) are employed in MD simulations to mimic an infinite system. However, in some cases in non-equilibrium molecular dynamics

(NEMD) simulations, such as the deformation method, classical PBC may not be a suitable choice. Nevertheless, an alternative approach exists, known as Lees-Edwards periodic boundary conditions (LEBCs), which offer a unique combination of laminar flow and periodic boundaries. LEBCs, also referred to as sliding brick boundary conditions, help simulate a large system experiencing a simple shear flow. This method is particularly valuable when simulating large systems undergoing a simple shear flow, and it deals well with finite size effects. The concept of LEBCs was introduced by Lees and Edwards in 1972, and since then, it has become a widely adopted technique in shear-flow MD simulations. By permitting particles to cross the simulation boundary, LEBCs enable changes in position and speed in response to the boundary movement, enhancing the realism of the simulation [134–136].

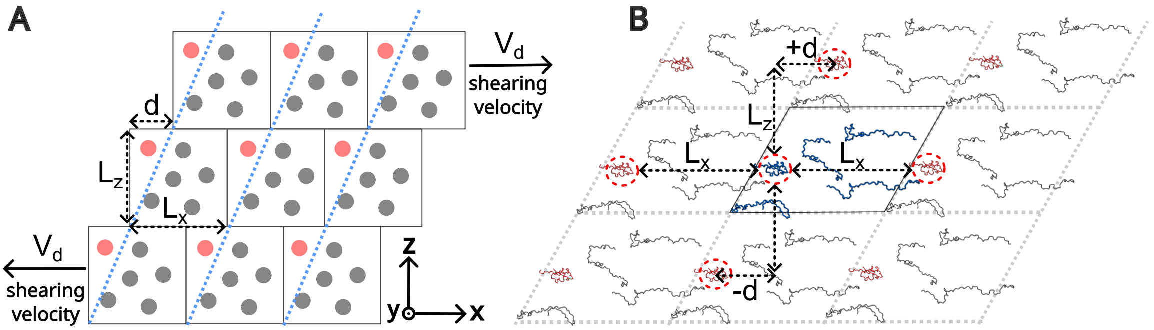

Figure 2.3. Lee-Edwards boundary conditions in 2D. A. Schematic visualization of Lee-Edwards Boundary Conditions (LEBCs). In the case of shear flow along the z-direction, the effect of LEBC will be similar to classical PBC when a particle crosses the edges from the x and y directions. However, LEBC will induce changes in the position and velocities of particles that cross the z-direction. **B.** Application of LEBCs in a MD simulations of polymers. Proteins are situated in the center, surrounded by their periodic images, showcasing the effectiveness of LEBCs in handling finite size effects and simulating shear-flow scenarios.

In Lee-Edwards boundary conditions, instead of replicating the simulation box in all three dimensions, PBC is applied only along two directions (typically x and y), while the positions and velocities in the third dimension (typically z) are updated differently. This modification accounts for the shearing motion introduced in the simulation, allowing the system to mimic the effects of continuous shear flow. Figure 2.3A illustrates a simple 2D representation of LEBC. Moreover, Figure 2.3B presents a real example of LEBC derived from molecular dynamics simulations of polymers. In LEBCs, particles are allowed to cross the simulation boundary. When a particle crosses the boundary, adjustments are made to its position and velocity to account for the boundary movement. This ensures that particles can move smoothly through the simulation box, mimicking the behavior expected in a system undergoing shear flow. Mathematically, the transformation of coordinates (x, y, z) to the new coordinates (x', y', z') under LEBCs can be expressed using a matrix: [redacted]

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & \pm\dot{\gamma}\Delta t L_z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{2.12}$$

Likewise, the transformation of velocity coordinates (v_x, v_y, v_z) to new coordinates (v'_x, v'_y, v'_z) under LEBC is expressed using a similar matrix:

$$\begin{pmatrix} v'_x \\ v'_y \\ v'_z \end{pmatrix} = \begin{pmatrix} 1 & 0 & \pm\dot{\gamma} L_z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}, \tag{2.13}$$

In equations 2.12 and 2.13, $\dot{\gamma}$ represents the shear rate and t denotes time. Shear rate is equivalent to the ratio of shearing velocity to the box length orthogonal to the shear flow ($\dot{\gamma} = V_d/L_z$). Additionally, in Figure 2.3, the variable d corresponds to $\dot{\gamma}\Delta t L_z$ [137–139]. However, it is crucial to emphasize that equations 2.12 and 2.13 are transformations specifically designed for a jump across the z-axis. The normal periodic boundary conditions apply for jumps across the x and y axes.

2.2 Viscosity

Viscoelastic materials possess both viscosity and elasticity characteristics. Viscosity, a fundamental attribute of fluids, is crucial for understanding and predicting the behavior of liquid flow. It represents a fluid's resistance to deformation or flow, reflecting the internal friction among its molecules, and is measured in pascal-seconds ($Pa.s$) in the SI units. Viscosity is a dynamic fluid property influenced by factors such as temperature, composition, molecular structure, concentration, pressure, and shear rate. The shear viscosity is a key parameter characterizing how a fluid responds under shear stress. Understanding these rheological behaviors is essential for various applications, ranging from industrial to biological systems. Fluids subjected to shear rates can display diverse responses, including Newtonian behavior, shear-thinning, or shear-thickening, contingent on their specific characteristics and external conditions. Figure 2.4 illustrates the response of a system to external shear flow. In Newtonian systems, viscosity remains independent of shear rate. Conversely, in shear-thinning and shear-thickening scenarios, viscosity decreases and increases with rising shear rates, respectively (see Figure 2.4) [140, 141].

The shear viscosity of liquids can be determined using different experimental methods, including capillary viscometers, rotational viscometers, and falling sphere viscometers [142, 143]. Apart from experimental approaches, MD simulations offer alternative

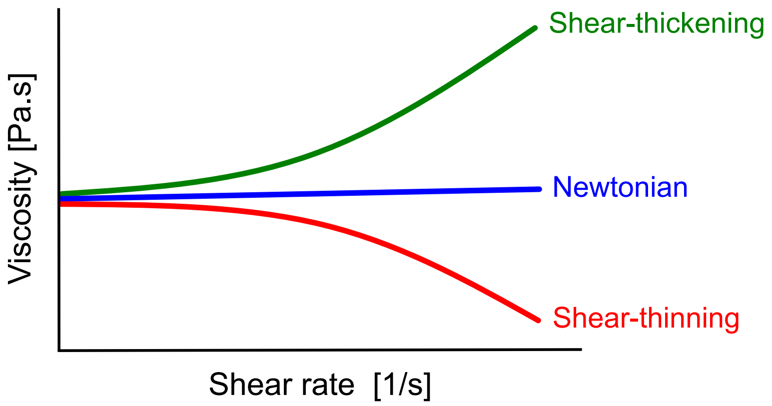

Figure 2.4. Viscosity behavior of fluids with respect to external shear rate The shear rate can have different impacts on fluid viscosity: no change Newtonian behavior (no change), shear-thickening (increase), and shear-thinning (decrease) responses.

means to calculate the viscosity of fluidic systems. In MD simulations, there are generally two main methods: equilibrium molecular dynamics (EMD) based on pressure or momentum fluctuations and nonequilibrium molecular dynamics (NEMD). EMD involves post-processing a standard MD trajectory, employing methods like Einstein–Helfand and Green–Kubo methods. NEMD methods, on the other hand, require adjustments to the fundamental equations of motion or system boundary conditions. Techniques like the SLLOD algorithm, periodic perturbation method, dragging parallel walls, and deforming the simulation box are utilized [144–146]. In our study, we employed two methods to determine viscosities. We used the Green-Kubo method for zero shear viscosity and the box deformation method to assess the viscosity under different shear rates.

2.2.1 Green-Kubo method for zero shear viscosity

Zero shear viscosity (η) characterizes a fluid's resistance to flow when there is no applied shear stress, essentially reflecting viscosity under conditions where the shear rate is zero. In EMD simulations, the Green-Kubo (GK) method is employed to determine the zero shear viscosity of a system [147–149]. The GK method involves simulating an equilibrium fluid under periodic boundary conditions and analyzing time-dependent stress fluctuations. By integrating the autocorrelation function (ACF) of pressure tensor components over time, the GK method yields a mathematical expression for this viscosity (η):

$$\eta = \lim_{t\to\infty} \eta(t) = \frac{V}{k_B T} \lim_{t\to\infty} \int_0^t \langle P_{\alpha\beta}(t) \cdot P_{\alpha\beta}(0) \rangle \, \mathrm{dt}, \tag{2.14}$$

where, η, V, k_B, T, $P_{\alpha\beta}$, and t represent the viscosity, system volume, Boltzmann constant, temperature, components of the pressure tensor, and time, respectively. In this equation, $\langle P_{\alpha\beta}(t) \cdot P_{\alpha\beta}(0) \rangle$ measures the ensemble average of the ACF of the pressure tensor elements, where α and β can take values x, y, or z.

To determine viscosity using the Green-Kubo method, multiple extensive EMD simulations must be conducted. The number and time of these simulations are depended on the size of the system and also the estimated viscosity. Subsequently, we can calculate the ACF of all six independent diagonal and off-diagonal components of the pressure tensor (P_{xy}, P_{xz}, P_{yz}, $(P_{xx} - P_{yy})/2$, $(P_{xx} - P_{zz})/2$, and $(P_{yy} - P_{zz})/2$). Finally, by averaging the viscosities obtained from different pressure elements and different replicas, the final value of viscosity can be obtained [150, 151].

2.2.2 Box deformation method for shear viscosity

In exploring the viscoelastic response of systems to external forces, non-equilibrium molecular dynamics (NEMD) simulations prove to be a powerful tool for understanding dynamic behaviors at the molecular level. NEMD, focusing on non-equilibrium phenomena, provides insights into various aspects, including viscosity. Several methods within NEMD, such as the SLLOD algorithm, box deformation, and dragging parallel walls, are employed to calculate viscosity under different shear rates. These methodologies have been instrumental in studying the molecular-level response of polymer solutions to shear forces [146, 152, 153]. In our study, we specifically utilized the box deformation method in shear-driven non-equilibrium MD simulations to elucidate the viscosity of the lubricin system.

The box deformation method is employed to initiate a planar Couette flow within a fluidic system, as illustrated in figure 2.5A. In this approach, by using Lee-Edwards boundary condition, the simulation box undergoes deformation by laterally displacing the upper wall, which moves consistently at a constant speed '$V_d = u$' along the x-axis. The distance between the upper and lower walls (L_z) is denoted as 'h'. The velocity component (V_x) of all particles within the simulation box is adjusted accordingly. This technique demonstrates effectiveness in generating planar Couette flow, causing shear in the x direction while aligning the velocity gradient with the z direction. Figure 2.5B shows the effect of the box deformation method on a system with polymers. As evident here, this method imposes a homogeneous shear flow on the system.

In Couette flow, the shear rate ($\dot{\gamma}$) is determined by the velocity of the moving wall and the separation between the walls, as given by

$$\dot{\gamma} = \frac{u}{h}, \tag{2.15}$$

and shear viscosity (η) is then calculated by assessing the ratio between the ensemble aver-

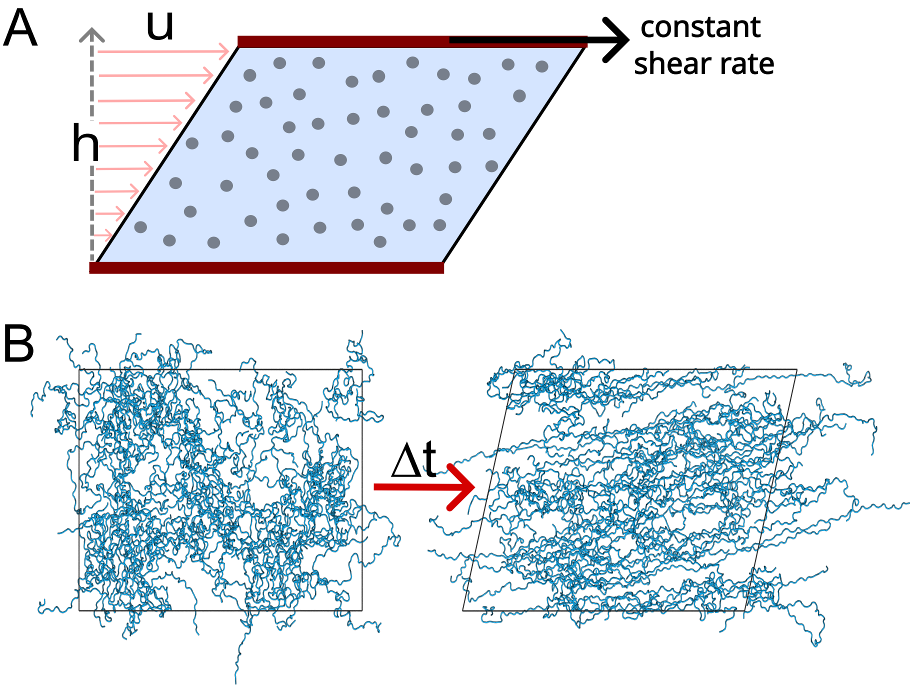

Figure 2.5. Visualization of the box deformation method A. It shows a simple illustration of the box deformation method, which imposes Couette flow through a constant shear rate on the system. **B.** Application of the box deformation method in a real molecular dynamics (MD) simulations with a polymer system.

age of the component of the pressure tensor $\langle P_{xz} \rangle$ and the fluid's shear rate, as expressed

$$\eta = \frac{\langle P_{xz} \rangle}{\dot{\gamma}}. \tag{2.16}$$

To deform the simulation box, we utilized the GROMACS-2023-dev version, incorporating the Lees-Edwards boundary condition, as detailed here. Further information on simulation details is available in Chapter 4.

2.3 Analysis of MD trajectories

Analyzing MD trajectories offers valuable insights into the dynamic behaviors of biological molecules at the atomic level. These trajectories, depicting the positions and velocities of atoms over time, are subjected to various analytical methods tailored to the specific goals and types of MD simulations. Structural analysis, dynamic behavior, thermodynamic analysis, free energy calculations, and so on are key analytical approaches. In our study, we employed the radius of gyration, hydrodynamic radius, end-to-end distance, persistence length, and solvent-accessible surface area (SASA) to assess the structural characteristics of single and multiple protein chains. Additionally, nematic order, p2 value, and SASA were employed to investigate the behavior of multiple protein chains under external forces, specifically shear flow conditions. In the upcoming subsections, I will provide fundamental details about the methods mentioned and guide on their application in the analysis of MD trajectories.

2.3.1 Radius of gyration and hydrodynamic radius

To assess and understand the structural characteristics of proteins in MD simulations, the radius of gyration (Rg) or hydrodynamic radii (Rh) can be used. These methods show the size of structures and are very helpful to determine whether proteins are ordered or disordered.

The radius of gyration (Rg) serves as a size metric for polymers, helping us understand how spread out or coiled up they are. Imagine a polymer as a chain of beads where each bead represents an atom. Rg computes the average distance of atoms from the center of mass of the entire polymer, taking into account the mass of each individual atom. This is expressed mathematically as:

$$Rg = \sqrt{\frac{\sum_i m_i(\boldsymbol{r}_i - \boldsymbol{r}_{cm})^2}{\sum_i m_i}}, \quad \boldsymbol{r}_{cm} = \frac{\sum_i m_i \boldsymbol{r}_i}{\sum_i m_i}, \tag{2.17}$$

where m_i is the mass of each atom, $\boldsymbol{r}_i$ is the position of each atom, and $\boldsymbol{r}_{cm}$ is the center of mass of the polymer [154].

In addition to Rg, hydrodynamic radii (Rh) can also be used to assess the structure of a polymer chain. The Rh, also known as the Stokes radius in polymer physics, refers to the effective size of a polymer chain as it behaves in a fluid or solvent environment. This parameter takes into account hydrodynamic interactions and the polymer's resistance to movement through a solution. The value of Rh is defined by the equation:

$$\frac{1}{Rh} \equiv \frac{1}{N^2} \langle \sum_{i \neq j} \frac{1}{r_{ij}} \rangle, \tag{2.18}$$

where r_{ij} is the distance between atom i and j and N is the total number of atoms. Additionally, the value of Rh may be obtained from the Stokes-Einstein equation:

$$Rh = \frac{k_{\mathrm{B}} T}{6\pi\eta D}, \tag{2.19}$$

here, k_{B} is Boltzmann's constant, T is the temperature, η is viscosity and D is the diffusion coefficient [155].

2.3.2 Persistence length

The persistence length (l_p) in polymer physics serves as a measure of the stiffness or flexibility of a polymer chain. It represents a characteristic length scale that quantifies how far a polymer chain can maintain a relatively straight configuration before undergoing significant bending due to thermal fluctuations. The l_p is determined by examining the rate of exponential decrease in the orientation correlation of bond vectors along the polymer chain:

$$\langle cos\theta(s)\rangle = \langle \boldsymbol{r}_i . \boldsymbol{r}_{i+s}\rangle / l_b^2 = \exp(-sl_b/l_p), \tag{2.20}$$

here, $\langle cos\theta(s)\rangle$ epresents the ensemble average of the cosine of the angle θ between bond vectors separated by a distance s along the polymer chain. $\langle \boldsymbol{r}_i . \boldsymbol{r}_{i+s}\rangle$ denotes the ensemble average of the dot product between the position vectors $\boldsymbol{r}_i$ and $\boldsymbol{r}_{i+s}$ corresponding to two segments separated by s steps along the chain and l_b is the average bond length. This value can also be obtained by considering worm-like chain theory and the mean square end-to-end distance [156–158].

2.3.3 Solvent-accessible surface area (SASA)

Solvent-accessible surface area (SASA) is a quantitative measure that provides insights into the exposure of atoms or residues within a biomolecule to the surrounding solvent. Typically measured in squared nanometers (nm^2), this metric is particularly useful in the study of various aspects of biomolecular behavior, including phenomena such as protein folding, aggregation, ligand binding, and conformational changes.

In our analysis, we obtained SASA values for both single-chain and multiple-chain systems. For single chains, we computed the SASA value for each individual chain, offering an overview of the protein's structure and its compactness. In the case of multiple chains, we calculated both the total SASA for all chains combined (S$_T$) and the SASA for each individual chain (S$_k$). To assess the interaction between chains and characterize their aggregation, we calculated the ratio of the total SASA to the sum of individual chain SASA values:

$$S = \frac{S_T}{\sum_{k=1}^{\#chains} S_k}. \tag{2.21}$$

A value of S equal to 1 indicates complete dissociation of the chains, implying non-interaction. Conversely, a lower S value approaching 0 suggests a higher degree of association among the chains, indicating the formation of a condensed or aggregated state. This approach provides valuable insights into the dynamic behavior and interactions within biomolecular systems during MD simulations [159].

GROMACS, by employing the double cubic lattice method (DCLM) [160], provides a dedicated module, gmx sasa, for computing this crucial structural feature.

2.3.4 Nematic correlation function

The nematic correlation function (NCF) is a mathematical expression that describes the correlation between the orientations of molecules at different positions. It provides information about how the orientation of one molecule influences another at a certain distance or position within the system. In our investigation, we calculated this value to assess how the chains tend to align in relation to each other.

Mathematically, the NCF, often represented as $\Pi_{ij}(r)$, is defined as the average of the product of the molecular orientation at one point and the molecular orientation at another point. The expression for NCF is given by

$$\Pi_{ij}(r) = \langle |\mathbf{t}_i \cdot \mathbf{t}_j| \rangle = \frac{\sum_I \sum_{J>I} \sum_{i\in I} \sum_{j\in J} |\mathbf{t}_i \cdot \mathbf{t}_j| \delta(|\mathbf{r}_i - \mathbf{r}_j| - r)}{\sum_I \sum_{J>I} \sum_{i\in I} \sum_{j\in J} \delta(|\mathbf{r}_i - \mathbf{r}_j| - r)}, \tag{2.22}$$

where i and j represent pairs of amino acids from two distinct protein chains I and J. The range of $\Pi_{ij}(r)$ spans from 0.5 to 1.0. A value of 0.5 denotes a total lack of alignment, indicating that the orientation of fragments is entirely random. In contrast, a value of 1.0 indicates perfect alignment, showcasing a uniform orientation of the fragments [158, 161].

2.3.5 The nematic order parameter P2

Alongside the NCF, the nematic order parameter P2 is another valuable metric for illustrating the orientation of a chain along a specific direction. The formula for P2 is given by:

$$P2 = \langle \frac{3}{2}\cos^2\theta - \frac{1}{2} \rangle = \frac{1}{N}\left[\sum_i^N \frac{3}{2}(\mathbf{u}_i.\mathbf{d})^2 - \frac{1}{2}\right]. \tag{2.23}$$

Here, θ denotes the angle between the average molecular orientation, $\langle ... \rangle$ signifies an average over all molecules in the system, $\mathbf{u}_i$ represents a unit vector connecting two amino acids, and $\mathbf{d}$ is a unit vector defining the preferred alignment direction. The resulting value of P2 spans from 0 to 1, where 0 implies a completely random orientation, and a positive value indicates nematic order or alignment [162–164].

3

Effect of O-glycans on lubricin's structural properties

This chapter delves into the initial phase of our research, where our primary objective was to unravel the influence of O-glycans on the structural properties of lubricin. To accomplish this objective, we conducted equilibrium MD simulations at an all-atom resolution. We combined parameters of IDPs and sugars to present a suitable force field for both protein and glycan atomic interactions. The resulting force field laid the groundwork for our simulations. In the forthcoming sections, we will investigate the construction of this detailed model, elucidating the methodology employed for its generation. Subsequently, we will begin on a comprehensive exploration of our findings and illustrate how O-glycans play a role in altering the structure of lubricin. Finally, we will wrap up with a section dedicated to discussion, conclusions, and a glimpse into future prospects.

3.1 Introduction

Lubricin, also known as PRG4, is a glycoprotein and structureless protein that plays a pivotal role in lubricating and protecting joints. Its primary function involves reducing friction in boundary conditions and facilitating smooth and pain-free joint movements [12–14].

Lubricin comprises 1404 amino acids, with a length of approximately 200 ± 50 nm.

Lubricin's molecular structure is characterized by three distinct parts: two end domains and a central part. The two folded terminal domains, referred to as the N- and C-terminal domains (SMB- and PEX-like domains), are globular proteins with a positive charge and lack glycosylation. On the other hand, the central domain, also known as the mucin-like domain, is highly glycosylated, negatively charged, and hydrophilic. It also lacks a specific structure and is rich in proline residues. Research indicates that the end domains are responsible for anchoring the protein onto cartilage surfaces, while the mucin-like domain primarily functions in providing excellent lubrication within joints. This lubricating property of lubricin is intricately connected to its three structural components [69–75], drawing our attention to the need for their more detailed exploration of these components.

Figure 1.4 shows a schematic illustration of lubricin, its three parts, and O-glycans. In the central region, core I and core II O-linked glycans are attached to the hydroxyl group of serine (Ser) or threonine (Thr) residues. These glycans constitute approximately 30-35% of lubricin's total structure. They also impart a negative charge to the protein through the presence of sialic acid [80–82]. More informatoin about lubricin and its srtucture can be found in section 1.3.

3.2 Atomistic model of lubricin and its O-glycans

In our investigation of the mucin-like domain of lubricin and the impact of O-glycans on this structure using equilibruim MD simulations, we encountered two primary challenges. The first challenge was related to the length of the sequence, while the second challenge was associated with the type of O-glycans and their parameters in the MD force field. This section outlines our approach in addressing and overcoming these challenges to gain insights into lubricin's molecular behavior.

3.2.1 Selecting lubricin fragments

The mucin-like domain of lubricin consists of approximately 800 amino acids [78]. Considering the entire sequence of the central domain in MD simulations would be prohibitively expensive. To address this challenge and make the simulation feasible for all-atom equilibrium MD simulations, we opted to select smaller segments representative of the entire chain. We considered the central domain from amino acid 220 to 1020 and divided this domain into 10 segments, each consisting of 80 residues.

In the next step, we determined the count of glycosylated, charged, and proline residues for each segment, utilizing the sequence provided by Ali Liaqat et al. [69]. We considered proline because previous studies have shown its contribution to structural disorder. Table 3.1

shows the abundance of selected residues in the segments. The data in Table 3.1 indicates no significant difference in the number of charged residues, while the abundance of glycosylated and proline residues varies considerably. To choose representative segments for the entire structure, we aimed to select specific segments based on the abundance of glycosylated and proline residues.

Table 3.1. Total number of negative, positive, Proline, and glycosylated residues in each lubricin segment

	Total number of			
Sequence	Negative	Positive	Proline	Glycosylated
221-300	10	9	7	15
301-380	9	14	14	24
381-460	7	11	25	27
461-540	8	11	26	25
541-620	7	12	25	25
621-700	11	8	26	8
701-780	7	12	22	15
781-860	7	10	23	12
861-940	13	12	12	11
941-1020	7	13	5	23

We have chosen five segments of lubricin, specifically 301–380, 461–540, 621–700, 781–860, and 941–1020. These segments collectively represent the physicochemical characteristics of the protein, incorporating aspects like glycosylation, charged residues, and proline content. For a visual representation of the selected segments, please refer to Figure 3.1, where they are delineated within grey boxes.

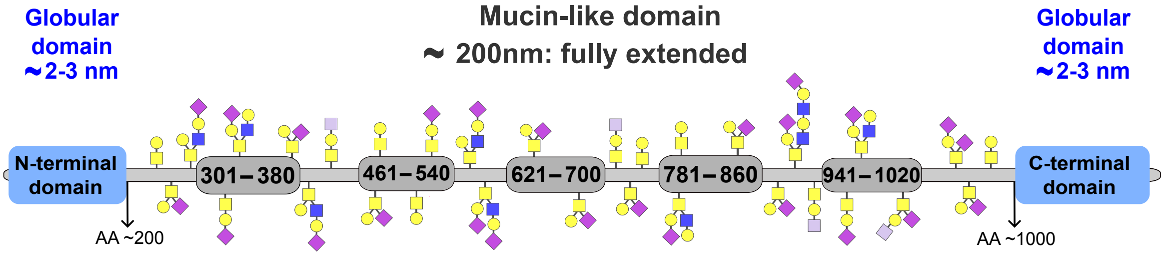

Figure 3.1. Selected segments of lubricin in the study. Five selected segments of lubricin, each comprising 80 amino acids, were chosen as representative segments capturing the essence of the entire mucin-like domain structure.

3.2.2 O-glycans of lubricin

As previously mentioned, the mucin-like domain of lubricin is heavily glycosylated, with O-glycans attaching to the hydroxyl groups of serine/threonine (Ser/Thr) side chains. In 2015, a study on synovial fluid from horses conducted by E. Svala and colleagues [86] proposed eleven distinct oligosaccharides, categorized into Core I (90%) and Core II (10%) structures. These glycans consist of sugars such as Gal (D-Galactose), GalNAc (N-Acetyl-D-Galactosamine), GlcNAc (N-Acetylglucosamine), Neu5Ac (N-Acetyl-Neuraminic Acid), and Neu5Gc (N-Glycolyl-Neuraminic acid). Notably, both Neu5Ac and Neu5Gc possess a negative charge. For a detailed representation, refer to Figure 1.5, which illustrates the intricate structure of these eleven glycans.

In this segment, for simplicity, we have opted for six specific O-glycan types from a total of eleven oligosaccharide variations. This choice considered the structure, composition, and net charge of the glycans. Figure 3.2A illustrates the categorization of these eleven glycans into six parts (a, b, c, d, e, and f), emphasizing their abundance (as a percentage) and net charge (in blue). Notably, glycans within categories c, d, e, and f share the same net charge and closely resemble each other in structure. Consequently, we selected only one type from each of these categories. The excluded ones are depicted with reduced opacity, while the chosen ones are displayed in a darker color. Figure 3.2B provides a visual representation of the selected glycans and their respective abundance.

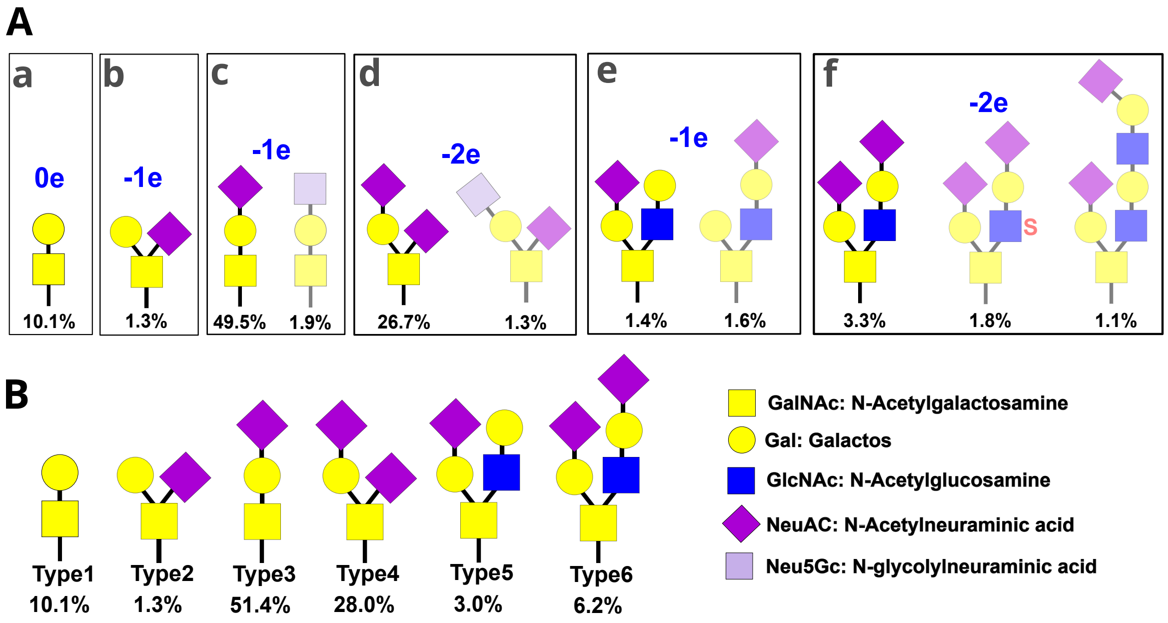

Figure 3.2. 11 Types of O-glycans of lubricin and Selected Ones for this Study. A. The 11 types of O-glaycans along with their charges and abundances. **B.** Six chosen types based on their structure and charge for this study.

From Ali Liaqat et al.'s study [69], we have information regarding the serine or threonine residues that are attached to glycans. Segments 301–380, 461–540, 621–700, 781–860, and 941–1020 contain 24, 25, 8, 12, and 23 glycosylated residues, respectively. However, the specific types of glycans binding to these glycosylated residues remain unclear, such as the exact O-glycan binding at each position. To address this uncertainty, we adopted a Monte Carlo sampling approach and leveraged the abundance information of each glycan (depicted in figure 3.2B) to present six different distributions of glycans for each fragment.

In our methodology, one of these six fragments was maintained without any glycan to serve as a non-glycosylated version. The other five fragments were randomly assigned each type of glycan based on the percentage contribution to the total O-glycans. Following this protocol, we generated one non-glycosylated and five distinct glycosylated versions for each lubricin segment. The specific sugars attached at each glycosylation site in the resulting 25 cases are detailed in Tables 3.2 – 3.6. The inclusion of Neu5Ac, which imparts a predominantly negative charge to most O-glycans, plays a substantial role in elevating the overall negative charge of the fragments post-glycan addition. This effect results in varied cumulative net-charge distributions, as depicted in Figure 3.3. For instance, fragments within the 301–380, 461–540, and 941–1020 regions exhibit a higher negative net charge compared to others, owing to their elevated number of glycosylated residues. Table 3.7 provides information on the net charge of all fragments.

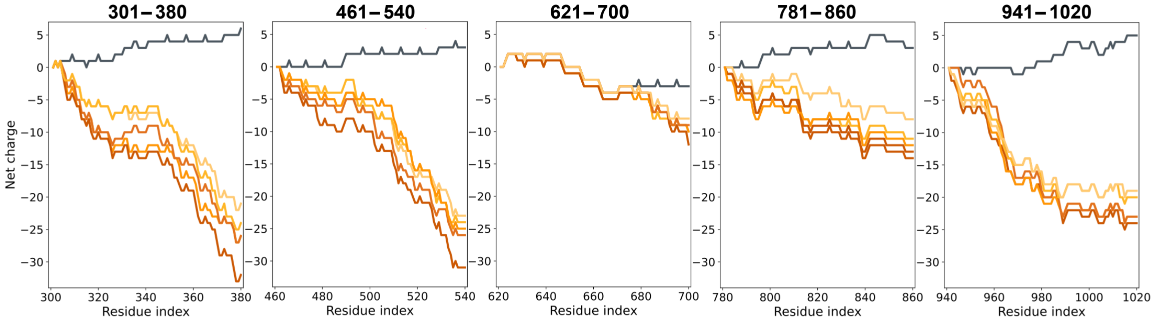

Figure 3.3. Net charge variation in lubricin fragments with added O-glycans. The non-glycosylated fragment of each segment is represented in grey, while colors ranging from yellow to red illustrate the impact of glycans on the negative charge of the fragments.

In this section, we presented six glycan distributions for each lubricin segment. However, for MD simulations, a comprehensive force field covering both proteins and glycans is essential. Due to the disordered nature of lubricin, we selected the amber99sb-star-ildnp force field [115,116] with TIP4P-D [165] water model, which lacked glycans parameters. To overcome this limitation, we integrated parameters from the glycans-specific GLYCAM06 force field [117]. The next section will provide details on the methodology used for this integration and our MD simulation approach.

3.3 Single-chain equilibrium MD simulations

As detailed in Section 2.1.1, the execution of an MD simulations requires an appropriate force field that encompasses all necessary parameters for the all materials within the system. Lubricin is a disordered and glycosylated protein [12, 69, 84, 166]. Given its disordered nature, we initially utilized the amber99sb-star-ildnp [115, 116] force field with TIP4P-D [165] water model, known for its efficacy in capturing the structural properties of intrinsically disordered proteins according to previous studies [167, 168]. Nevertheless, the original version of this force field did not support the glycans in the system. To address this gap, we modified the force field, introducing a new version to cover both proteins and glycans comprehensively. For the glycans, the GLYCAM06 force field supplied the required parameters. As this force field was in Amber format, the ACPYPE script [169, 170] was employed to convert the parameters into GROMACS format. In the subsequent step, these parameters were added to amber99sb-star-ildnp to support 12 different types of glycosylated versions of Ser and Thr. Through this process, we developed a unified force field that was employed for all our simulations.

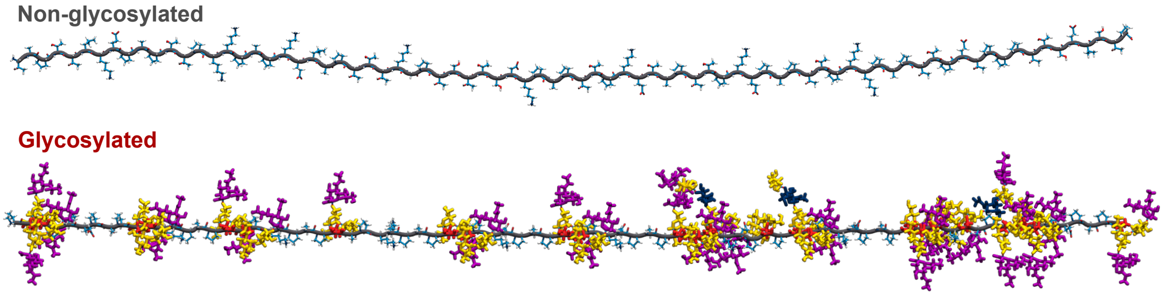

Figure 3.4. Initial structure of lubricin fragment. The upper structure depicts a non-glycosylated fragment, while the lower one shows the lubricin fragment with glycans in yellow, blue, and purple. Additionally, in glycosylated one, Ser and Thr residues are colored in red.

The initial fully elongated linear configuration of lubricin segments, representing a non-glycosylated fragment, was generated using the Avogadro software package [171]. Subsequently, O-glycans were added to the glycosylated residues of fragments based on their glycan types, as detailed in Tables 3.2 to 3.6, using the glycam.org server. Figure 3.4 illustrates examples of the non-glycosylated and glycosylated initial structures.

In this phase of the study, all MD simulations were executed utilizing the GROMACS software suite (version 2020) [172]. The initial structures were positioned within the center

Table 3.2. Glycosylated residues of the 301–380 segment and their five distinct O-glycan types across five different fragments.

Glycosylation site	O-glycan type				
	fragment1	fragment2	fragment3	fragment4	fragment5
304	Type4	Type4	Type2	Type4	Type4
305	Type3	Type3	Type3	Type4	Type3
306	Type5	Type4	Type4	Type4	Type1
310	Type1	Type4	Type4	Type3	Type3
312	Type4	Type4	Type4	Type4	Type6
316	Type3	Type1	Type3	Type4	Type3
317	Type1	Type4	Type4	Type3	Type1
324	Type3	Type4	Type3	Type3	Type3
325	Type3	Type3	Type3	Type4	Type3
332	Type4	Type4	Type1	Type6	Type3
345	Type3	Type3	Type4	Type3	Type4
346	Type3	Type1	Type3	Type3	Type2
351	Type3	Type3	Type4	Type4	Type3
353	Type3	Type2	Type1	Type3	Type4
354	Type3	Type4	Type3	Type3	Type5
360	Type3	Type3	Type3	Type4	Type3
361	Type3	Type4	Type3	Type4	Type4
362	Type5	Type4	Type4	Type3	Type3
367	Type3	Type3	Type6	Type3	Type4
369	Type4	Type1	Type3	Type4	Type4
370	Type4	Type3	Type4	Type4	Type3
373	Type3	Type3	Type5	Type3	Type6
376	Type1	Type3	Type5	Type4	Type3
377	Type4	Type4	Type4	Type4	Type3

Table 3.3. Glycosylated residues of the 461-540 segment and their five distinct O-glycan types across five different fragments.

Glycosylation site	O-glycan type				
	fragment1	fragment2	fragment3	fragment4	fragment5
462	Type3	Type4	Type4	Type3	Type3
463	Type4	Type2	Type4	Type3	Type3
470	Type1	Type3	Type3	Type1	Type1
471	Type1	Type3	Type3	Type3	Type4
477	Type3	Type3	Type4	Type1	Type1
478	Type2	Type1	Type3	Type3	Type3
485	Type2	Type3	Type2	Type1	Type3
493	Type6	Type3	Type3	Type4	Type3
494	Type3	Type3	Type3	Type4	Type1
501	Type3	Type4	Type4	Type2	Type1
502	Type1	Type4	Type3	Type5	Type3
509	Type4	Type4	Type4	Type3	Type4
510	Type4	Type3	Type3	Type4	Type3
511	Type3	Type4	Type6	Type5	Type4
515	Type4	Type3	Type3	Type3	Type4
517	Type3	Type1	Type5	Type4	Type6
518	Type3	Type4	Type3	Type3	Type3
525	Type1	Type6	Type3	Type4	Type6
526	Type3	Type3	Type4	Type3	Type3
527	Type3	Type3	Type3	Type4	Type3
529	Type3	Type1	Type6	Type3	Type4
532	Type3	Type6	Type4	Type4	Type4
533	Type3	Type3	Type3	Type3	Type3
534	Type4	Type3	Type4	Type1	Type3
540	Type4	Type4	Type4	Type4	Type3

Table 3.4. Glycosylated residues of the 621-700 segment and their five distinct O-glycan types across five different fragments.

Glycosylation site	O-glycan type				
	fragment1	fragment2	fragment3	fragment4	fragment5
627	Type1	Type3	Type2	Type1	Type1
676	Type3	Type3	Type1	Type3	Type3
683	Type4	Type3	Type3	Type3	Type1
684	Type4	Type3	Type6	Type4	Type4
691	Type1	Type3	Type4	Type3	Type4
692	Type3	Type3	Type5	Type3	Type1
699	Type3	Type3	Type4	Type1	Type1
700	Type3	Type3	Type3	Type4	Type4

Table 3.5. Glycosylated residues of the 781-860 segment and their five distinct O-glycan types across five different fragments.

Glycosylation site	O-glycan type				
	fragment1	fragment2	fragment3	fragment4	fragment5
781	Type5	Type3	Type1	Type3	Type6
784	Type3	Type3	Type3	Type3	Type5
785	Type3	Type4	Type3	Type3	Type4
792	Type3	Type3	Type4	Type4	Type4
793	Type3	Type4	Type1	Type4	Type3
805	Type3	Type3	Type1	Type3	Type3
811	Type6	Type3	Type3	Type1	Type3
812	Type4	Type4	Type3	Type6	Type1
813	Type3	Type6	Type3	Type4	Type3
829	Type3	Type3	Type4	Type3	Type1
837	Type3	Type3	Type1	Type4	Type4
838	Type3	Type4	Type4	Type3	Type6

Table 3.6. Glycosylated residues of the 941-1020 segment and their five distinct O-glycan types across five different fragments.

Glycosylation site	O-glycan type				
	fragment1	fragment2	fragment3	fragment4	fragment5
941	Type4	Type3	Type1	Type3	Type3
943	Type3	Type5	Type1	Type3	Type3
944	Type3	Type4	Type1	Type3	Type3
945	Type4	Type3	Type4	Type3	Type4
954	Type3	Type1	Type3	Type3	Type3
956	Type3	Type4	Type3	Type1	Type3
957	Type6	Type4	Type3	Type4	Type3
958	Type3	Type3	Type3	Type3	Type3
961	Type3	Type3	Type4	Type3	Type2
962	Type1	Type6	Type3	Type3	Type3
963	Type3	Type2	Type6	Type4	Type1
964	Type3	Type4	Type4	Type3	Type3
965	Type4	Type1	Type3	Type3	Type3
968	Type3	Type3	Type4	Type1	Type3
975	Type6	Type4	Type3	Type4	Type4
978	Type1	Type2	Type5	Type3	Type3
979	Type4	Type3	Type6	Type3	Type1
980	Type1	Type3	Type3	Type4	Type4
986	Type4	Type3	Type3	Type6	Type4
987	Type3	Type3	Type4	Type1	Type3
988	Type6	Type3	Type3	Type1	Type1
1012	Type4	Type3	Type3	Type4	Type1
1014	Type3	Type6	Type4	Type3	Type4

Table 3.7. Net charge comparison between glycosylated and non-glycosylated fragments for different lubricin segments.

Segment sequences	Non glycosylated	Glycosylated fragment1	fragment2	fragment3	fragment4	fragment5
301-380	5	-22	-27	-27	-33	-25
461-540	3	-25	-28	-33	-26	-26
621-700	-3	-11	-11	-13	-11	-10
781-860	3	-11	-14	-8	-13	-12
941-1020	6	-23	-22	-22	-19	-18

of a simulation box shaped like a dodecahedron and solvated with TIP4P-D [165] water molecules, accompanied by 150 mM NaCl ions. Additional ions were introduced to counterbalance the net charge of the fragments. The resultant systems were composed of approximately 0.36 to 0.92 million atoms, with detailed atom count information available in Table 3.8.

The systems underwent energy minimization employing a steepest descent algorithm until the maximum atomic force fell below 1000 kJmol^{-1} nm^{-1}. Subsequently, thermalization took place in the NVT ensemble at 310 K using the velocity rescaling thermostat [131] over 1 ns (with a coupling time of 0.1 ps). The solvent was then relaxed in the NPT ensemble at 1 atm with the Parinello–Rahman barostat [132] for 2 ns (using a coupling constant of 2.0 ps and reference compressibility of 4.5×10^{-5} bar^{-1}). During both equilibration stages, a harmonic force (with an elastic constant of 1000 kJmol^{-2}) was applied to restrain the position of the glycosylated-protein heavy atoms.

For production runs under the NPT ensemble, periodic boundary conditions were applied, and position restraints on heavy atoms were released. The velocity rescale thermostat and the Parrinello-Rahman algorithm were utilized during production runs to maintain a constant temperature and pressure. Each of the 30 different fragments underwent three replicas, each lasting 200 ns (see Table 3.8). Electrostatic interactions were taken into account in our simulations by the Particle Mesh Ewald algorithm [125]. Short-range interactions were simulated using a Lennard Jones potential, with the calculations limited at a cutoff distance of 1.0 nm. Constraints were applied to bonds involving hydrogen atoms within the fragments, and this was achieved through the utilization of the LINCS algorithm [173]. The Verlet Buffer was employed to manage neighbors, incorporating a tolerance of 0.005 kJmol^{-1} ps^{-1}. Updates to neighboring interactions were executed at regular intervals, precisely every 10 steps. The numerical integration of the equations of motion was accomplished using the Leap Frog algorithm, employing discrete time steps set at 2 fs.

3.4 Results

To explore the influence of O-glycans on the structural attributes of lubricin, we introduced a diverse set of 30 lubricin fragments, each characterized by distinct glycan distributions (Figure 3.5A). Each system underwent an equilibrium MD simulation, extending over a duration of 200 ns. Notably, for each fragment, three independent replicas were executed to ensure robust and reliable outcomes (Figure 3.5B). This process yielded equilibrium structural ensembles, focusing on individual fragments. Figure 3.5 illustrates the outlined protocol for this section.

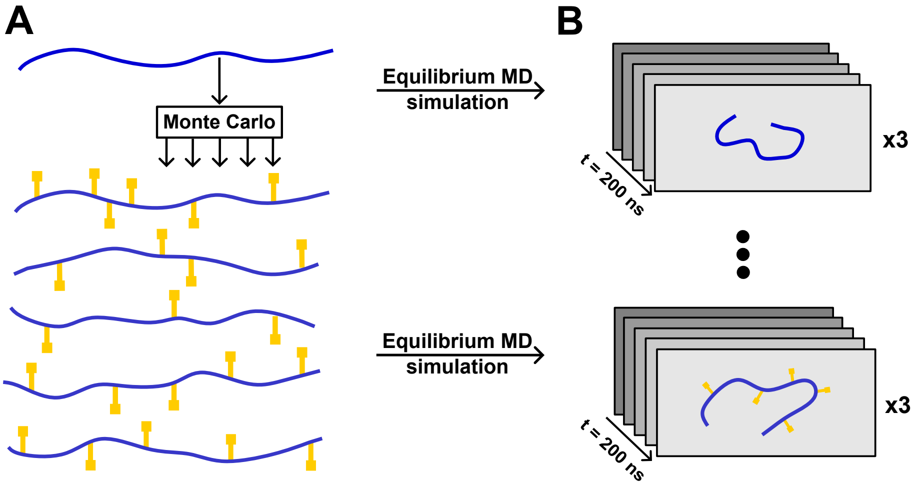

Figure 3.5. Molecular dynamics (MD) protocol to analyze O-glycan impact on lubricin structure. A. Five distinct glycosylation distributions for each of the five 80-amino acid long lubricin segments were introduced using Monte Carlo sampling. The non-glycosylated fragment is represented by the blue line, while the yellow symbols illustrate the O-glycans. **B.** Three independent equilibrium MD simulations, each lasting for 200 ns, were conducted for all fragments.

Following the completion of our simulations, we conducted an in-depth exploration of the acquired data. In the following, it will be shown how O-glycans influence the conformational dynamics of lubricin. This analysis focused on key parameters such as radius of gyration, end-to-end distance, solvent-accessible surface area (SASA), and persistence length.

Table 3.8. Details of the equilibrium molecular dynamics simulations for single-chain systems.

Peptide	Glycosylation	Atoms of peptide (no.)	Atoms of system (no.)	Simulation replicas (no.)	Simulation length (ns)
301-380	non-gly.	1223	0.58M	3	200
	fr.1	3443	0.66M		
	fr.2	3527	0.57M		
	fr.3	3671	0.55M		
	fr.4	3791	0.92M		
	fr.5	3599	0.72M		
461-540	non-gly.	1182	0.43M	3	200
	fr.1	3438	0.80M		
	fr.2	3594	0.78M		
	fr.3	3822	0.76M		
	fr.4	3522	0.40M		
	fr.5	3522	0.60M		
621-700	non-gly.	1182	0.56M	3	200
	fr.1	1854	0.53M		
	fr.2	1854	0.50M		
	fr.3	2022	0.52M		
	fr.4	1854	0.61M		
	fr.5	1818	0.41M		
781-860	non-gly.	1177	0.49M	3	200
	fr.1	2353	0.50M		
	fr.2	2413	0.36M		
	fr.3	2149	0.36M		
	fr.4	2377	0.40M		
	fr.5	2437	0.67M		
941-1020	non-gly.	1255	0.52M	3	200
	fr.1	3547	0.72M		
	fr.2	3511	0.53M		
	fr.3	3511	0.79M		
	fr.4	3307	0.71M		
	fr.5	3223	0.59M		

3.4.1 O-glycans expand the structure of lubricin fragments

For the initial investigation of the impact of glycans on the structure of lubricin, we calculated the radius of gyration (R_g), end-to-end (E2E) distance, and solvent-accessible surface area (SASA). These parameters provide insights into the size, shape, and structural characteristics of the fragments. Beyond their direct relevance to our study objectives, these observations also serve to assess the equilibration of the fragments over the 200 ns simulation period.

Radius of gyration

The initial metric under consideration was the R_g, a crucial parameter in assessing the compactness of the fragments. The influence of glycans on the overall structure of a fragment is illustrated in Figure 3.4A and B. The presence of bulky glycans contributes to alterations in the overall size of the fragments, attributed to the bulky side chains. To facilitate a meaningful comparison between non-glycosylated and different type of glycosylated fragments, our analysis focused exclusively on the backbone atoms for R_g calculation, omitting the contribution of glycans.

We calculated the R_g for each fragment at every time step, capturing the transition from fully extended to equilibrated structures. Figure 3.6 displays these dynamic changes over time. Notably, the fragments underwent a partial collapse within the initial tens of nanoseconds, ultimately adopting conformations within a range of approximately 2.3 to 3.9 nm.

Figure 3.7 displays the distribution of R_g for non-glycosylated and glycosylated fragments. For every 20 ns of simulation, the cumulative distribution of whole glycosylated fragments was calculated, and the same procedure was applied to non-glycosylated fragments. The figure reveals that the average R_g of glycosylated fragments is higher compared to non-glycosylated ones. This observation suggests that O-glycans possess the potential to expand the structure of lubricin fragments.

In addition to the R_g values, Figures 3.6 and 3.7 offer insights into whether the simulated time of 200 ns is adequate for the investigation. Both figures indicate that, on average, the fragments reached an equilibrium state after approximately 100 ns. Therefore, the first half of the simulation can be considered for achieving equilibrium, while the second half can be utilized for analyzing and studying the impact of O-glycans on lubricin structure.

Hence, to evaluate the influence of O-glycans on the structure of each lubricin fragment, the average R_g was calculated from the data in the second half of the simulations (from 100 to

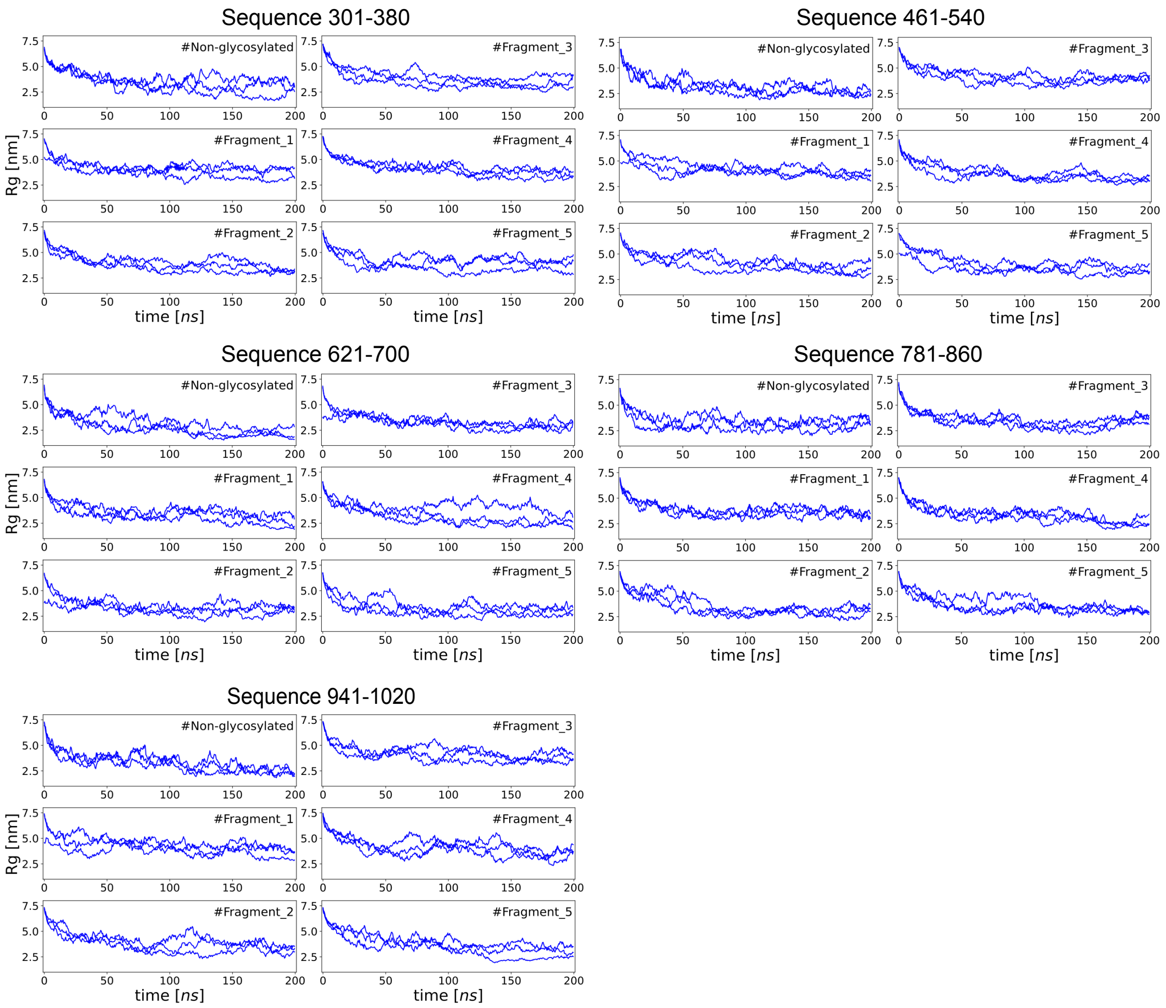

Figure 3.6. Evolution of the radius of gyration over time for lubricin fragments with different glycosylation levels. Equilibrium MD simulations were conducted for five different segments of lubricin, each with six distinct fragments featuring different distributions of O-glycans (including one non-glycosylated and five glycosylated chains). Three independent replicas were executed for each system, each extending over 200 ns.

200 ns). Figure 3.8 illustrates the relationship between the number of glycosylated residues, the net charge of the fragment, and the average R_g.

For a more in-depth exploration of the effect of O-glycans, we compared the equilibrium R_g of fragments to two essential parameters: their total number of glycosylated residues and their net charge, as depicted in Figure 3.8. Figure 3.8A illustrates that as more glycans are added to the lubricin fragments, their conformation becomes less compact, indicating a more extended structure. This direct relationship is attributed to the bulky side chains of O-glycans and their negative charge (see Figure 3.2). In the subsequent analysis of net charge, Figure 3.8B reveals the correlation between the net charge of lubricin fragments and their level of compactness. This figure shows that an increase in charge corresponds to an

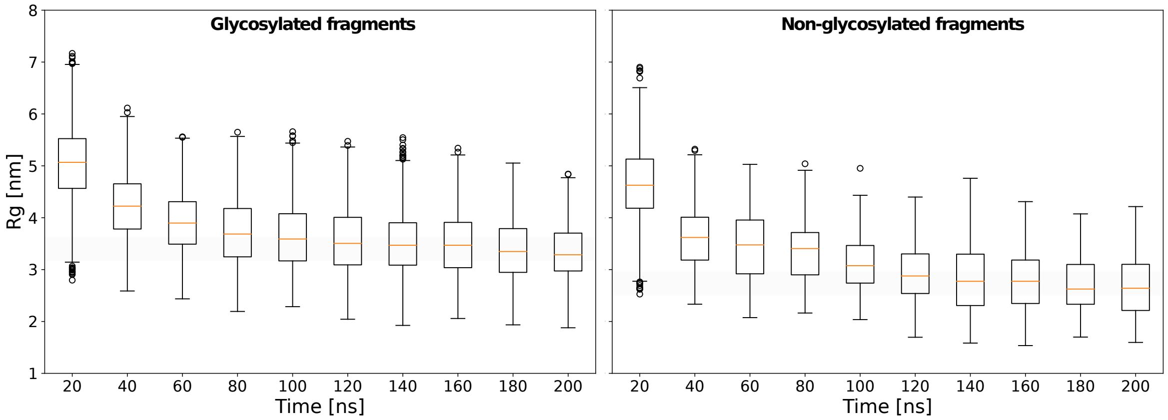

Figure 3.7. Radius of gyration distribution for glycosylated and non-glycosylated fragments over time. The boxplots illustrate the distribution of R_g values throughout the simulations. For each 20 ns interval of the simulation, the cumulative distribution of entire fragments was computed. Both glycosylated and non-glycosylated fragments show a trend towards equilibrium after around 100 ns, with glycosylated fragments exhibiting a higher average, indicating the expansion of their structural dynamics.

increase in the size of the fragments, following a V-shaped pattern. This pattern aligns with observations in other intrinsically disordered proteins (IDPs) [167, 168].

In Figure 3.8B, a linear regression analysis of the data, expressed as $R_g = a_q|q| + b_q$, resulted in a slope of $a_q = 0.036\,\text{nm/e}$, a value approximately one-fourth smaller than that reported in previous studies on other IDPs phosphorylation (0.048 nm/e) [167, 168]. The increased ionic strength (150 mM) employed in this study, compared to the 100 mM used in [167, 168], may account for the observed weaker response of r_g to changes in charge. Another possible factor could be the larger size of O-glycans compared to phosphoryl (PO_3) groups. This size effect might reduce the strength of the impact change on the conformational changes of the fragments. The intercept, $b_q = 2.72\,\text{nm}$, closely aligns with the estimate for a random coil adapted to IDPs [174] ($R_c = R_0 \cdot N^{0.588} = 2.61\,\text{nm}$, where $R_0 = 0.19\,\text{nm}$, and N is the number of amino acids, specifically $N = 80$ in our case). Previous studies reported a significantly smaller intercept (1.7 nm) [167, 168]. The intercept serves as an indicator of the overall size of the chains under neutral conditions. These results indicate that conformational changes are related to both the bulky size of sugars and their charge. In this scenario, the bulky sugars, potentially acting as spacers, tend to promote more expanded conformations of the lubricin backbone fragments compared to other IDP fragments with comparably smaller standard [168] or phosphorylated [167] amino acid side chains. For further assessment, we also calculated the end-to-end distance and SASA for all fragments. Both parameters exhibit a similar trend as the radius of gyration. However, SASA provides a more detailed overview of the bulky side chains of glycans and their im-

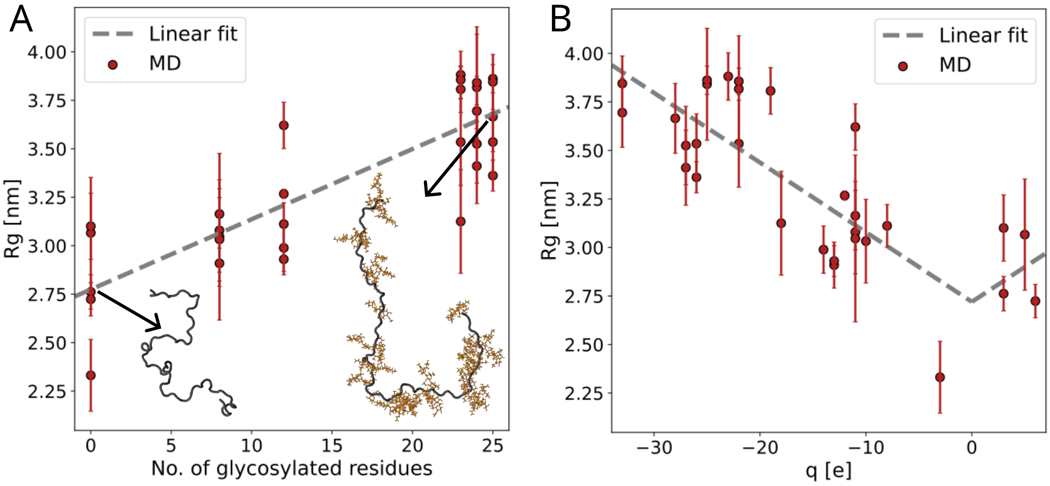

Figure 3.8. Impact of O-glycans on the extension of lubricin fragments. Radius of gyration R$_g$ as a function of **(A)** the number of glycosylated residues and **(B)** the net charge of the fragment q. Each data point represents values obtained from three independent MD simulations (average ± s.e., $n = 3$). Dashed lines represent linear regressions of the data, described by $R_g = a_N N + b_N$ (A) and $R_g = a_q|q| + b_q$ (B), with fitting parameters $a_N =$ 0.036 [nm], $b_N = 2.77$ [nm], $a_q = 0.036$ [nm/e], and $b_q = 2.72$ [nm]. Cartoons illustrate the conformation of a glycosylated and a nonglycosylated fragment (protein: grey, sugars: orange).

pact on the structural changes. In the next section, more information about these parameters will be provided.

E2E distance and SASA

In Figure 3.9, the results of the E2E distance of lubricin fragments are presented, which align well with the findings for R$_g$. Figure 3.9A displays the data distribution of E2E distance for each 20 ns interval of glycosylated and nonglycosylated fragments. This figure shows that the fragments undergo a collapsing process during the initial 100 ns, reaching a roughly constant value thereafter, indicative of an equilibrated structure. Moreover, on average, glycosylated fragments exhibit a higher E2E distance than nonglycosylated ones. For a more detailed analysis, the average E2E distance of each fragment was calculated from the second half of the simulation. Figures 3.9B left and right illustrate the correlation between E2E distance and the number of glycosylated residues, as well as the absolute charge of fragments. These results closely align with the R$_g$ findings, indicating that O-glycans can indeed expand the conformation of lubricin fragments.

In the subsequent analysis to highlight the impact of the bulky side chains of glycans, we computed the SASA values for the fragments. In contrast to the radius of gyration, we

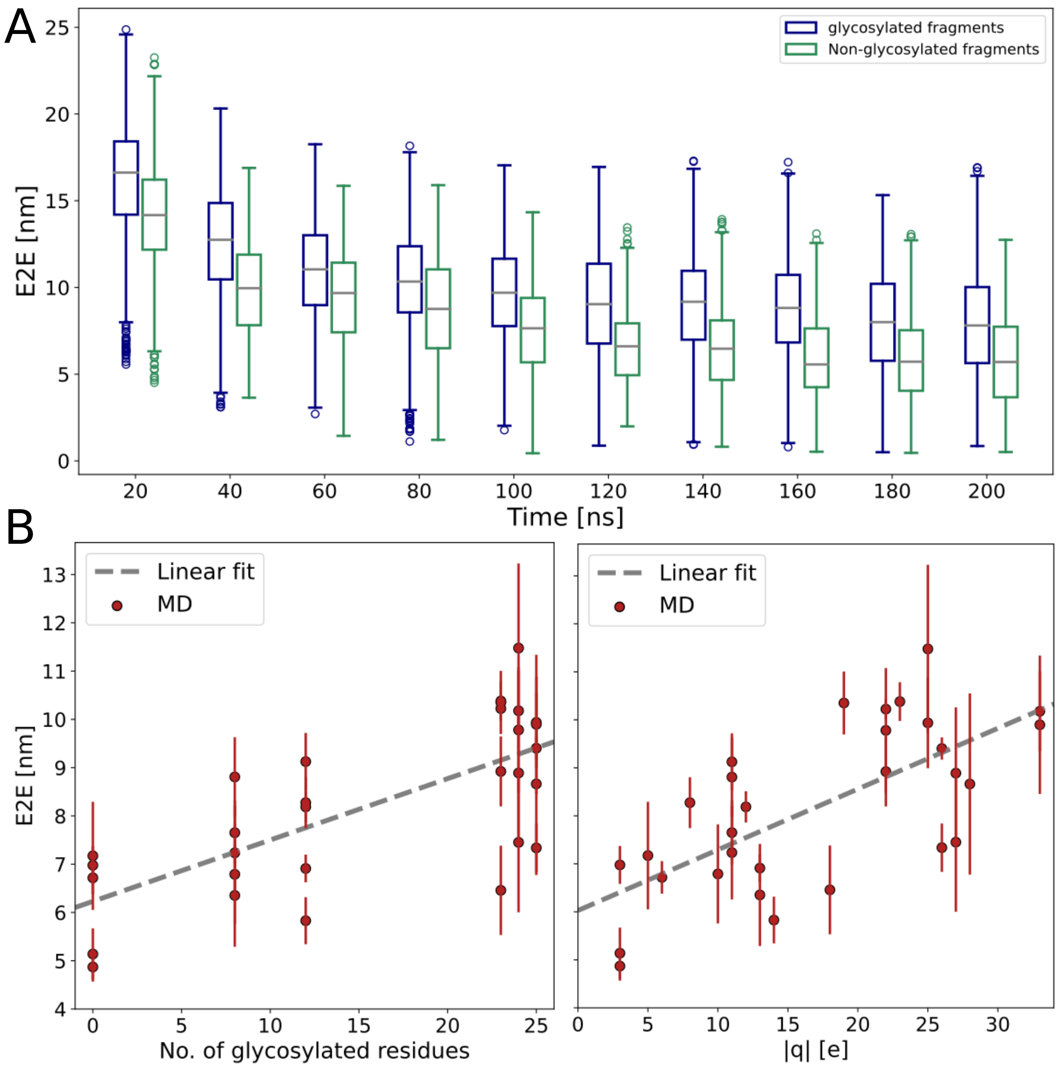

Figure 3.9. End-to-end distance analysis for glycosylated and non-glycosylated lubricin fragments. **A.** Data distribution of E2E distance for each 20 ns interval of glycosylated and non-glycosylated fragments. The collapsing process during the initial 100 ns is followed by a stable equilibrated structure. Non-glycosylated fragments, on average, exhibit a shorter E2E distance compared to glycosylated fragments. **B.** Correlation between E2E distance and the number of glycosylated residues (left) and absolute charge of fragments (right). Each data point represents values obtained from three independent MD simulations (average ± s.e., $n = 3$).

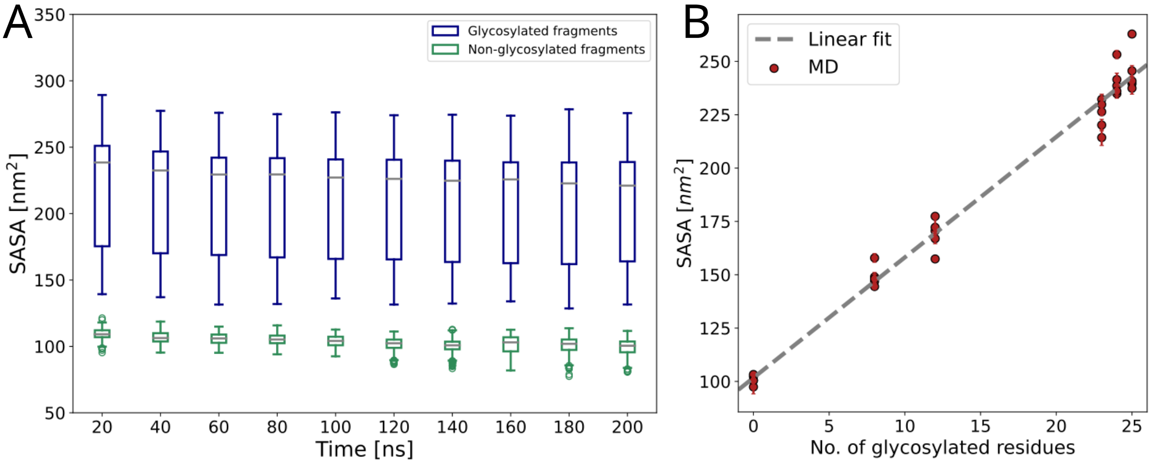

Figure 3.10. SASA analysis for glycosylated and non-glycosylated lubricin fragments. **A.** Distribution of SASA values for both glycosylated (in green) and non-glycosylated (in blue) fragments. **B.** Correlation between SASA and the number of glycosylated residues. Each data point represents values obtained from three independent MD simulations (average $\pm$ s.e., $n = 3$).

considered the entire structure of lubricin fragments, including both the backbone and its O-glycans. The calculations were focused on the second half of the simulation, where the fragments reached an equilibrated state. Figure 3.10 illustrates the obtained SASA results.

The distribution of SASA results for glycosylated and non-glycosylated fragments is depicted in Figure 3.10A. Notably, a significant disparity exists between the non-glycosylated and glycosylated fragments, underscoring the substantial impact of O-glycans. This disparity reflects the presence of O-glycans, signifying their role as bulky side chains. For a more detailed exploration of the influence of O-glycans, Figure 3.10B presents the correlation between SASA and the number of glycosylated residues. This figure illustrates a clear trend where an increase in the number of sugar residues leads to a dramatic rise in the SASA of lubricin fragments. This observation is well-matched with the results from the experimental study conducted by Brito and colleagues [175]. They demonstrated that the presence of glycans leads to a higher SASA value, and also decreases the aggregation propensity of glycosylated chains. These findings provide additional insights into the significant effect of O-glycans on the surface accessibility of the fragments.

3.4.2 O-glycans enhance the stiffness of lubricin fragments

Another aspect that we were interested in exploring was how glycans could alter the stiffness of lubricin. To evaluate this property, we calculated the persistence length of lubricin fragments, as detailed in Section 2.3.2. Figure 3.11 presents the correlation between the stiffness

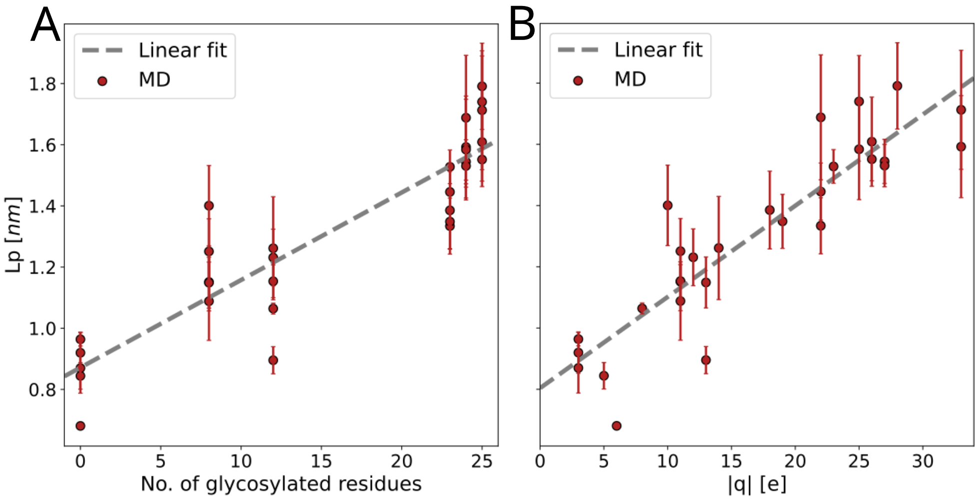

Figure 3.11. Persistence length of lubricn fragments. Persistence length as a function of the number of glycosylated residues **(A)** and the absolute net charge of fragments **(B)**. Each data point represents values obtained from three independent MD simulations (average ± s.e., $n = 3$).

of lubricin fragments and both the number of glycosylated residues (3.11A) and the absolute net charge of the fragments (3.11B). The results reveal a clear relationship wherein the number of glycans and the net charge of fragments can influence the stiffness of lubricin. An increase in the number of glycans corresponds to an increase in the SASA value (indicative of steric interaction) and also an increase in the net charge of the fragments. Thus, glycans, through their steric and electrostatic interactions, have the potential to impact the stiffness of lubricin. In summary, the presence of O-glycans contributes to increased stiffness in the structure of lubricin.

3.5 Discussion

This chapter aimed to understand how O-glycans impact the structural properties of lubricin. The protein features a disordered region extensively adorned with sugars (Figure 1.4). To delve into this, we employed all-atom equilibrium MD simulations, combining force fields for disordered proteins and glycans. In modeling the mucin-like domain of lubricin, we strategically selected five segments. For each segment, six different fragments were presented, incorporating various O-glycan types (Figure 3.1). Subsequently, we conducted three independent simulations for each of the 30 distinct lubricin fragments, each spanning 200 ns. This comprehensive approach allowed us to thoroughly explore the dynamics and interactions within the disordered region of lubricin under the influence of O-glycans. To gain a deeper understanding of the glycan effect, we calculated R_g, E2E, SASA, and persistence length after conducting simulations. The results revealed that O-glycans can both expand the conformational structure and increase its stiffness. These changes arise from the strong steric and electrostatic interactions of the bulky side chains of the glycans. These simulations provide valuable insights into the structural changes induced by O-glycans.

The dynamic conformations of IDPs are intricately regulated by various post-translational modifications, including phosphorylation [167,168,176], methylation [177], and glycosylation [175,178]. One distinguishing characteristic of glycosylation, setting it apart from other post-translational modifications, lies in the size of the attached part. In contrast to phosphorylation or methylation, where relatively small side chains are added to protein residues, glycosylation involves the attachment of large and bulky glycans. This distinctive feature becomes particularly evident in proteins such as lubricin, which exhibits a complex and dense glycosylation pattern (Figure 3.8B). As a result, glycosylation induced more extended conformations in lubricin fragments compared to those observed in other IDPs with a similar sequence length [167, 168] (Figure 3.8B). However, the size of the fragments exhibits a linear increase with their net charge, resembling a V-pattern [167, 168]. This underscores the significant influence that electrostatics wield in dictating the conformation of IDPs and highlights the capacity of post-translational modifications to finely tune these conformations.

In addition to exploring electrostatic interactions, we investigated the influence of bulky side chains by analyzing SASA (see section 3.4.1). The alterations in conformational properties observed in our study are intricately linked to both steric and electrostatic effects. However, in our current investigation, we are unable to discern between these two factors. To address this, a potential solution could involve neutralizing the glycans and conducting additional MD simulations for glycosylated fragments of lubricin with uncharged glycans. This proposed protocol aims to specifically unveil the impact of steric interactions, providing valuable insights into the distinct contributions of steric and electrostatic factors on the structural properties of lubricin.

In our investigation, we opted for smaller fragments of lubricin, aiming to choose representative segments that capture the essence of the entire structure. However, it is crucial to acknowledge that lubricin possesses two globular end domains, potentially exerting an influence on its overall structural properties. To comprehensively consider the complete structure of lubricin, a higher resolution is required (lubricin has approximately 1400 amino acids and 185 O-glycans [69]). The Martini3 force field [179] has demonstrated its capability to replicate the properties of intrinsically disordered proteins. Nevertheless, when it comes to glycans, there is a lack of appropriate and well-defined parameters. Therefore, to encompass the entirety of lubricin's structure, it becomes imperative to establish a suitable model and parameters specifically tailored for the glycans. This approach ensures a more accurate representation of the interplay between lubricin and glycans in the context of structural dynamics. In addition to martini, coarse-grained models employing one bead per amino acid, like CALVADOS [180, 181] have been quite successfull for IDPs. If it is only about the equilibrium structural properties of glycosylated IDPs, these models could be considered.

In addition to the intriguing results obtained from equilibrium MD simulations, we now have 30 distinct equilibrated structures of lubricin fragments. These equilibrated structures will be used to study and understand the viscoelastic properties of lubricin. In the upcoming chapter, our focus will be on investigating the viscosity of lubricin fragments and the influence of glycans on these properties. Two different methods were employed in MD simulations—Green-Kubo method and shear-driven non-equilibrium MD simulation—to calculate both zero and non-shear viscosities.

4

O-glycans and lubricin's viscosity and shear thinning

In the previous chapter, we explored the outcomes concerning individual lubricin fragments and their structural properties in relation to O-glycans. These findings revealed that O-glycans have the capacity to expand lubricin fragments and increase their stiffness, owing to their steric and electrostatic interactions. In this chapter, our focus shifts to a multi-chain system of lubricin fragments. Our primary objective is to determine the viscosity of lubricin and understand the molecular factors influencing lubricin's role in minimizing viscosity within synovial joints, as well as how O-glycans contribute to the viscosity mediated by lubricin. To attain this goal, we employed both equilibrium and shear-driven non-equilibrium MD simulations to calculate zero shear viscosity and shear viscosities across a range of shear rates [6, 145, 182].

In the upcoming sections, we will explore the equilibrium molecular dynamics (EMD) and non-equilibrium molecular dynamics (NEMD) simulation methods utilized for viscosity determination. Subsequently, we will present simulation results that provide insight into the influence of O-glycans on the viscosity of lubricin. Finally, we will discuss the outcomes, shedding light on the response of the multi-chain system to external shear flow.

4.1 Introduction

Shear-thinning is a rheological property commonly observed in various non-Newtonian and complex fluids like polymers, colloids, and certain biological substances. In simple terms, it refers to the phenomenon where the viscosity of a solution decreases as the applied shear rates increase [183]. This behavior is widespread in both industrial and biological systems, ranging from molecular inks and Nafion to cellulose hydrogels, mucus, and blood [5–7,184]. Notably, synovial fluid, which lubricates joints, exhibits shear-thinning behavior [42, 185, 186].

The flow properties and shear-thinning characteristics of synovial fluid are intricately linked to its diverse components, including aggrecan, hyaluronic acid, lubricin, and others (see Table 1.1 for details). These components collaborate with proteins on the cartilage surface, such as collagens, proteoglycans, and noncollagenous proteins, facilitating the smooth movement of bones against each other [187–189]. To optimize lubrication and provide excellent boundary lubrication, during bone movement, a reduction in viscosity is essential. Lubricin, an intricate glycosylated and disordered protein, plays a pivotal role in friction reduction. While its globular end domains easily attach to the cartilage surface, the central part (mucin-like domain) extends into the synovial fluid, predominantly governing this lubrication process. However, the detailed mechanisms at the atomistic level, especially regarding the function of the glycosylated central part of lubricin, remain unclear (see section 1.3 for more details).

To address this question, we investigated the impact of O-glycans on the viscoelastic behavior of lubricin using the Green-Kubo [147, 148] and box deformation [146, 190, 191] methods in EMD and NEMD simulations.

Our findings reveal that glycosylation expands lubricin fragments and reduces intermolecular clustering. In terms of viscosity, O-glycans alter the viscoelastic behavior of lubricin. They enable lubricin to reduce the viscosity of the fluid and also exhibit a shear-thinning response to external shear stress. This implies that glycosylated lubricin exhibits lower overall viscosities when compared at similar mass densities, and furthermore, these viscosities exhibit a less pronounced decrease under shear conditions. By elucidating the structural and rheological characteristics of glycosylated lubricin, this study contributes to a deeper understanding of synovial joint dynamics.

4.2 Methods

To achieve our objective, we utilized both equilibrium and non-equilibrium MD simulations. EMD simulations provide a comprehensive overview of the system in equilibrium. Specifically, by utilizing the Green-Kubo method, we can calculate the zero shear viscosity. To understand the behavior of lubricin fragments under shear flow, we utilized the box deformation method to impose a planar Couette flow. This method allows us to obtain viscosities under different shear rates. Both methods will be elucidated in detail in the upcoming sections.

4.2.1 Multi-chain equilibrium MD simulation

To establish how lubricin glycosylation affects the medium viscosity, we considered five distinct lubricin systems with varying levels of glycosylation, molar concentration, and mass density. We considered a system containing 30 non-glycosylated fragments (w.o. ρ=66.5 kg/m^3). To provide a comparative extreme state, we introduced another system with 30 highly glycosylated fragments of lubricin (w.+ ρ=230.6 kg/m^3). Both of these systems had the same molar concentration (N=30). For comparison, we included a third system with 101 non-glycosylated fragments (w.o. ρ=226.6 kg/m^3) to match the mass density of the glycosylated system. This arrangement allowed us to compare a glycosylated system to two non-glycosylated systems with the same mass and molar density. To create a system that more closely resembles reality, we opted for a configuration comprising 30 randomly glycosylated peptides (w. ρ=160. kg/m^3). Additionally, to specifically study the influence of glycans, we established a system with 73 non-glycosylated peptides, ensuring it had the same mass density (w.o. ρ=158.6 kg/m^3). Further simulation details can be found in Table 4.1.

Here, all MD simulations were conducted using the GROMACS software suite (version 2020) [172]. For each system, peptides (obtained from the single-chain MD simulations) were placed at the center of a simulation box, which had a cubic shape with dimensions of 25 × 16 × 16 nm. The system was solvated with TIP4P-D [165] water molecules and supplemented with 150 mM NaCl ions. Additional ions were added to neutralize the net charge of the systems. The resulting systems consisted of approximately 0.8 million atoms, and detailed atom count information is available in Table 4.1. The systems underwent energy minimization employing a steepest descent algorithm until the maximum atomic force fell below 1000 kJmol^{-1} nm^{-1}. Subsequently, thermalization took place in the NVT ensemble at 310 K using the velocity rescaling thermostat [131] over 4 ns (with a coupling time of 0.1 ps). The solvent was then relaxed in the NPT ensemble at 1 atm with

Table 4.1. Details of the equilibrium molecular dynamics for multi-chain systems.

Glycosylation level	Chains	Number of chains	ρ [kg/m^3]	Atoms of system system (no.)	Simulation replicas (no.)	Simulation length (ns)
Non	non-gly. of 301-380	30	66.5	0.83M	4	200
Non	random from all non-gly. fragments	73	158.6	0.82M	4	200
Non	non-gly. of 301-380	101	226.6	0.81M	4	200
Medium	random from all fragments	30	160.5	0.82M	4	200
Highly	fragment4 of 301-380	30	230.6	0.81M	4	200

the Parinello–Rahman barostat [132] for 4 ns (using a coupling constant of 2.0 ps and reference compressibility of 4.5×10^{-5} bar^{-1}). During both equilibration stages, a harmonic force (with an elastic constant of 1000 kJmol^{-2}) was applied to restrain the position of the glycoscylated and non-glycosylated protein heavy atoms.

For production runs under the NPT ensemble, periodic boundary conditions were applied, and position restraints on heavy atoms were released. The velocity rescale thermostat and the Parrinello-Rahman algorithm were utilized during production runs to maintain a constant temperature and pressure. Each system underwent four replicas, each lasting 200 ns (see Table 4.1). Electrostatic interactions were taken into account in our simulations by the Particle Mesh Ewald algorithm [125]. Short-range interactions were simulated using a Lennard Jones potential, with the calculations limited at a cutoff distance of 1.0 nm. Constraints were applied to bonds involving hydrogen atoms within the fragments, and this was achieved through the utilization of the LINCS algorithm [173]. The Verlet Buffer was employed to manage neighbors, incorporating a tolerance of 0.005 kJmol^{-1} ps^{-1}. Updates to neighboring interactions were executed at regular intervals, precisely every 10 steps. The numerical integration of the equations of motion was accomplished using the Leap Frog algorithm, employing discrete time steps set at 2 fs.

After the equilibration phase, the final conformations were chosen for subsequent viscosity calculations. Figure 4.1 A and B schematically illustrate this process for both non-

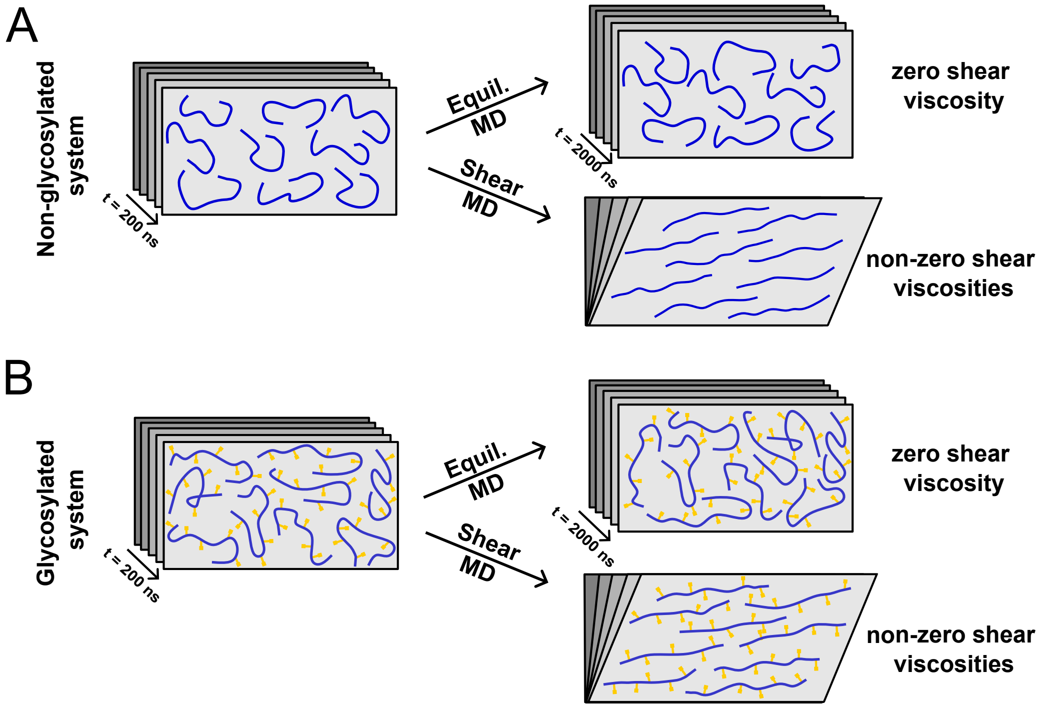

Figure 4.1. Equilibrium and non-equilibrium MD simulation protocol for multi-chain systems. The protocol was applied to both glycosylated and non-glycosylated fragments. Initially, a 200 ns MD simulation was conducted under equilibrium conditions. Subsequently, an extensive equilibration simulation was performed to obtain zero shear viscosity. Finally, the box deformation method was employed to impose Couette flow and determine viscosities under shear flow. Here, the non-glycosylated fragment is depicted by the blue line, while the yellow symbols illustrate the O-glycans.

glycosylated and glycosylated systems. In each case, an initial 200 ns simulation was conducted, followed by two separate simulations to determine zero shear viscosity and viscosities under shear flow. we conducted two equilibrium MD simulations, each lasting 2000 ns, for every multi-chain system to evaluate zero shear viscosity. Additionally, we performed shear-driven non-equilibrium MD simulations to assess viscosities under shear flow. Both parts will be elaborated upon in the following sections.

4.2.2 Shear-driven non-equilibrium MD simulation

The shear-driven non-equilibrium MD simulations followed the equilibrium MD simulations for multi-chain systems. These simulations aimed to investigate the impact of shear flow and glycosylation on the viscosity of lubricin fragments assemblies, utilizing the shear deformation method.

In the box deformation method, the imposition of a planar Couette flow on the sys-

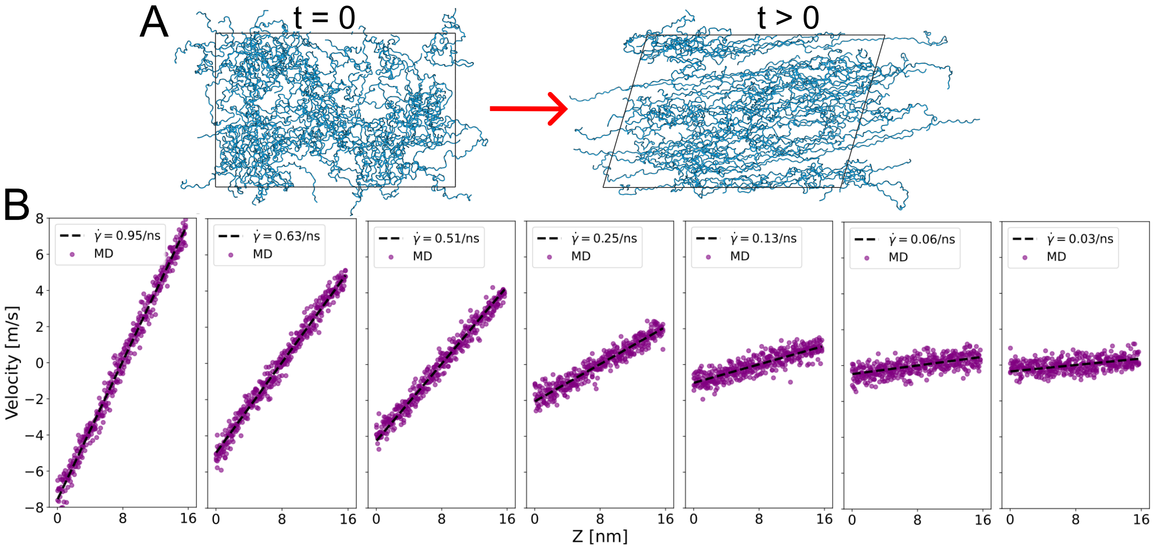

Figure 4.2. Box deformation method for viscosity calculations under shear flows. A. To induce Couette flow and assess shear viscosity, we employed the box deformation method, deforming the simulation box at a shear rate $\dot{\gamma}$. Initial system configurations from equilibrium MD simulations (t = 0) and subsequent snapshots at t > 0 are illustrated, with lubricin fragments highlighted in blue. **B.** Velocity profiles for each shear rate were derived through box deformation method (in purple color). The dashed black line represents the linear fit. The output shear rates (u/h) are indicated in the legend. Additionally, the anticipated shear rates are 0.94, 0.625, 0.50, 0.25, 0.125, 0.06, and 0.03 ns^{-1}.

tem [192] enables us to determine the viscosity of the system. In this approach, the simulation box experiences deformation as the upper wall is laterally displaced at a constant speed u along the x-axis (Figure 2.5A), where the upper wall is located at $z = h$, and h denotes the box size along the z-axis. Additionally, to generate a planar Couette flow, Lees-Edwards periodic boundary conditions [134,135] are employed. Figure 4.2 provides an overview of this method, illustrating the deformation of the simulation box in a specific direction (Figure 4.2 A) and inducing a linear shear flow (Figure 4.2 B).

In the context of Couette flow, the shear rate ($\dot{\gamma}$) is determined by both the velocity of the moving wall and the separation between the walls, as described by Equation 4.1. Subsequently, the shear viscosity (η) is calculated by examining the relationship between the component of the pressure tensor $\langle P_{xz} \rangle$ and the shear rate of the fluid, as expressed in Equation 4.2.

$$\dot{\gamma} = \frac{u}{h}. \tag{4.1}$$

$$\eta = \frac{\langle P_{xz} \rangle}{\dot{\gamma}}. \tag{4.2}$$

During this phase of the investigation, we employed the GROMACS-2023-dev version,

Table 4.2. Parameters of shear-driven non-equilibrium MD simulation

Deforming speeds (nm/ns)	Shear rates (ns^{-1})	Simulation length (ns)
0.5	0.03	600
1.0	0.06	400
2.0	0.13	200
4.0	0.25	100
8.0	0.51	100
10.0	0.63	50
15.0	0.95	50

which integrates the Lees-Edwards boundary condition, to compute the viscosities of lubricin fragments under shear flow using the box deformation method. Since the initial systems were derived from the last conformation of the previous simulation, which was in equilibrium MD, we only repeated the equilibration step in the NPT ensemble to be sure our systems are fully equilibrated. The solvent underwent relaxation in the NPT ensemble at 1 atm with the Parinello–Rahman barostat [132] for 4 ns, using a coupling constant of 2.0 ps and a reference compressibility of 4.5×10^{-5} bar^{-1}. Throughout both equilibration stages, a harmonic force (with an elastic constant of 1000 kJmol^{-2}) was applied to restrain the position of all heavy atoms.

In the box deformation method, the use of the NPT ensemble is not feasible. Instead, we employed the NVT ensemble for production runs, applying the Lees-Edwards boundary condition and releasing position restraints on heavy atoms. During production runs, the velocity rescale thermostat was exclusively used to ensure a constant temperature. Electrostatic interactions were taken into account in our simulations by the Particle Mesh Ewald algorithm [125]. Short-range interactions were simulated using a Lennard Jones potential, with the calculations limited at a cutoff distance of 1.0 nm. Constraints were applied to bonds involving hydrogen atoms within the fragments, and this was achieved through the utilization of the LINCS algorithm [173]. The Verlet Buffer was employed to manage neighbors, incorporating a tolerance of 0.005 kJmol^{-1} ps^{-1}. Updates to neighboring interactions were executed at regular intervals, precisely every 10 steps. The numerical integration of the equations of motion was accomplished using the Leap Frog algorithm, employing discrete time steps set at 2 fs.

For each of the five different systems outlined in Table 4.1, we subjected them to seven distinct shear rates, with three replicas conducted for each shear rate. The duration of the simulations ranged from 50 to 600 ns, varying based on the applied shear rate. To achieve these shear rates, we employed seven distinct deforming speeds: $u = 0.5$, 1.0, 2.0, 4.0, 8.0, 10.0, and 15.0 nm/ns. Additionally, the h value was maintained at 15.9 ± 0.1 nm across all systems. Further details can be found in Table 4.2. Upon completion of the simulations, we obtained the pressure elements P_{xz} using the GROMACS tool, gmx energy. The ensemble average pressure elements $\langle P_{xz} \rangle$ were computed using Python. Subsequently, employing Equation 4.2, we calculated the viscosity for each system.

4.2.3 Zero shear viscosity

In NEMD simulations, calculating viscosities at extremely low shear rates is highly challenging due to the significant time investment. To bridge the gap between experimental studies (typically below $10^5 s^{-1}$) and our simulations, we focused on determining the zero shear viscosity (η) of lubricin. This parameter represents a fluid's resistance to flow in the absence of applied shear stress, providing insights into the impact of O-glycans on viscoelastic behavior, particularly at low shear rates.

Here, we computed the zero shear viscosity (η) using equilibrium MD simulations and the Green-Kubo method [147–149]. This method necessitates an equilibrium MD simulation under periodic boundary conditions, where viscosity is determined by analyzing time-dependent stress fluctuations. Mathematically, viscosity is obtained by integrating the autocorrelation function (ACF) of the pressure tensor components over time, as expressed in the following equation [145, 149]:

$$\eta = \lim_{t \to \infty} \eta(t) = \frac{V}{k_{\mathrm{B}} T} \lim_{t \to \infty} \int_0^t \langle P_{\alpha\beta}(t) \cdot P_{\alpha\beta}(0) \rangle \, \mathrm{dt}, \tag{4.3}$$

where, η, V, k_{B}, T, $P_{\alpha\beta}$, and t represent the viscosity, system volume, Boltzmann constant, temperature, components of the pressure tensor, and time, respectively. The term $\langle P_{\alpha\beta}(t) \cdot P_{\alpha\beta}(0) \rangle$ measures the ensemble average of the autocorrelation function (ACF) of the pressure tensor elements, where α and β can take values x, y, or z. In our calculation for better accuracy, we considered all six independent diagonal and off-diagonal components of the pressure tensor ($P_{\alpha\beta} = P_{xy}$, P_{xz}, P_{yz}, $(P_{xx} - P_{yy})/2$, $(P_{xx} - P_{zz})/2$, and $(P_{yy} - P_{zz})/2$) [150, 151]. Figure 4.3 illustrates these three steps.

Upon analyzing the results of non-zero viscosities, the systems with 73 and 101 non-glycosylated chains are expected to have a high zero shear viscosity. Given that the Green-Kubo method is particularly effective in fluids with relatively low viscosity, typically below 20 mPa·s [144, 193], we decided to exclude non-glycosylated systems with medium and high mass density from this calculation, as their expected viscosity was anticipated to be high. Therefore, we selected three specific systems: 30 non-glycosylated (w.o.), 30 medium

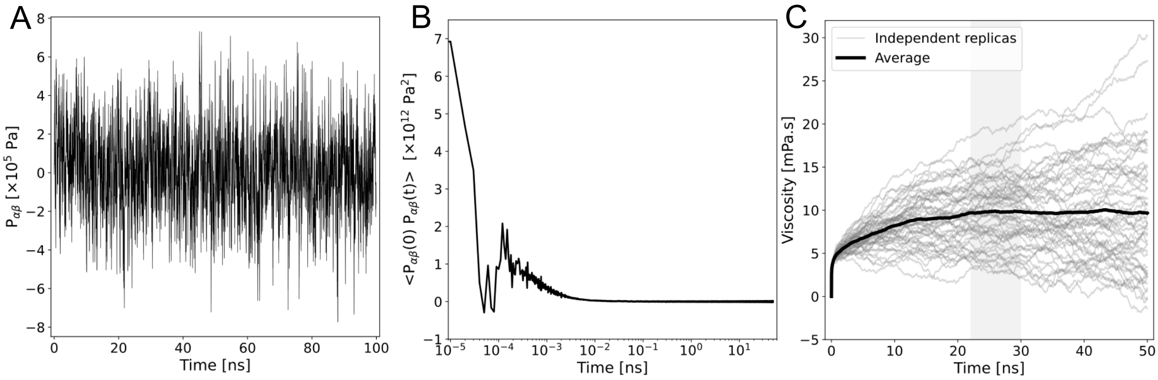

Figure 4.3. Zero shear viscosity calculation using the Green-Kubo method. A. Pressure components ($P_{\alpha\beta}$) were extracted from equilibrium MD simulations. From each simulation, we have six different components, and here, the time-trace of one component is shown. **B.** From the six pressure components, the average autocorrelation is computed for each independent simulation. **C** The zero shear viscosity is determined from the average autocorrelation of all pressure components using Equation 3 (refer to methods). Zero shear viscosity is derived by solving equation 4.3 of averaged the autocorrelation of all pressure components. The gray curves depict the viscosities of all 50 independent replicas, and the black curve represents the overall average. The viscosity is extracted from the plateau-highlighted region, with values reported as the average ± standard error.

glycosylated (w.), and 30 highly glycosylated (w.+) systems to determine the zero shear viscosities.

For equilibration, as the initial configurations were derived from the final conformation of the equilibrium MD simulations of multi-chains, we skipped energy minimization and NVT ensemble considerations. However, to ensure a fully equilibrated system, we conducted two prolonged simulations (each lasting 2000 ns) for each system within the NPT ensemble. Under the NPT ensemble, periodic boundary conditions were applied, and position restraints on heavy atoms were released. The velocity rescale thermostat and the Parrinello-Rahman algorithm were employed during production runs to maintain a constant temperature and pressure. Electrostatic interactions were considered in our simulations using the Particle Mesh Ewald algorithm [125]. Short-range interactions were modeled using a Lennard Jones potential, with calculations limited to a cutoff distance of 1.0 nm. Constraints were applied to bonds involving hydrogen atoms within the fragments, achieved through the utilization of the LINCS algorithm [173]. The Verlet Buffer was utilized to manage neighbors, with a tolerance of 0.005 kJmol^{-1} ps^{-1}. Updates to neighboring interactions were performed at regular intervals, precisely every 10 steps. The numerical integration of the equations of motion was accomplished using the Leap Frog algorithm, employing discrete time steps set at 2 fs.

We computed the viscosity, by extracting 25 conformations at 8 ns intervals from the last 200 ns of each of the two 2000 ns equilibrium simulations. This resulted in a total of 50 extracted conformations for each of the three systems under consideration. To initiate the equilibration of these configurations, a 2 ns NPT ensemble simulation was conducted. Subsequently, production run simulations were executed in the NVT ensemble, spanning a duration of 20-100 ns for each configuration, with the objective of obtaining the zero shear viscosity. In these production run simulations, only the velocity rescale thermostat and periodic boundary conditions were applied. To ensure a reliable estimate of the autocorrelation function (ACF) due to the high fluctuations of the pressure tensor, the pressure tensor components were recorded with a higher output frequency of 10 fs. Subsequently, the ACF (Figure 4.3B) was computed using the pressure elements. For each of the 50 independent simulations, the zero shear viscosity (η) was determined by averaging the results of the six distinct pressure components and solving equation 4.3. The final viscosity of each lubricin system was obtained from the average over the 50 values (Figure 4.3C).

4.3 Results

The results of this study can be categorized into two main sections: viscosity and rheological properties of lubricin systems under external shear stress. In this section, we will initially focus on the viscosity of lubricin and investigate how O-glycans can alter the viscosity of lubricin. First, we compute the viscosity of water to validate our methodology. Subsequently, we explore the rheological changes in lubricin under external shear stress with different shear rates. This investigation involved the calculation of various parameters, including the radius of gyration, end-to-end distance, solvent-accessible surface area (SASA), nematic order, and the P2 value.

4.3.1 Viscosity of pure water

To validate our simulation protocol, we examined the viscosity of TIP4P-D water molecules [165], which was used throughout our simulations. Due to the Newtonian behavior of water, i.e. its viscosity does not depend on the applied shear, we anticipated a constant value for both zero shear or under sheared conditions.

Given the relatively short simulation time required for water systems and our interest in understanding the impact of system size, we opted for various box sizes: 3×3×3, 4×4×4, 5×5×5, and 6×6×6 nm. These boxes were filled with TIP4P-D water molecules, and following equilibration processes (energy minimization, NVT, and NPT), five independent

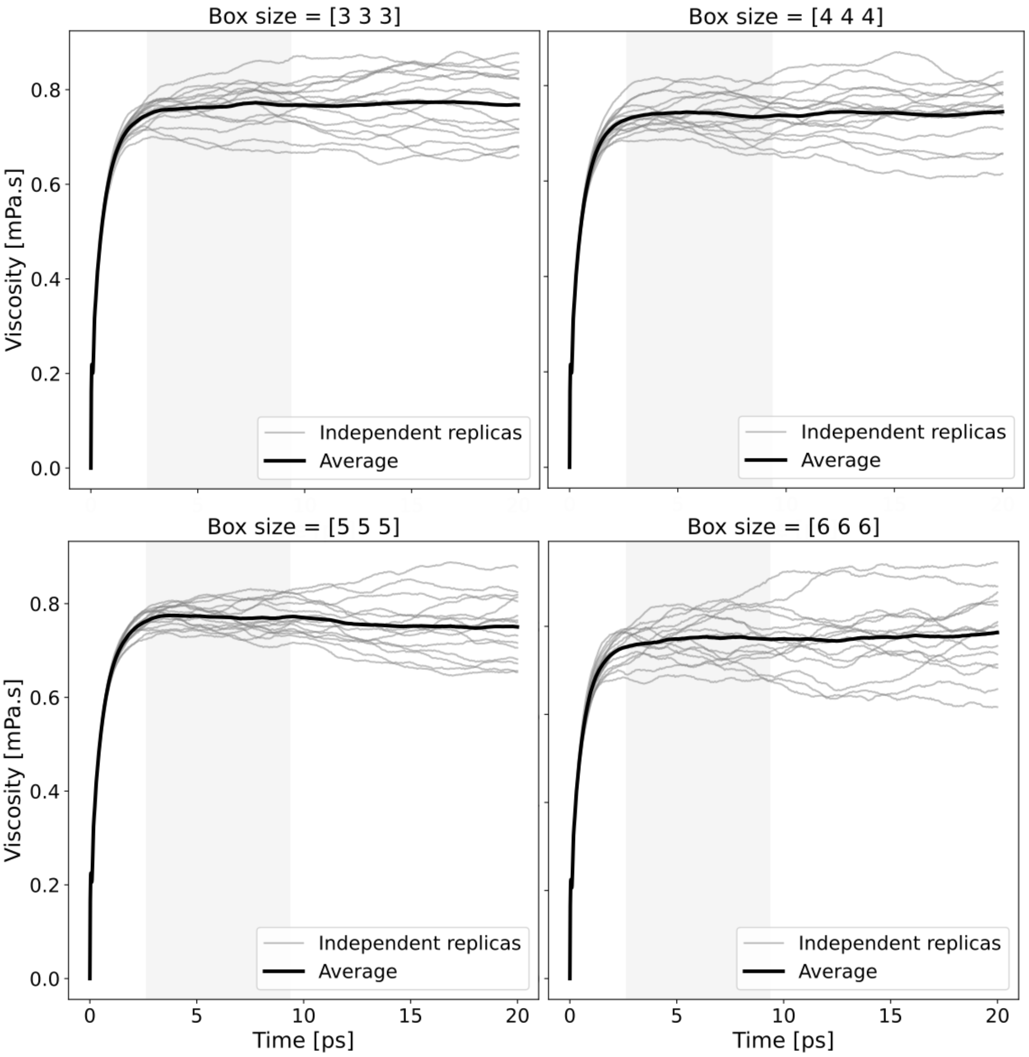

Figure 4.4. Zero shear viscosity of TIP4P-D water molecules. To determine the zero shear viscosity of TIP4P-D water molecules, four different simulation boxes were selected, and five independent simulations were conducted for each system. The zero shear viscosity was calculated by solving Equation 4.3 using the autocorrelation of three off-diagonal pressure components (gray curves). The black curve represents the overall average. The viscosity values are extracted from the plateau-highlighted region, with reported values as the average ± standard deviation.

Table 4.3. Viscosities obtained for water at the indicated shear rates. For comparison experimental value and previous estimates are also shown.

Shear velocity (nm/ns)	Shear rate (ns^{-1})	Viscosity (mPa·s)
10	2.48	0.74 ± 0.03
8	1.99	0.76 ± 0.08
4	0.99	0.99 ± 0.29
2	0.50	0.68 ± 0.17
1	0.25	0.95 ± 0.42
Experimental data [194]	0	0.69
Computational data [195]	0	$\sim 0.9 \pm 0.1$

replicas were conducted for each system, each lasting 10 ns. Due to the simplicity of the water system, we used only three off-diagonal components of the pressure tensor (P_{xy}, P_{xz}, P_{yz}) to calculate viscosity. Figure 4.4 illustrates the temporal evolution of viscosities for each system. We obtained 0.76 ± 0.04, 0.74 ± 0.04, 0.77 ± 0.03, and 0.77 ± 0.01 mPa·s for the zero shear viscosity of the 3×3×3, 4×4×4, 5×5×5, and 6×6×6 boxes, respectively.

In the next step, to calculate non-zero shear viscosities, an $8\times4\times4$ nm^3 box was filled with 4300 TIP4P-D water molecules. We selected five different shear rates (Table 4.3), and for each rate, three simulation replicas were conducted. To equilibrate, all systems underwent energy minimization, NVT, and NPT ensembles. Finally, for the shearing part, simulations were conducted in the NVT ensemble at a temperature of 310 K, following the same simulation protocol as explained in Section 4.2.2. After finishing the simulation, we calculated the viscosities using pressure elements and Equation 4.2. Table 4.3 shows the viscosity for each shear rate.

Our results affirm the anticipated Newtonian behavior of the water system, demonstrating that the viscosity remains unaltered under shear stress. Furthermore, the viscosities obtained from our simulations closely align with both experimental values [194] and other computational results [195]. This outcome validates our simulation protocol. It also establishes a reference for evaluating the viscosity of systems containing lubricin fragments.

4.3.2 O-glycans decrease viscosity of lubricin

After validation of our simulation protocol to compute the viscosity, we proceeded with the calculation of the viscosity for systems containing O-glycosylated lubricin fragments (Figure 4.5).

We obtained the zero shear viscosity for three systems (with 30 chains, N=30), as depicted in Figure 4.5A, B, and C. Each figure shows the viscosities of 50 different replicas in grey, with the overall average presented in black. The final viscosity value was selected based on the point where the average curve reached a plateau state. The vertical grey column indicates the region used to calculate viscosity. Notably, in these three cases, the systems had 30 chains with the same molar mass but varied in glycosylation level and mass density. An increase in glycosylation level or mass density necessitated a longer simulation time to reach the plateau state.

For non-zero shear viscosities, for all considered systems, the box deformation method induced a linear velocity profile (Figure 4.2B). Viscosities could then be calculated from the resulting terminal velocities using Equation 4.2. Figure 4.5D presents the viscosities under various shear rates. To provide an overall context, results from zero shear viscosity and water (in grey color) viscosities are also included. Here, Carreau and Power-law fits were employed to describe the shear rate dependence of the system [196, 197].

Figure 4.5D presents the viscosities of both water and multi-fragment lubricin systems at zero shear and various shear rates. In the case of the pure water system, viscosities exhibit an independent response to external shear stress, aligning with the expected behavior of Newtonian fluids where viscosity remains unchanged in the presence of shear. These viscosities are consistent with previous studies [194, 195]. However, unlike the water system, there is a non-Newtonian and shear-thinning behavior for all systems containing lubricin fragments. This indicates that the viscosity of these systems changes with each shear rate, decreasing from zero to high shear rates. In general, lubricin systems present higher viscosity values compared to pure water systems, indicating that they are more viscous.

In the context of lubricin systems, viscosities exhibit variations at different shear rates, in contrast to the behavior observed in pure water. As the shear rate increases, the calculated viscosity decreases, signifying a shift towards a non-Newtonian regime characterized by shear-thinning behavior. Additionally, the response of viscosity to shear shows significant variations across different systems. Both the level of glycosylation and mass density play a crucial role in altering the viscoelastic properties of lubricin. These findings highlight that an increase in protein mass density corresponds to an elevation in viscosity, as evident in the comparison of curves representing different densities (indicated by various colors). Moreover, the results indicate that glycosylation contributes to a reduction in viscosity. Within the same mass density, glycosylated systems (orange and green solid lines) exhibit lower viscosity compared to their non-glycosylated counterparts (dashed lines). Additionally, glycosylation decreases the shear-thinning behavior of the system. That is, in both highly and medium glycosylated systems, the slope in the viscosity-shear curve is lower than that of

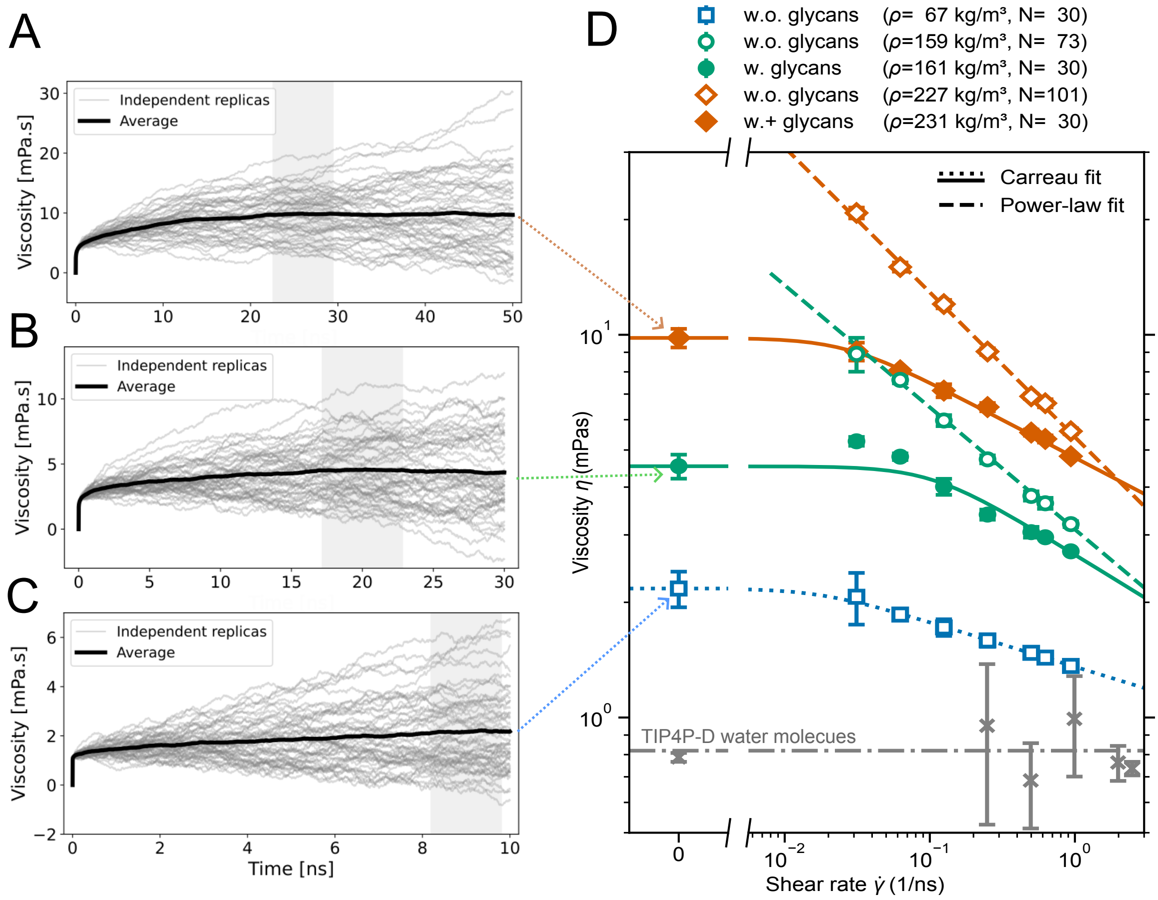

Figure 4.5. Viscosity and shear thinning behavior of lubricin multi-fragments. Zero shear viscosity of 30 non-glycosylated **(A)**, 30 medium glycosylated **(B)**, and 30 highly glycosylated **(C)** fragments is obtained using the Green-Kubo method. The black curve displays the average of 50 independent replicas for each system. Additional technical details can be found in section 4.2.3. **(D)** Non-zero shear viscosities of lubricin fragments were obtained from shear-driven non-equilibrium MD simulations (see more details in section 4.2.2). Five systems were selected with different glycosylation levels and mass densities. Open symbols (w.o) represent non-glycosylated systems with varying mass densities, while closed symbols (w. and w.+) denote glycosylated systems. TIP4P-D water viscosity is also displayed in grey for comparison.

their non-glycosylated counterparts.

In this section, we explored how O-glycans impact the viscoelastic behavior of lubricin. Our investigation revealed that systems containing glycosylated lubricin fragments display a strong shear-thinning behavior and changes in mass density and glycosylation level significantly alter such behavior. To explore deeper into the molecular mechanisms underlying these viscosity modifications, we examined various rheological properties of lubricin fragments. In the following section, we elucidate how shear stress and glycosylation dynamically influence the viscosity of lubricin.

4.3.3 Structural changes of lubricin under shear stress

Having established the influence of O-glycans on the viscosity and shear-thinning behavior of lubricin fragments, we proceeded to investigate the molecular factors underlying these alterations. Our analysis centered on specific structural characteristics within these mixtures, delving into aspects such as elongation, aggregation, and alignment responses of the lubricin fragments when subjected to external shear stress.

Elongation

To assess the elongation of individual chains, we computed the radius of gyration. In Figure 4.6A, the impact of the shear on the R_g is illustrated. Under zero shear viscosity and in the same molar mass (N=30), glycosylation increases the R_g of the system, while an increase in mass density for non-glycosylated systems results in a lower R_g. After applying external shear stress to the systems, all chains adopt more elongated conformations, and an increase in shear rate induces a further increase in R_g indicating that the chains adopt more elongated conformations under shear stress. Highly and medium glycosylated systems have higher R_g than non-glycosylated systems. Due to the influence of external shear flow, chains exert force on each other, compelling alignment along the shear flow, which results in non-glycosylated systems with higher mass density exhibiting a higher R_g compared to zero shear rate conditions. Note that other elongation metrics, such as the end-to-end distance (Figure 4.6B), extension along the flow direction (Figure 4.6C), and hydrodynamic radius (Figure 4.6D) further support the findings of Rg.

In summary, at zero shear rate, the elongation properties of lubricin fragments are significantly influenced by mass density and glycosylation. Glycosylation enhances elongation, while an increase in mass density in non-glycosylated systems leads to more compact fragments. Under shear stress, elongation is determined by both mass density and glycosylation level. In systems with the same mass density, glycosylation extends the fragments, causing them to become more expanded—an effect attributed to the electrostatic repulsion between negatively-charged sugars. Interestingly, non-glycosylated systems lack this repulsion, yet mass density still regulates the system's response to shear stress.

Aggregation

In the subsequent analysis to comprehend the influence of glycosylation on the aggregation tendency of lubricin systems, we computed the solvent-accessible surface area (SASA) for both glycosylated and non-glycosylated systems under various shear stresses. Figures 4.7A and B illustrate the results of SASA for the entire protein conglomerate and individual chains, respectively. Glycosylation primarily regulates the SASA of lubricin fragments, with an increase in glycosylation level corresponding to an increase in the SASA of chains (Figure 4.7A). In non-glycosylated systems, an increase in mass density does not alter the SASA of individual chains, consistent with the results for individual chains. For both glycosylated and non-glycosylated cases, shear stress induces a small increase in SASA, but an increase in shear rate does not impose any significant changes in SASA.

However, for the SASA of the entire system (Figure 4.7B), an increase in mass density results in a higher value for SASA. In systems with the same mass density, non-glycosylated systems exhibit a higher value of SASA due to the greater number of chains in the system (101 and 73 compared to 30 chains). Another significant result from this figure is the change in SASA for non-glycosylated systems, which is not observed in glycosylated systems. In high mass density systems, shear stress increases the SASA of the entire system.

Subsequently, to investigate and comprehend the aggregation tendencies of the systems, we calculated the ratio of the SASA of the entire protein conglomerate to the sum of the SASA of each individual chain (Equation 2.21). Figure 4.7C illustrates the impact of shear stress on aggregation and the interaction behavior of the lubricin mixtures. Here, when the value of the ratio is equal to 1, it indicates complete dissociation of the chains, implying non-interaction. Conversely, a lower value approaching 0 suggests a higher degree of association among the chains, indicating the formation of a condensed or aggregated state (additional details are provided in section 2.3.3). In non-glycosylated systems, changes in mass density influence the aggregation and intermolecular interactions between fragments. As the mass density increases, chains tend to aggregate and exhibit more interactions. However, in glycosylated systems at the same mass density as non-glycosylated systems, there are no significant changes. This implies that glycosylated systems remain dispersed, and increasing their mass density does not lead to increased aggregation. Another noteworthy result is the distinct response of glycosylated and non-glycosylated mixtures to shear. In all cases, after the imposition of shear stress, there is an increase in aggregation or intermolecu-

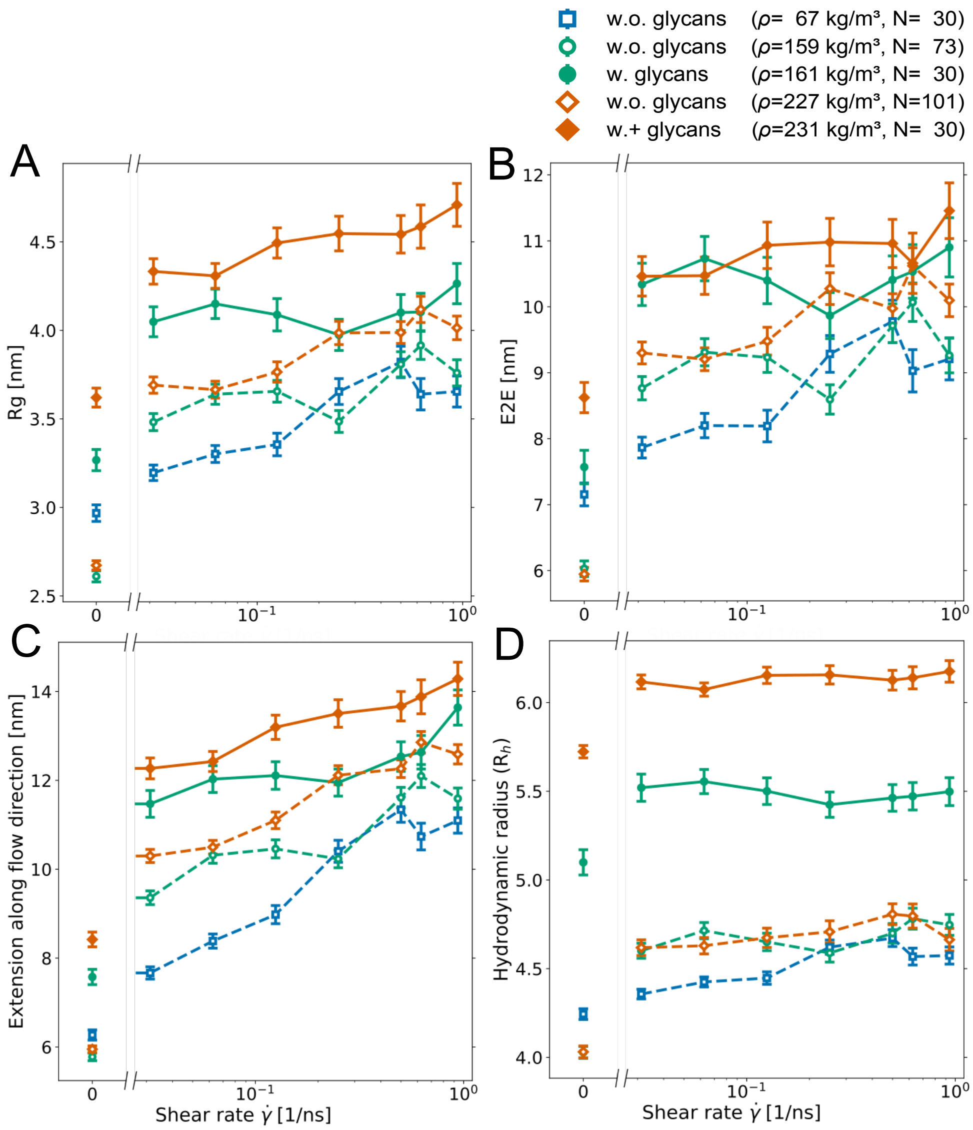

Figure 4.6. Impact of shear stress on the elongation of lubricin fragments. Radius of gyration **(A)**, end-to-end distance **(B)**, extension along shear flow **(C)**, and hydrodynamic radius **(D)** as a function of shear rate ($\dot{\gamma}$). Symbols represent data obtained from Equilibrium (zero shear rate) and shear-driven non-equilibrium (non-zero shear rates) MD simulations (average ± standard error, n = 4). These five different systems have varying mass density and glycosylation levels. Dashed lines represent non-glycosylated systems (w.o.), while solid lines indicate glycosylated ones (medium glycosylated: w. and highly glycosylated: w.+). The same color indicates the same mass densities.

lar interactions between fragments. However, non-glycosylated chains tend to disaggregate and have fewer interactions with increasing shear stress. Conversely, glycosylated systems display insensitivity to shear, and their aggregation tendency remains unchanged.

Alignment

For further investigation, we explored the alignment of lubricin fragments with different glycosylation levels and mass densities. This involved computing both the Nematic Correlation Function (NCF) and the ordering parameter (P2 value). These parameters offer insights into how the fragments align relative to other chains. The NCF tracks the coordinated orientation of all chains, considering their radial distribution function (refer to section 2.3.4 for details). The P2 value quantifies the orientational order in a system and provides information about the bulk alignment (see section 2.3.5).

Figure 4.8 depicts the Nematic Correlation Function (NCF) for all systems across various shear rates, where values of 0.5 and 1 represent fully random and fully aligned chain configurations, respectively. This figure shows that external shear stress enhances the alignment of chains in all cases, indicating an increased alignment with higher shear rates. In non-glycosylated systems, mass density significantly influences fragment alignments, resulting in a higher degree of alignment at increased density. This effect is observed both locally, considering first-neighbor chains (depicted in the grey column), and for distant chains (illustrated by purple double-headed arrows). On the other hand, in glycosylated systems, mass density does not significantly alter alignment properties for distant chains, but it contributes to higher alignment at short distances. Glycosylation itself does not have discernible effects on the alignment of systems. However, it can influence the first-neighbor distance of chains or create more spacing between these chains due to the presence of bulky glycan side chains.

In the subsequent analysis to scrutinize the alignment of lubricin fragments more closely, we divided the NCF into short ($d < 4nm$) and long ($d > 4nm$) distances. Figure 4.9A and 4.9B illustrate the short and long alignment of different systems under various shear rates, respectively. In both cases, shear rates enhance chain alignment. For short-range alignment, the alignment of non-glycosylated systems increases with mass density, while glycosylated systems exhibit lower values of NCF compared to their counterparts. In the long range, it is evident that mass density regulates the alignment and orientation of systems, while glycosylation does not play a significant role. The results from the P2 value of the ordering parameter in Figure 4.9C align well with the results of NCF in the long range, indicating the substantial influence of mass density on the alignment of lubricin fragments.

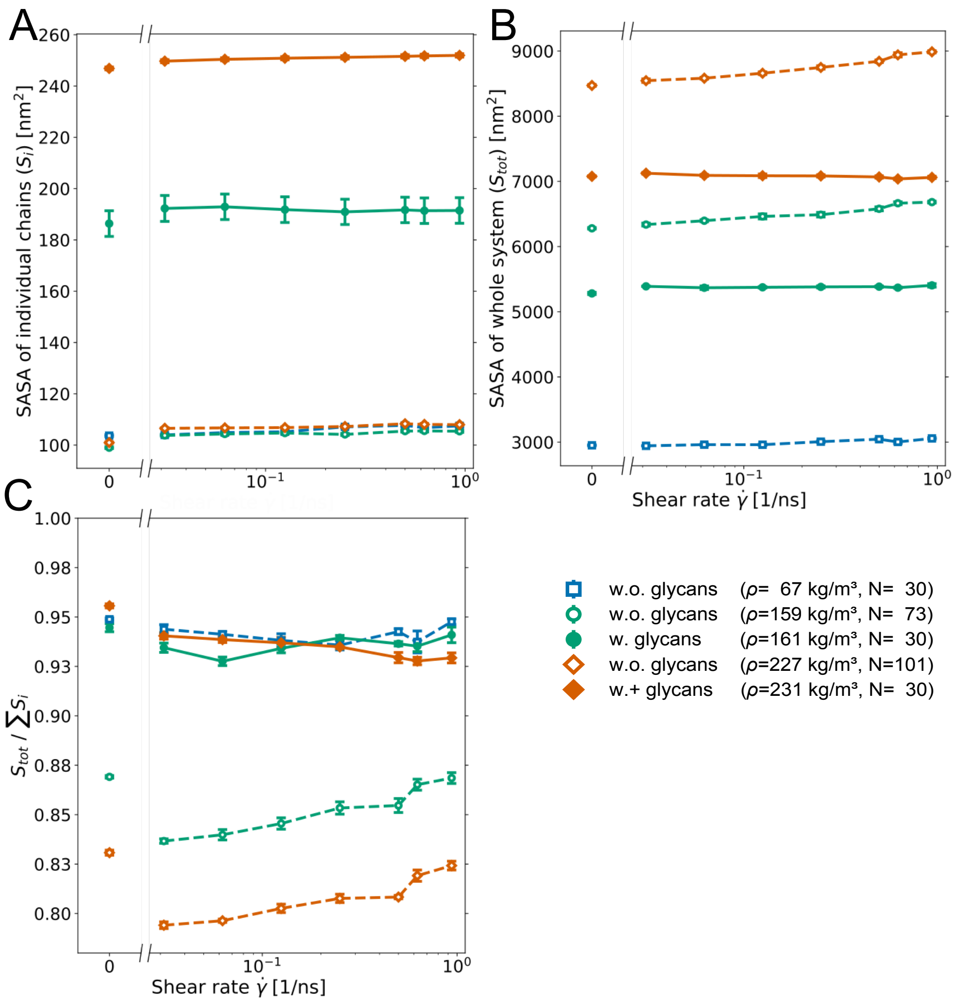

Figure 4.7. Impact of shear stress on the aggregation and interaction of the lubricin fragments. The averaged SASA of individual fragments **(A)** and the sasa of whole chains altogether **(B)**, as a function of shear rate ($\dot{\gamma}$), are calculated for five different systems. **C.** Aggregation tendency ($S_{tot}/\sum S_i$) is obtained by using equation 2.21 Symbols represent data obtained from Equilibrium (zero shear rate) and shear-driven non-equilibrium (non-zero shear rates) MD simulations (average ± standard error, n = 4). Dashed lines represent non-glycosylated systems (w.o.), while solid lines indicate glycosylated ones (medium glycosylated: w. and highly glycosylated: w.+). The same color indicates the same mass densities.

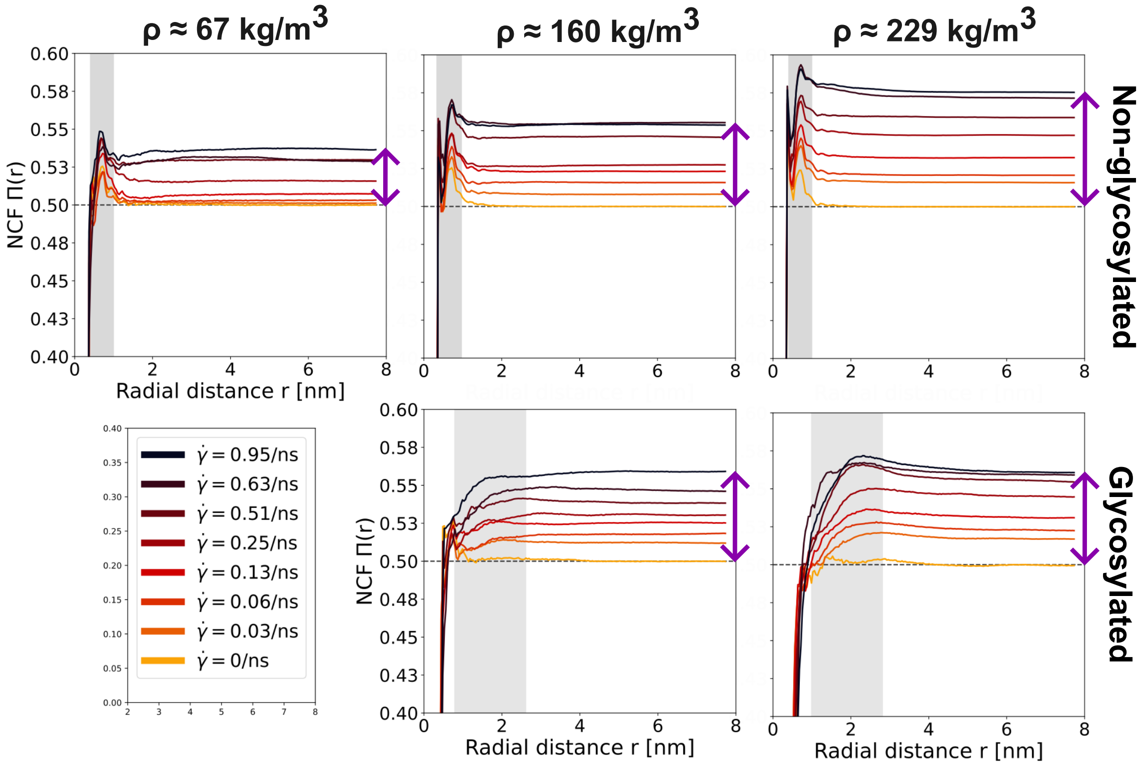

Figure 4.8. Impact of shear stress on the alignment of lubricin fragments. Nematic Correlation Function (NCF) depicting the alignment of non-glycosylated (top row) and glycosylated (bottom row) systems with different mass concentrations (columns) as a function of radial distance. Distinct shear rates are represented by different colors, from yellow to black. An NCF value of 0.5 signifies a random chain orientation, while 1.0 indicates full alignment. The gray area corresponds to the region of nearest neighbors, displaying heightened alignment.

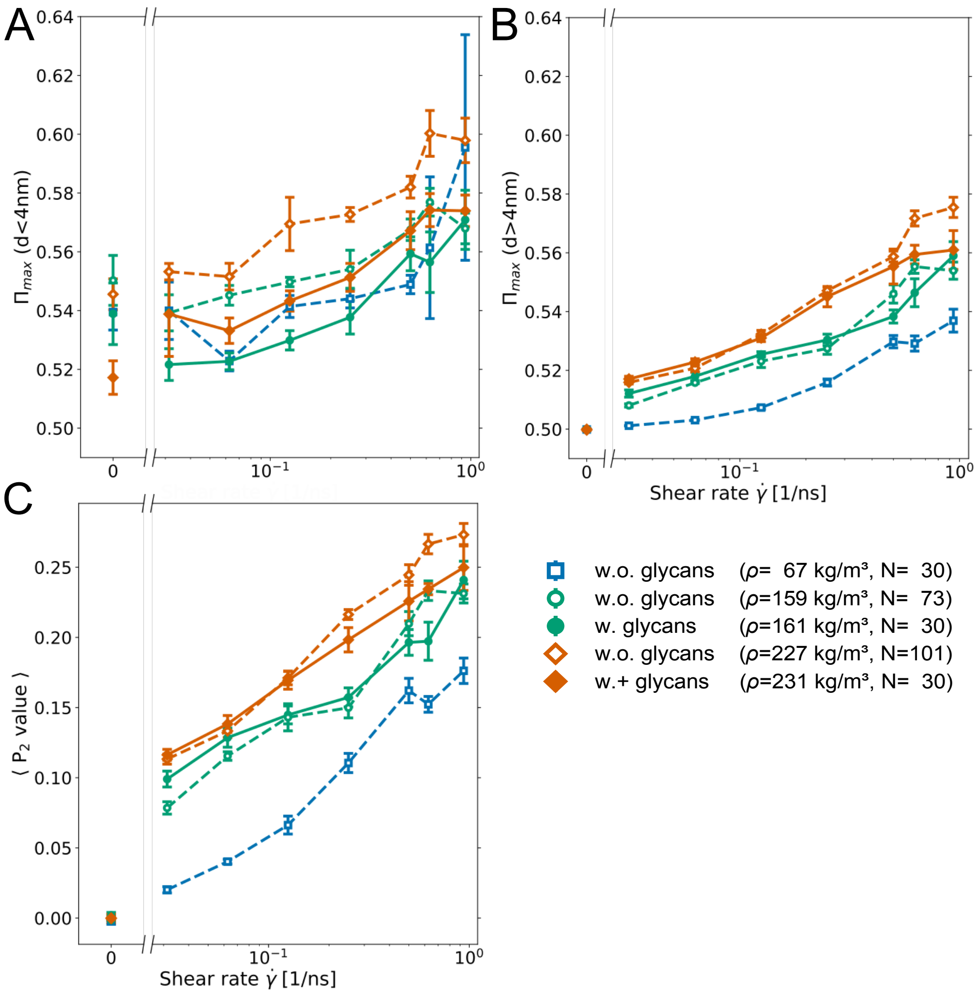

Figure 4.9. short and long range alignment of lubricin fragments. The NCF of short distances **(A)**, long distances **B** ans also p2 value of ordering parameters **C** as a function of the radial distance. Symbols represent data obtained from Equilibrium (zero shear rate) and shear-driven non-equilibrium (non-zero shear rates) MD simulations (average ± standard error, n = 4). Dashed lines represent non-glycosylated systems (w.o.), while solid lines indicate glycosylated ones (medium glycosylated: w. and highly glycosylated: w.+). The same color indicates the same mass densities.

4.3.4 Viscosity and structural characteristics

Having analyzed the structural changes of lubricin in detail, we next asked which of these structural properties can explain the change in viscosity across shear flow. We assessed the viscosity ratio between glycosylated and non-glycosylated systems ($\eta(w.)/\eta(w.o.)$) and compared it with the ratios of the radius of gyration (Figure 4.10A), solvent-accessible surface area (Figure 4.10B), and nematic order parameter (Figure 4.11) between glycosylated and non-glycosylated systems ($X(w.)/X(w.o.)$).

Figure 4.10C-D and Figure 4.11C-D illustrate the relationship between viscosity and structural properties in both the medium (green) and high (orange) mass density regimes. These results emphasize a linear correlation with the aggregation or inter-chain interactions of lubricin fragments. This correlation is much weaker for other structural observables. Therefore, the shear-thinning behavior of lubricin fragments is more correlated with the clustering of chains than with other structural properties. Thus, in conclusion, shear thinning can be explained by that lubricin aggregates less when exposed to higher shear. In addition, the response of lubricin multi-chain system to shear stress is significantly influenced by both the protein density and the level of glycosylation. We can also conclude that this tendency to cluster is not only shear-dependent but also strongly altered by protein density and the level of glycosylation.

4.4 Discussion

In this chapter, we investigated how O-glycans influence the viscoelastic and rheological properties of multi-chain systems consisting of lubricin fragments. To assess the viscosity response to shear, we employed both equilibrium and non-equilibrium MD simulations to calculate the viscosity at zero and non-zero shear rates. For obtaining zero shear viscosity, the Green-Kubo method was utilized in equilibrium MD simulations. On the other hand, to determine the viscosity of lubricin systems under varying shear rates, we employed the box deformation method in shear-driven non-equilibrium MD simulations. To investigate the impact of lubricin and its glycosylation on medium viscosity, we opted for a computationally efficient strategy. Rather than examining the entire disordered region, comprising 800 amino acids and 185 glycans, we concentrated on short glycosylated fragments extracted from it. This approach enables the replication of lubricin content in a compact volume unit, marking a promising initial step in this research direction. A key finding of our investigation is that O-glycans have the capacity to decrease viscosity and mitigate shear thinning behavior in systems when compared to their non-glycosylated counterparts with the same mass density. Our results indicate that the presence of bulky side chains of O-glycans, influenced

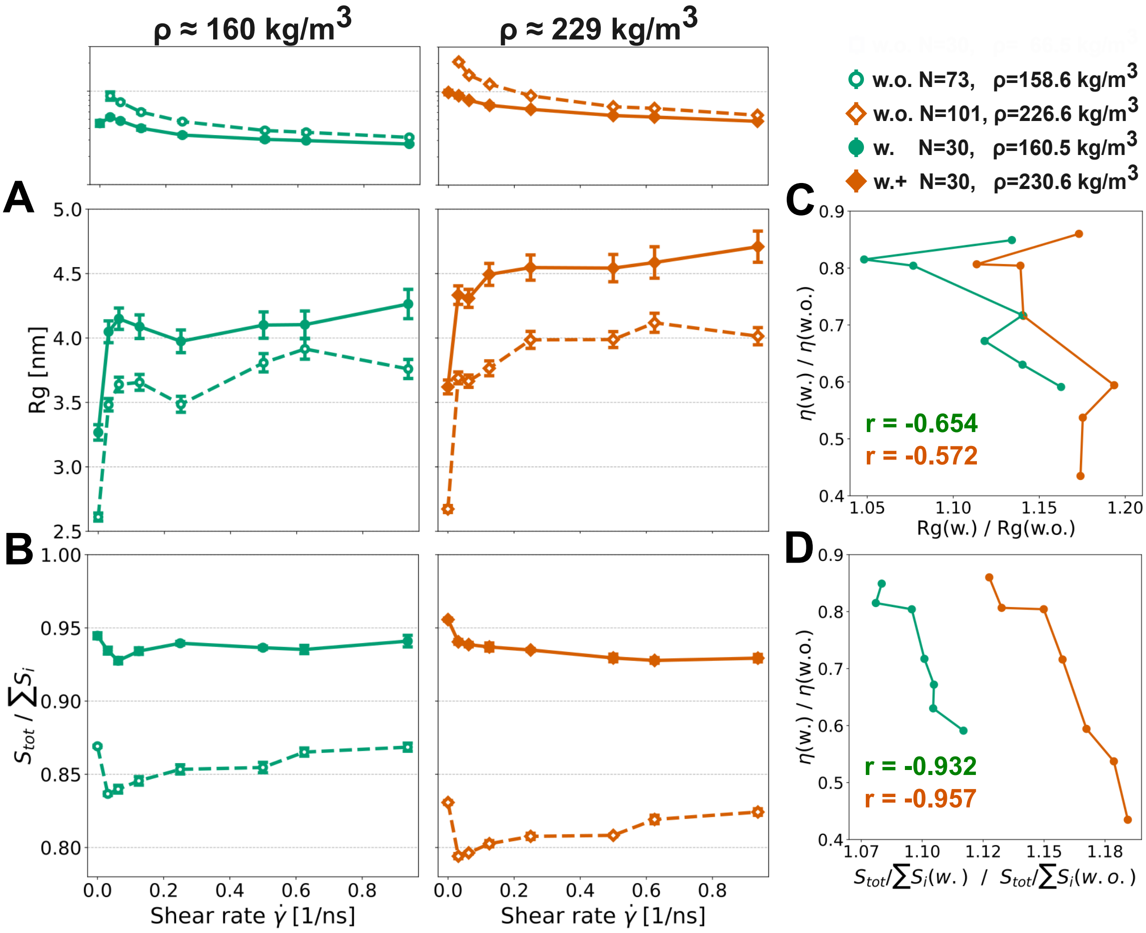

Figure 4.10. Rheological properties and viscosity of lubricin fragments. A. Radius of gyration (R_g) and **B.** aggregation tendency ($S_{tot}/\sum S_i$) of medium (green) and highly (orange) glycosylated systems and their non-glycosylated counterparts as a function of shear rate ($\dot{\gamma}$). Symbols represent data obtained from equilibrium (zero shear rate) and shear-driven non-equilibrium (non-zero shear rates) MD simulations (average ± standard error, n = 4). Dashed lines represent non-glycosylated systems (w.o.), while solid lines indicate glycosylated ones (medium glycosylated: w. and highly glycosylated: w.+). The same color indicates the same mass densities. **C-D.** The viscosity ratio $\eta(w.)/\eta(w.o.)$ is plotted against the ratio of radius of gyration and aggregation tendency. Correlation coefficients (r) are shown for each dataset.

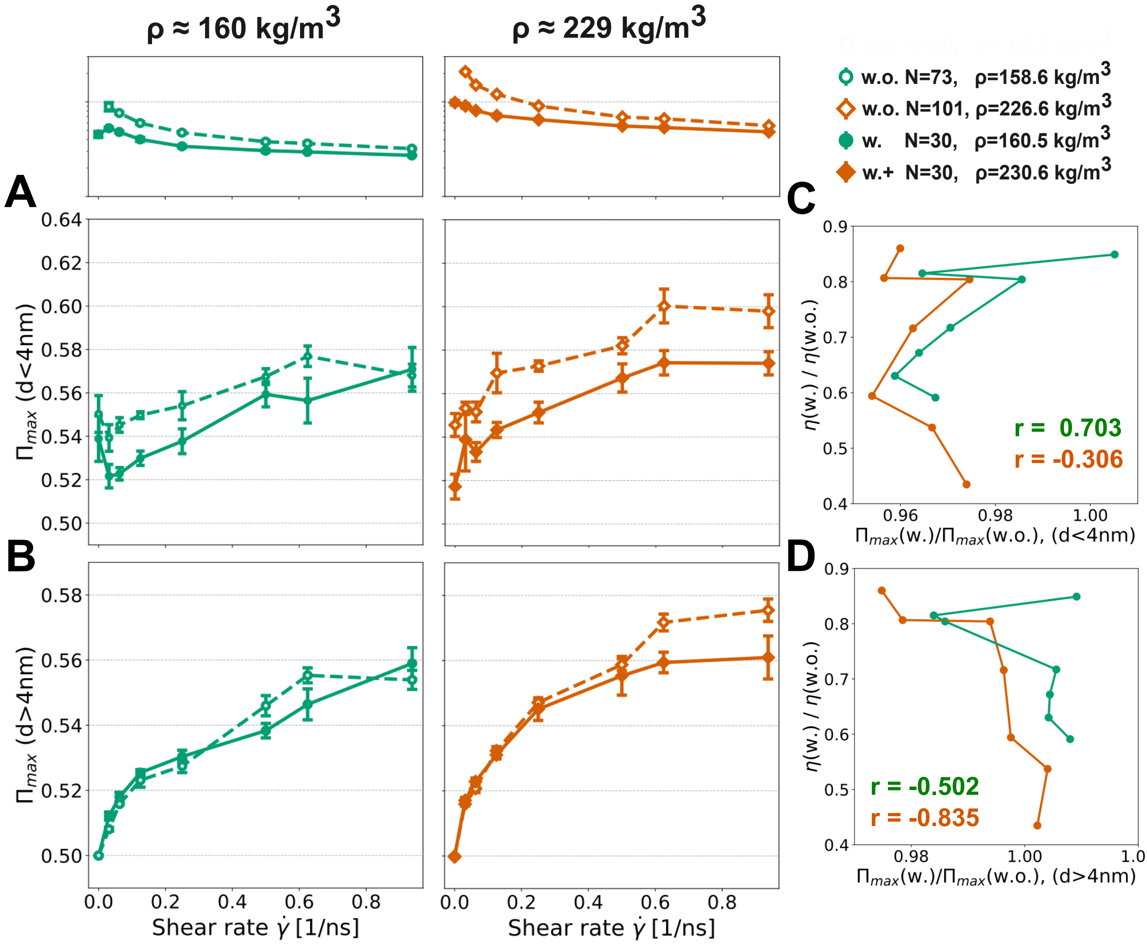

Figure 4.11. Alignment properties and viscosity of lubricin fragments. Nematic correlation function in short ($d < 4\ nm$) **(A)** and long **(B)** distance as a function of shear rate ($\dot{\gamma}$) in both the medium (green) and high (orange) mass density regimes. Symbols represent data obtained from equilibrium (zero shear rate) and shear-driven non-equilibrium (non-zero shear rates) MD simulations (average ± standard error, n = 4). Dashed lines represent non-glycosylated systems (w.o.), while solid lines indicate glycosylated ones (medium glycosylated: w. and highly glycosylated: w.+). The same color indicates the same mass densities. **C-D.** The viscosity ratio $\eta(w.)/\eta(w.o.)$ is plotted as a function of the ratio $X(w.)/X(w.o.)$, with $X = NCF_{\max}$, i.e. the maximum of the NCF, at short-range (C) and the long-range (D) separations. Correlation coefficients (r) are shown for each dataset.

by steric and electrostatic interactions, reduces the aggregation tendency. This reduction in aggregation tendency contributes to a weakening of the shear-thinning behavior in response to external shear stress.

To validate our viscosity protocol, we initially calculated the viscosities of a pure water system. Figure 4.5D illustrates that the viscosity of water remains independent of external shear stress and exhibits consistency. This outcome aligns well with the expected Newtonian behavior of water systems. Furthermore, the computed viscosity values for water are in agreement with experimental data [198] and another computational model [195]. Consequently, the TIP4P-D water model [165] proves to be a suitable choice for investigating IDPs and their viscosity.

To evaluate the viscosity of lubricin fragments, we employed both equilibrium and shear-driven non-equilibrium MD simulations. For non-zero shear viscosities and to achieve a linear velocity profile representative of Couette shear flow, we applied the box deformation method along with Lees-Edwards periodic boundary conditions [135]. The GROMACS package [172], specifically the GROMACS-2023-dev version (GROMACS-2023-dev version), which incorporates these methods, was utilized for our simulations (see section 4.2.2). This method offers ease of setup compared to alternatives, such as systems with two moving walls. Additionally, it aligns more closely with biological or experimental settings, where walls are typically absent. Figure 4.2B depicts the velocity profile obtained from our simulations and confirms that this methodology is capable of reproducing sustained planar Couette flows. However, there are limitations to this method. Due to the need for extended simulation times, especially for large systems, and the requirement for substantial deformation velocities to achieve a linear velocity profile, it becomes challenging to compute viscosities at small shear rates. To overcome this limitation, EMD simulations and the Green-Kubo method [145, 199] can be employed. In this method, viscosity at zero shear can be obtained from the autocorrelation of pressure tensor components, as detailed in Section 4.2.3. These methods help bridge the gap between experimental observations [200] and our shear-driven non-equilibrium MD simulation results. Nevertheless, this method is only applicable to systems with a zero shear viscosity below 20–30 mPa.s [144, 193]. Therefore, we employed this approach exclusively for systems expected to exhibit low viscosity values, such as medium and highly glycosylated systems and the dilute system of non-glycosylated ones.

Lubricin is known to play a crucial role in providing effective lubrication in synovial joints [9, 11–13, 45]. Our findings reinforce this by revealing that O-glycans, through their steric and electrostatic interactions, contribute to reducing the viscosity of lubricin. The viscosity results obtained for lubricin fragments closely align with those reported in experimental studies conducted by Ludwig et al. [200]. Furthermore, our simulations confirm a Newtonian behavior in pure glycosylated lubricin, consistent with experimental observations [200]. However, there are minor discrepancies between our estimates for glycosylated systems and the experimental results. One contributing factor is our focus on small lubricin fragments, while the experimental study considered the entire chain. This discrepancy may result in some entanglement or surface attachment of lubricin proteins in the experimental setting, which is not captured in our study, where all fragments are fully unentangled. Additionally, there is a difference in the mass density of proteins; in our study, mass densities (ranging from 60 to 230 kg/m^3) are at least 500 times higher than the experimental protein density [200]. A potential solution to address these differences and enhance accuracy is to employ a coarse-grained model of lubricin and consider the entire chain in future investigations. However, as mentioned in Chapter 3, the utilization of the coarse-grained model in Martini3 necessitates the parameterization of glycans. This is essential due to the current lack of appropriate and well-defined parameters for glycans in the force field.

Our findings demonstrate that O-linked glycosylation plays a crucial role in regulating viscoelastic behavior of lubricin under different shear stresses. To understand how O-glycans impact the viscosity and shear-thinning response of lubricin fragments at the atomistic level, we examined the elongation, aggregation, and alignment of lubricin fragments under various shear stresses. The presence of glycans prevents the formation of conglomerates, likely due to steric and electrostatic repulsion between fragments. This leads to increased inter-fragment spacing, accommodating water, and reduced intermolecular interactions. In contrast, non-glycosylated systems form conglomerates with limited water spacing and pronounced intermolecular interactions after shear stress-induced extension (Figure 4.10D). Intriguingly, neither the elongation (Figure 4.10A) of the chains nor their alignment (Figure 4.11C, D) directly correlates with this mitigation. Overall, the reduction in viscosity due to shear is suggested to be linked to the amount of space between chains available to the solvent. It is plausible that other highly glycosylated systems, such as mucin or aggrecan, exhibit a similar behavior when subjected to shear [201, 202]. In our study, we demonstrated that lubricin, particularly through its O-glycans, has the capacity to decrease the viscosity of synovial fluid.

It is important to recognize that effective lubrication in joints is a intricate interplay involving various components, including both cartilage and synovial fluid. Furthermore, the rheological properties and surface characteristics of cartilage can exert a significant influence on lubrication, especially in boundary conditions. To gain a more comprehensive understanding of the precise mechanisms governing joint function, further research becomes essential. These systems could consider longer lubricin chains and other synovial components, such as hyaluronic acids, known to impact fluid viscosity [200]. Additionally, incorporating boundary effects like surface attachment [42] would enhance the realism of such simulations.

5

Conclusions and future perspectives

Lubricin is an essential constituent of synovial joints, whose function is to act as lubricant within joints. It is a multi-domain protein connected by a long intrinsically disordered glycosylated region. The O-glycans are believed to play a pivotal role in the lubrication function of this protein, by reducing the viscosity in a shear-dependent fashion [12–14, 68, 83–85]. However, the molecular mechanism underlying this response has remained largely unkown. The main goal of this thesis was to elucidate the molecular mechanisms underlying the lubrication response of lubricin, particularly focusing on the role of the glycans.

To attain this objective, this dissertation comprised two main parts. In the first part, we delved into modeling glycosylated lubricin fragments, at an atomistic level of resolution, considering an appropriate molecular interaction force field for intrinsically disordered glycoproteins, which could subsequently be utilized in equilibrium Molecular Dynamics simulations. Lubricin, with approximately 1400 amino acids and 185 O-glycans [69], is a large and complex protein. Due to computational constraints, we selected smaller fragments of lubricin to make the simulations feasible, focusing on representative segments that encapsulate the essential physicochemical features of the entire structure. Then, we generated six different glycosylation fragments for each segment using the Monte Carlo sampling approach. Accordingly, we systematically explored the influence glycans had on the structure of lubricin fragments. In the second part of our investigation, we focused on evaluating the viscoelastic properties of systems containing lubricin. This involved employing both the Green-Kubo [147–149] and box deformation [146, 152, 153] methods, respectively, to determine the viscosity of the medium in response to shear, the glycosylation and density of lubricin. Through these approaches, we aimed to unravel the intricate interplay of lubricin and its O-glycans, shedding light on their collective role in determining the viscoelastic behavior of synovial fluids.

First, we investigated the influence of O-glycans on the structural properties of single-chain short lubricin fragments, capturing the essential physico-chemical ingredients of its intrinsically disordered region, such as proline content, charged, hydrophilic, and glycosylated residues (chapter 3). Through the calculation of different structural indicators, including the radius of gyration, end-to-end distance, solvent-accessible surface area (SASA), and persistence length, we we demonstrate that O-glycans have the potential to expand the structure of lubricin. Specifically, a higher abundance of glycans correlated with a more elongated structures. Glycans contribute to an increased exposure of lubricin to the surrounding solution molecules, by steric and electrostatic repulsion. The presence of O-glycans also increased the persistence length of the fragments, thereby indicating these type of carbohydrates have an impact increasing the stiffness of lubricin too.

In the second part of this thesis (chapter 4), we focused on the viscosity a direct descriptor of the viscoelastic behavior of lubricin, which connects with its main biological function acting as lubricant. In this section, to validate our protocol for computing zero shear viscosity and viscosities under different shear rates, we obtained the viscosity of pure water. As anticipated, pure water exhibited a Newtonian behavior, i.e. the obtained viscosities were independent of shear stress. Furthermore, the obtained value closely aligned with experimental estimates [194] and other computational studies [195]. In particular, we recapitulated the viscosity predicted for the TIP4P-D water model, the water model that we used throughout this thesis and which has been specifically designed for the simulation of intrinsically disordered segments [165]. These findings provided confirmation of the validity of our protocol to compute the viscosity. After confirming the validity of our protocol, we proceeded to calculate the viscosity of different lubricin systems with varying levels of glycosylation, as well as different molar and mass densities. Contrary to water, the presence of glycosylated lubricin fragments induced a pronounced shear-thinning response. The results revealed that glycosylation and mass density play a crucial role in regulating the viscosity. An increase in mass density induced an elevation in system viscosity. However, under the same mass density, O-glycans were found to attenuate the viscosity and shear-thinning behavior of the system compared to their non-glycosylated counterparts. To elucidate the exact mechanism behind this viscosity reduction, we examined various elongation, aggregation, and alignment parameters of lubricin fragments. Our results indicate that the presence of O-glycans prevents aggregation and reduces the interaction between chains. Additionally, this leads to a more open space between chains, facilitating solvent molecule penetration. We believe these physico-chemical structural features, that we unravel through this study, give lubricin the ability to respond to external shear flows to thereby regulate the medium viscosity and thereby influence the lubrication of the synovial joint.

Encouragingly, the viscosities we estimated for glycosylated systems are comparable to the experimental results reported by Ludwig et al. [200]. However, it is important to note that our study employed a higher density and shorter protein chains compared to the experimental study. Despite these differences, we successfully captured the shear-thinning behavior of the lubricin system.

It is essential to acknowledge that our study concentrated on small fragments of lubricin and selected glycans, offering a limited perspective on the comprehensive mechanism behind excellent joint lubrication. Effective joint lubrication is a synergistic outcome involving contributions from both cartilage and synovial fluid components. Moreover, the rheology and surface characteristics of cartilage play a pivotal role in influencing lubrication, especially in boundary conditions. Therefore, further research is imperative to achieve a comprehensive understanding of the precise mechanisms governing joint function. In the upcoming section, we will outline some ideas and viewpoints to better understand how synovial joints work.

Due to computational limitations, in our study, we focused on small parts of lubricin that capture essential aspects of its structure. However, it is important to note that lubricin has two globular end domains that could influence its overall structure. To fully understand lubricin's structure, we need higher-resolution methods (given its 1400 amino acids and 185 O-glycans [69]). While coarse-grained models like Martini3 [179] or one bead per amino acid (CALVADOS [180, 181]) exist for disordered proteins, there is a lack of parameters for glycans. Developing these parameters would enable a more comprehensive study of lubricin's complete structure. Moreover, this parameterization could help include other joint components (like hyaluronic acid, aggrecan, collagen, etc.) in our analysis.

To enhance viscosity calculation, employing a coarse-grained approach allows us to model the entire length of lubricin with all its glycans. This expanded scope facilitates viscosity calculations at lower shear rates. Another avenue for refinement involves introducing collagen or other cartilage components into the system [76–79]. Binding lubricin to these components before subjecting them to shear flow can offer a more realistic simulation.

Moreover, including additional synovial joint components provides a comprehensive overview of the lubrication mechanism. Components such as hyaluronic acid and aggrecan play crucial roles in joint function and should be considered for a more holistic understanding [40, 50, 54, 55, 202].

The insights derived from this study significantly deepen our understanding of the molecular intricacies of lubricin. These findings underscore the importance and impact of O-glycans on lubricin, highlighting their crucial role in contributing to the proper lubrication mechanism of joints. Moreover, the results shed light on the intricate interplay between lubricin's conformational dynamics and its vital role in finely modulating lubrication within the complex context of synovial joints.

Bibliography

[1] John D Ferry. *Viscoelastic properties of polymers*. John Wiley & Sons, 1980.

[2] Masao Doi. *Soft matter physics*. Oxford University Press, USA, 2013.

[3] Shaobai Li and Xuelin Sun. Droplet shape and drag coefficients in non-newtonian fluids: A review. *ChemBioEng Reviews*, 2023.

[4] Narendra Kumar, Amit Verma, and Ajay Mandal. Formation, characteristics and oil industry applications of nanoemulsions: A review. *Journal of Petroleum Science and Engineering*, 206:109042, 2021.

[5] Sahar Sultan, Gilberto Siqueira, Tanja Zimmermann, and Aji P Mathew. 3d printing of nano-cellulosic biomaterials for medical applications. *Current Opinion in Biomedical Engineering*, 2:29–34, 2017.

[6] Nicholas Michelarakis, Florian Franz, Konstantinos Gkagkas, and Frauke Gräter. Longitudinal strand ordering leads to shear thinning in nafion. *Physical Chemistry Chemical Physics*, 23(45):25901–25910, 2021.

[7] Richard A Cone. Barrier properties of mucus. *Advanced Drug Delivery Reviews*, 61(2):75–85, 2009.

[8] S More, A Kotiya, A Kotia, SK Ghosh, Leonidas A Spyrou, and Ioannis E Sarris. Rheological properties of synovial fluid due to viscosupplements: A review for osteoarthritis remedy. *Computer Methods and Programs in Biomedicine*, 196:105644, 2020.

[9] Gavin Gonzales, Stefan Zauscher, and Shyni Varghese. Progress in the design and synthesis of viscosupplements for articular joint lubrication. *Current Opinion in Colloid & Interface Science*, page 101708, 2023.

[10] Roman J Krawetz, Yiru Elizabeth Wu, Karri L Bertram, Anchita Shonak, Anand O Masson, Guomin Ren, Catherine Leonard, Mohit Kapoor, John R Matyas, and Paul T Salo. Synovial mesenchymal progenitor derived aggrecan regulates cartilage homeostasis and endogenous repair capacity. *Cell Death & Disease*, 13(5):470, 2022.

[11] Jacob Klein. Molecular mechanisms of synovial joint lubrication. *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, 220(8):691–710, 2006.

[12] Yunsup Lee, Jaehoon Choi, and Nathaniel S Hwang. Regulation of lubricin for functional cartilage tissue regeneration: a review. *Biomaterials Research*, 22:1–10, 2018.

[13] Gregory D Jay and Kimberly A Waller. The biology of lubricin: near frictionless joint motion. *Matrix Biology*, 39:17–24, 2014.

[14] GD Jay, JR Torres, ML Warman, MC Laderer, and KS17404241 Breuer. The role of lubricin in the mechanical behavior of synovial fluid. *Proceedings of the National Academy of Sciences*, 104(15):6194–6199, 2007.

[15] GERARD A Ateshian, VC Mow, and R Huiskes. Friction, lubrication, and wear of articular cartilage and diarthrodial joints. *Basic Orthopaedic Biomechanics and Mechano-biology*, 3:447–494, 2005.

[16] Alexander Y Hui, William J McCarty, Koichi Masuda, Gary S Firestein, and Robert L Sah. A systems biology approach to synovial joint lubrication in health, injury, and disease. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 4(1):15–37, 2012.

[17] Gregory M Williams, Elaine F Chan, Michele M Temple-Wong, Won C Bae, Koichi Masuda, William D Bugbee, and Robert L Sah. Shape, loading, and motion in the bioengineering design, fabrication, and testing of personalized synovial joints. *Journal of Biomechanics*, 43(1):156–165, 2010.

[18] Konstantine A Fetfatsidis, Lisa M Gamache, Jennifer L Gorczyca, James A Sherwood, David Jauffrès, and Julie Chen. Design of an apparatus for measuring tool/fabric and fabric/fabric friction of woven-fabric composites during the thermostamping process. *International Journal of Material Forming*, 6:1–11, 2013.

[19] Bruno Zappone, Marina Ruths, George W Greene, Gregory D Jay, and Jacob N Israelachvili. Adsorption, lubrication, and wear of lubricin on model surfaces: polymer brush-like behavior of a glycoprotein. *Biophysical Journal*, 92(5):1693–1708, 2007.

[20] Lin Han, Alan J Grodzinsky, and Christine Ortiz. Nanomechanics of the cartilage extracellular matrix. *Annual Review of Materials Research*, 41:133–168, 2011.

[21] Debby P Chang, Nehal I Abu-Lail, Farshid Guilak, Gregory D Jay, and Stefan Zauscher. Conformational mechanics, adsorption, and normal force interactions of lubricin and hyaluronic acid on model surfaces. *Langmuir*, 24(4):1183–1193, 2008.

[22] Nam Vo, Laura J Niedernhofer, Luigi Aurelio Nasto, Lloydine Jacobs, Paul D Robbins, James Kang, and Christopher H Evans. An overview of underlying causes and animal models for the study of age-related degenerative disorders of the spine and synovial joints. *Journal of Orthopaedic Research*, 31(6):831–837, 2013.

[23] David T Felson. Osteoarthritis of the knee. *New England Journal of Medicine*, 354(8):841–848, 2006.

[24] Verna Wright and Eric L Radin. *Mechanics of Human Joints: Physiology: Pathophysiology, and Treatment*. CRC Press, 2020.

[25] DV Davies. Synovial membrane and synovial fluid of joints. *The Lancet*, 248(6432):815–819, 1946.

[26] F Eckstein, M Winzheimer, J Hohe, K-H Englmeier, and M Reiser. Interindividual variability and correlation among morphological parameters of knee joint cartilage plates: analysis with three-dimensional mr imaging. *Osteoarthritis and Cartilage*, 9(2):101–111, 2001.

[27] Marian W Ropes, Elsie C Rossmeisl, Walter Bauer, et al. The origin and nature of normal human synovial fluid. *The Journal of Clinical Investigation*, 19(6):795–799, 1940.

[28] Abu Ilius Faisal, Sumit Majumder, Tapas Mondal, David Cowan, Sasan Naseh, and M Jamal Deen. Monitoring methods of human body joints: State-of-the-art and research challenges. *Sensors*, 19(11):2629, 2019.

[29] Van C Mow and Rik Huiskes. *Basic orthopaedic biomechanics & mechano-biology*. Lippincott Williams & Wilkins, 2005.

[30] D Dowson. Paper 12: modes of lubrication in human joints. In *Proceedings of the Institution of Mechanical Engineers, Conference Proceedings*, volume 181, pages 45–54. SAGE Publications Sage UK: London, England, 1966.

[31] ZM Jin and D Dowson. Micro-elastohydrodynamic squeeze-film lubrication of compliant layered surfaces firmly bonded to a rigid substrate. In *Tribology Series*, volume 32, pages 361–369. Elsevier, 1997.

[32] CW McCutchen. Mechanism of animal joints: sponge-hydrostatic and weeping bearings. *Nature*, 184(4695):1284–1285, 1959.

[33] Charles W McCutchen. The frictional properties of animal joints. *Wear*, 5(1):1–17, 1962.

[34] Van C Mow and W Michael Lai. Recent developments in synovial joint biomechanics. *Siam Review*, 22(3):275–317, 1980.

[35] Teruo MURAKAMI. The lubrication in natural synovial joints and joint prostheses. *JSME International Journal. Ser. 3, Vibration, Control Engineering, Engineering for Industry*, 33(4):465–474, 1990.

[36] Teruo Murakami, Hidehiko Higaki, Yoshinori Sawae, N Ohtsuki, S Moriyama, and Y Nakanishi. Adaptive multimode lubrication in natural synovial joints and artificial joints. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 212(1):23–35, 1998.

[37] GA Ateshian. A theoretical formulation for boundary friction in articular cartilage. 1997.

[38] GA Ateshian, Huiqun Wang, and WM Lai. The role of interstitial fluid pressurization and surface porosities on the boundary friction of articular cartilage. 1998.

[39] H Forster and J Fisher. The influence of loading time and lubricant on the friction of articular cartilage. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 210(2):109–119, 1996.

[40] Sierra G Cook, Ya Guan, Noah J Pacifici, Cory N Brown, Evan Czako, Mihir S Samak, Lawrence J Bonassar, and Delphine Gourdon. Dynamics of synovial fluid aggregation under shear. *Langmuir*, 35(48):15887–15896, 2019.

[41] Karl Meyer, Elizabeth M Smyth, and Martin H Dawson. The isolation of a mucopolysaccharide from synovial fluid. *Journal of Biological Chemistry*, 128(1):319–327, 1939.

[42] Edward D Bonnevie, Devis Galesso, Cynthia Secchieri, Itai Cohen, and Lawrence J Bonassar. Elastoviscous transitions of articular cartilage reveal a mechanism of synergy between lubricin and hyaluronic acid. *PloS One*, 10(11):e0143415, 2015.

[43] David W Smith, Bruce S Gardiner, Lihai Zhang, and Alan J Grodzinsky. *Articular cartilage dynamics*. Springer, 2019.

[44] Alice J Sophia Fox, Asheesh Bedi, and Scott A Rodeo. The basic science of articular cartilage: structure, composition, and function. *Sports Health*, 1(6):461–468, 2009.

[45] Sabrina Jahn, Jasmine Seror, and Jacob Klein. Lubrication of articular cartilage. *Annual Review of Biomedical Engineering*, 18:235–258, 2016.

[46] JinJing Liao, David W Smith, Saeed Miramini, Bruce S Gardiner, and Lihai Zhang. Investigation of role of cartilage surface polymer brush border in lubrication of biological joints. *Friction*, pages 1–18, 2021.

[47] Taku Saito. The superficial zone of articular cartilage. *Inflammation and Regeneration*, 42(1):1–6, 2022.

[48] A Robin Poole, Toshi Kojima, Tadashi Yasuda, Fackson Mwale, Masahiko Kobayashi, and Sheila Laverty. Composition and structure of articular cartilage: a template for tissue repair. *Clinical Orthopaedics and Related Research*, 391:S26–S33, 2001.

[49] James T Cassidy, Ross E Petty, Ronald M Laxer, and Carol B Lindsley. *Textbook of pediatric rheumatology E-Book*. Elsevier Health Sciences, 2010.

[50] L Martin-Alarcon and TA Schmidt. Rheological effects of macromolecular interactions in synovial fluid. *Biorheology*, 53(2):49–67, 2016.

[51] H Fam, JT Bryant, and M Kontopoulou. Rheological properties of synovial fluids. *Biorheology*, 44(2):59–74, 2007.

[52] AÖ Bingöl, D Lohmann, K Püschel, and W-M Kulicke. Characterization and comparison of shear and extensional flow of sodium hyaluronate and human synovial fluid. *Biorheology*, 47(3-4):205–224, 2010.

[53] AJ Palfrey and M Newton. The viscosity of synovial fluid at high shear rates. *Journal of Anatomy*, 106(Pt 2):404–404, 1970.

[54] D Prekasan and KK Saju. Review of the tribological characteristics of synovial fluid. *Procedia Technology*, 25:1170–1174, 2016.

[55] Tamer Mahmoud Tamer. Hyaluronan and synovial joint: function, distribution and healing. *Interdisciplinary Toxicology*, 6(3):111, 2013.

[56] Karl Meyer and John W Palmer. The polysaccharide of the vitreous humor. *Journal of Biological Chemistry*, 107(3):629–634, 1934.

[57] Panyakorn Taweechat, Ras B Pandey, and Pornthep Sompornpisut. Conformation, flexibility and hydration of hyaluronic acid by molecular dynamics simulations. *Carbohydrate Research*, 493:108026, 2020.

[58] Eleni Papakonstantinou, Michael Roth, and George Karakiulakis. Hyaluronic acid: A key molecule in skin aging. *Dermato-endocrinology*, 4(3):253–258, 2012.

[59] Paul Smith, Robert M Ziolek, Elena Gazzarrini, Dylan M Owen, and Christian D Lorenz. On the interaction of hyaluronic acid with synovial fluid lipid membranes. *Physical Chemistry Chemical Physics*, 21(19):9845–9857, 2019.

[60] Balaji Nagarajan, Samuel G Holmes, Nehru Viji Sankaranarayanan, and Umesh R Desai. Molecular dynamics simulations to understand glycosaminoglycan interactions in the free-and protein-bound states. *Current Opinion in Structural Biology*, 74:102356, 2022.

[61] István Hargittai and Magdolna Hargittai. Molecular structure of hyaluronan: an introduction. *Structural Chemistry*, 19:697–717, 2008.

[62] P. A. NETTI L. AMBROSIO, A. BORZACCHIELLO and L. NICOLAIS. Rheological study on hyaluronic acid and its derivative solutions. *Journal of Macromolecular Science, Part A*, 36(7-8):991–1000, 1999.

[63] Jihoon Kim, Ji-Youn Chang, Yoon-Young Kim, Moon-Jong Kim, and Hong-Seop Kho. Effects of molecular weight of hyaluronic acid on its viscosity and enzymatic activities of lysozyme and peroxidase. *Archives of Oral Biology*, 89:55–64, 2018.

[64] Tamaki Miyazaki, Chikako Yomota, and Satoshi Okada. Change in molecular weight of hyaluronic acid during measurement with a cone-plate rotational viscometer. *Journal of Applied Polymer Science*, 67(13):2199–2206, 1998.

[65] Atoosa Maleki, Anna-Lena Kjøniksen, and Bo Nyström. Effect of ph on the behavior of hyaluronic acid in dilute and semidilute aqueous solutions. In *Macromolecular symposia*, volume 274, pages 131–140. Wiley Online Library, 2008.

[66] Jason A Burdick and Glenn D Prestwich. Hyaluronic acid hydrogels for biomedical applications. *Advanced Materials*, 23(12):H41–H56, 2011.

[67] Petr Snetkov, Kseniia Zakharova, Svetlana Morozkina, Roman Olekhnovich, and Mayya Uspenskaya. Hyaluronic acid: The influence of molecular weight on structural, physical, physico-chemical, and degradable properties of biopolymer. *Polymers*, 12(8):1800, 2020.

[68] DA Swann, FH Silver, HS Slayter, W Stafford, and E Shore. The molecular structure and lubricating activity of lubricin isolated from bovine and human synovial fluids. *Biochemical Journal*, 225(1):195–201, 1985.

[69] Liaqat Ali, Sarah A Flowers, Chunsheng Jin, Eric Paul Bennet, Anna-Karin H Ekwall, and Niclas G Karlsson. The o-glycomap of lubricin, a novel mucin responsible for joint lubrication, identified by site-specific glycopeptide analysis. *Molecular & Cellular Proteomics*, 13(12):3396–3409, 2014.

[70] Khaled A Elsaid, Gregory D Jay, Ru Liu-Bryan, and Robert Terkeltaub. Proteoglycan 4 (prg4)/lubricin and the extracellular matrix in gout. *Gout, Urate, and Crystal Deposition Disease*, 1(3):122–136, 2023.

[71] Federico Roggio, Luca Petrigna, Bruno Trovato, Michelino Di Rosa, and Giuseppe Musumeci. The role of lubricin, irisin and exercise in the prevention and treatment of osteoarthritis. *International Journal of Molecular Sciences*, 24(6):5126, 2023.

[72] Xiaoyan Liu and Per M Claesson. Bioinspired bottlebrush polymers for aqueous boundary lubrication. *Polymers*, 14(13):2724, 2022.

[73] George W Greene, Rajiv Thapa, Stephen A Holt, Xiaoen Wang, Christopher J Garvey, and Rico F Tabor. Structure and property changes in self-assembled lubricin layers induced by calcium ion interactions. *Langmuir*, 33(10):2559–2570, 2017.

[74] Roberto C Andresen Eguiluz, Sierra G Cook, Cory N Brown, Fei Wu, Noah J Pacifici, Lawrence J Bonassar, and Delphine Gourdon. Fibronectin mediates enhanced wear protection of lubricin during shear. *Biomacromolecules*, 16(9):2884–2894, 2015.

[75] Bruno Zappone, George W Greene, Emin Oroudjev, Gregory D Jay, and Jacob N Israelachvili. Molecular aspects of boundary lubrication by human lubricin: effect of disulfide bonds and enzymatic digestion. *Langmuir*, 24(4):1495–1508, 2008.

[76] Sarah A Flowers, Agata Zieba, Jessica Örnros, Chunsheng Jin, Ola Rolfson, Lena I Björkman, Thomas Eisler, Sebastian Kalamajski, Masood Kamali-Moghaddam, and Niclas G Karlsson. Lubricin binds cartilage proteins, cartilage oligomeric matrix protein, fibronectin and collagen ii at the cartilage surface. *Scientific Reports*, 7(1):13149, 2017.

[77] Aled RC Jones, Jason P Gleghorn, Clare E Hughes, Lori J Fitz, Richard Zollner, Shane D Wainwright, Bruce Caterson, Elisabeth A Morris, Lawrence J Bonassar, and Carl R Flannery. Binding and localization of recombinant lubricin to articular cartilage surfaces. *Journal of Orthopaedic Research*, 25(3):283–292, 2007.

[78] George W Greene, Lisandra L Martin, Rico F Tabor, Agnes Michalczyk, Leigh M Ackland, and Roger Horn. Lubricin: A versatile, biological anti-adhesive with properties comparable to polyethylene glycol. *Biomaterials*, 53:127–136, 2015.

[79] GE Nugent-Derfus, AH Chan, BL Schumacher, and RL Sah. Prg4 exchange between the articular cartilage surface and synovial fluid. *Journal of Orthopaedic Research*, 25(10):1269–1276, 2007.

[80] Kirk J Samaroo, Mingchee Tan, Roberto C Andresen Eguiluz, Delphine Gourdon, David Putnam, and Lawrence J Bonassar. Tunable lubricin-mimetics for boundary lubrication of cartilage. *Biotribology*, 9:18–23, 2017.

[81] Ruby P Estrella, John M Whitelock, Nicolle H Packer, and Niclas G Karlsson. The glycosylation of human synovial lubricin: implications for its role in inflammation. *Biochemical Journal*, 429(2):359–367, 2010.

[82] Heidi L Reesink, Edward D Bonnevie, Sherry Liu, Carolyn R Shurer, Michael J Hollander, Lawrence J Bonassar, and Alan J Nixon. Galectin-3 binds to lubricin and reinforces the lubricating boundary layer of articular cartilage. *Scientific Reports*, 6(1):25463, 2016.

[83] Hari G Garg, David A Swann, and Lowrie R Glasgow. The structure of the o-glycosylically-linked oligosaccharide chains of lpg-i, a glycoprotein present in articular lubricating fraction of bovine synovial fluid. *Carbohydrate Research*, 78(1):79–88, 1980.

[84] Sarah A Flowers, Catherine S Lane, and Niclas G Karlsson. Deciphering isomers with a multiple reaction monitoring method for the complete detectable o-glycan repertoire of the candidate therapeutic, lubricin. *Analytical chemistry*, 91(15):9819–9827, 2019.

[85] Gregory D Jay, Darcy A Harris, and Chung-Ja Cha. Boundary lubrication by lubricin is mediated by o-linked β (1-3) gal-galnac oligosaccharides. *Glycoconjugate Journal*, 18:807–815, 2001.

[86] E Svala, C Jin, U Rüetschi, S Ekman, A Lindahl, NG Karlsson, and E Skiöldebrand. Characterisation of lubricin in synovial fluid from horses with osteoarthritis. *Equine Veterinary Journal*, 49(1):116–123, 2017.

[87] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

[88] A Keith Dunker, Celeste J Brown, J David Lawson, Lilia M Iakoucheva, and Zoran Obradović. Intrinsic disorder and protein function. *Biochemistry*, 41(21):6573–6582, 2002.

[89] A Keith Dunker, J David Lawson, Celeste J Brown, Ryan M Williams, Pedro Romero, Jeong S Oh, Christopher J Oldfield, Andrew M Campen, Catherine M Ratliff, Kerry W Hipps, et al. Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling*, 19(1):26–59, 2001.

[90] Pedro Romero, Zoran Obradovic, Xiaohong Li, Ethan C Garner, Celeste J Brown, and A Keith Dunker. Sequence complexity of disordered protein. *Proteins: Structure, Function, and Bioinformatics*, 42(1):38–48, 2001.

[91] Antonio Deiana, Sergio Forcelloni, Alessandro Porrello, and Andrea Giansanti. Intrinsically disordered proteins and structured proteins with intrinsically disordered regions have different functional roles in the cell. *PLoS One*, 14(8):e0217889, 2019.

[92] Robin Van Der Lee, Marija Buljan, Benjamin Lang, Robert J Weatheritt, Gary W Daughdrill, A Keith Dunker, Monika Fuxreiter, Julian Gough, Joerg Gsponer, David T Jones, et al. Classification of intrinsically disordered regions and proteins. *Chemical Reviews*, 114(13):6589–6631, 2014.

[93] M Madan Babu, Robin van der Lee, Natalia Sanchez de Groot, and Jörg Gsponer. Intrinsically disordered proteins: regulation and disease. *Current Opinion in Structural Biology*, 21(3):432–440, 2011.

[94] Frederik Lermyte. Roles, characteristics, and analysis of intrinsically disordered proteins: A minireview. *Life*, 10(12):320, 2020.

[95] Robert Schneider, Damien Maurin, Guillaume Communie, Jaka Kragelj, D Flemming Hansen, Rob WH Ruigrok, Malene Ringkjøbing Jensen, and Martin Blackledge. Visualizing the molecular recognition trajectory of an intrinsically disordered protein using multinuclear relaxation dispersion nmr. *Journal of the American Chemical Society*, 137(3):1220–1229, 2015.

[96] Christopher M Baker and Robert B Best. Insights into the binding of intrinsically disordered proteins from molecular dynamics simulation. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(3):182–198, 2014.

[97] Wenning Wang. Recent advances in atomic molecular dynamics simulation of intrinsically disordered proteins. *Physical Chemistry Chemical Physics*, 23(2):777–784, 2021.

[98] Andrew R Leach. *Molecular modelling: principles and applications*. Pearson Education, 2001.

[99] Daan Frenkel and Berend Smit. *Understanding molecular simulation: from algorithms to applications*. Academic Press San Diego, 2002.

[100] Michael P Allen and Dominic J Tildesley. *Computer simulation in chemical physics*, volume 397. Springer Science & Business Media, 2012.

[101] Harold A Scheraga, Mey Khalili, and Adam Liwo. Protein-folding dynamics: overview of molecular simulation techniques. *Annu. Rev. Phys. Chem.*, 58:57–83, 2007.

[102] Ashutosh Singh, Sai Kranthi Vanga, Valerie Orsat, and Vijaya Raghavan. Application of molecular dynamic simulation to study food proteins: A review. *Critical Reviews in Food Science and Nutrition*, 58(16):2779–2789, 2018.

[103] Eric Paquet, Herna L Viktor, et al. Molecular dynamics, monte carlo simulations, and langevin dynamics: a computational review. *BioMed Research International*, 2015, 2015.

[104] Berni Julian Alder and Thomas Everett Wainwright. Phase transition for a hard sphere system. *The Journal of Chemical Physics*, 27(5):1208–1209, 1957.

[105] J Andrew McCammon, Bruce R Gelin, and Martin Karplus. Dynamics of folded proteins. *Nature*, 267(5612):585–590, 1977.

[106] Max Born. Born-oppenheimer approximation. *Ann. Phys*, 84:457–484, 1927.

[107] Paul Ehrenfest. Bemerkung über die angenäherte gültigkeit der klassischen mechanik innerhalb der quantenmechanik. *Zeitschrift für Physik*, 45(7-8):455–457, 1927.

[108] van der Spoel D. Lindahl E. Hess B. Abraham, M. and the GROMACS development team. Gromacs 2023.3 manual. 2023.

[109] Miguel A González. Force fields and molecular dynamics simulations. *École thématique de la Société Française de la Neutronique*, 12:169–200, 2011.

[110] Lauren Boldon, Fallon Laliberte, and Li Liu. Review of the fundamental theories behind small angle x-ray scattering, molecular dynamics simulations, and relevant integrated application. *Nano Reviews*, 6(1):25661, 2015.

[111] Viktor Hornak, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3):712–725, 2006.

[112] Bernard R Brooks, Charles L Brooks III, Alexander D Mackerell Jr, Lennart Nilsson, Robert J Petrella, Benoît Roux, Youngdo Won, Georgios Archontis, Christian Bartels, Stefan Boresch, et al. Charmm: the biomolecular simulation program. *Journal of Computational Chemistry*, 30(10):1545–1614, 2009.

[113] William L Jorgensen, David S Maxwell, and Julian Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45):11225–11236, 1996.

[114] Chris Oostenbrink, Alessandra Villa, Alan E Mark, and Wilfred F Van Gunsteren. A biomolecular force field based on the free enthalpy of hydration and solvation: the gromos force-field parameter sets 53a5 and 53a6. *Journal of Computational Chemistry*, 25(13):1656–1676, 2004.

[115] Robert B Best and Gerhard Hummer. Optimized molecular dynamics force fields applied to the helix- coil transition of polypeptides. *The Journal of Physical Chemistry B*, 113(26):9004–9015, 2009.

[116] Abil E Aliev, Martin Kulke, Harmeet S Khaneja, Vijay Chudasama, Tom D Sheppard, and Rachel M Lanigan. Motional timescale predictions by molecular dynamics simulations: case study using proline and hydroxyproline sidechain dynamics. *Proteins: Structure, Function, and Bioinformatics*, 82(2):195–215, 2014.

[117] Karl N Kirschner, Austin B Yongye, Sarah M Tschampel, Jorge González-Outeiriño, Charlisa R Daniels, B Lachele Foley, and Robert J Woods. Glycam06: a generalizable biomolecular force field. carbohydrates. *Journal of Computational Chemistry*, 29(4):622–655, 2008.

[118] Roger W Hockney and James W Eastwood. *Computer simulation using particles*. Crc Press, 2021.

[119] Daan Frenkel and Berend Smit. *Understanding molecular simulation: from algorithms to applications*. Elsevier, 2023.

[120] Roger Williams Hockney, SP Goel, and JW Eastwood. Quiet high-resolution computer models of a plasma. *Journal of Computational Physics*, 14(2):148–158, 1974.

[121] Berk Hess, Henk Bekker, Herman JC Berendsen, and Johannes GEM Fraaije. Lincs: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12):1463–1472, 1997.

[122] Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman JC 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, 1977.

[123] Shuichi Miyamoto and Peter A Kollman. Settle: An analytical version of the shake and rattle algorithm for rigid water models. *Journal of Computational Chemistry*, 13(8):952–962, 1992.

[124] Loup Verlet. Computer" experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical Review*, 159(1):98, 1967.

[125] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An n log n method for ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089–10092, 1993.

[126] 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):8577–8593, 1995.

[127] Jaewoon Jung, Chigusa Kobayashi, Kento Kasahara, Cheng Tan, Akiyoshi Kuroda, Kazuo Minami, Shigeru Ishiduki, Tatsuo Nishiki, Hikaru Inoue, Yutaka Ishikawa, et al. New parallel computing algorithm of molecular dynamics for extremely huge scale biological systems. *Journal of Computational Chemistry*, 42(4):231–241, 2021.

[128] Herman JC Berendsen, JPM van Postma, Wilfred F Van Gunsteren, ARHJ DiNola, and Jan R Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684–3690, 1984.

[129] Shūichi Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 52(2):255–268, 1984.

[130] William G Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31(3):1695, 1985.

[131] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1), 2007.

[132] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182–7190, 1981.

[133] Mattia Bernetti and Giovanni Bussi. Pressure control using stochastic cell rescaling. *The Journal of Chemical Physics*, 153(11), 2020.

[134] AW Lees and SF Edwards. The computer study of transport processes under extreme conditions. *Journal of Physics C: Solid State Physics*, 5(15):1921, 1972.

[135] Sebastian Bindgen, Florian Weik, Rudolf Weeber, Erin Koos, and Pierre de Buyl. Lees–edwards boundary conditions for translation invariant shear flow: Implementation and transport properties. *Physics of Fluids*, 33(8), 2021.

[136] Denis J Evans and GP Morriss. Nonlinear-response theory for steady planar couette flow. *Physical Review A*, 30(3):1528, 1984.

[137] Nathan Berry, Yonghao Zhang, and Sina Haeri. Lees-edwards boundary conditions for the multi-sphere discrete element method. *Powder Technology*, 389:292–308, 2021.

[138] Thomas A Hunt and BD Todd. On the arnold cat map and periodic boundary conditions for planar elongational flow. *Molecular Physics*, 101(23-24):3445–3454, 2003.

[139] Sanjeev R Rastogi and Norman J Wagner. A parallel algorithm for lees-edwards boundary conditions. *Parallel Computing*, 22(6):895–901, 1996.

[140] Dabir S Viswanath, Tushar K Ghosh, Dasika HL Prasad, Nidamarty VK Dutt, and Kalipatnapu Y Rani. *Viscosity of liquids: theory, estimation, experiment, and data*. Springer Science & Business Media, 2007.

[141] Rajendra P Chhabra. Non-newtonian fluids: an introduction. *Rheology of Complex Fluids*, pages 3–34, 2010.

[142] RF Brooks, AT Dinsdale, and PN Quested. The measurement of viscosity of alloys—a review of methods, data and models. *Measurement Science and Technology*, 16(2):354, 2005.

[143] AT Dinsdale and PN Quested. The viscosity of aluminium and its alloys–a review of data and models. *Journal of Materials Science*, 39:7221–7228, 2004.

[144] Edward J Maginn, Richard A Messerly, Daniel J Carlson, Daniel R Roe, and J Richard Elliot. Best practices for computing transport properties 1. self-diffusivity and viscosity from equilibrium molecular dynamics [article v1. 0]. *Living Journal of Computational Molecular Science*, 1(1):6324–6324, 2019.

[145] Berk Hess. Determining the shear viscosity of model liquids from molecular dynamics simulations. *The Journal of Chemical Physics*, 116(1):209–217, 2002.

[146] Andrew Zgorski, Richard W Pastor, and Edward Lyman. Surface shear viscosity and interleaflet friction from nonequilibrium simulations of lipid bilayers. *Journal of Chemical Theory and Computation*, 15(11):6471–6481, 2019.

[147] Melville S. Green. Markoff random processes and the statistical mechanics of time-dependent phenomena. II. Irreversible processes in fluids. *The Journal of Chemical Physics*, 22(3):398–413, March 1954.

[148] Ryogo Kubo. Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems. *Journal of the Physical Society of Japan*, 12(6):570–586, June 1957.

[149] Yong Zhang, Akihito Otani, and Edward J Maginn. Reliable viscosity calculation from equilibrium molecular dynamics simulations: A time decomposition method. *Journal of chemical theory and computation*, 11(8):3537–3546, 2015.

[150] Sören von Bülow, Marc Siggel, Max Linke, and Gerhard Hummer. Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. *Proceedings of the National Academy of Sciences*, 116(20):9843–9852, 2019.

[151] Tobias M Prass, Patrick Garidel, Michaela Blech, and Lars V Schäfer. Viscosity prediction of high-concentration antibody solutions with atomistic simulations. *Journal of Chemical Information and Modeling*, 63(19):6129–6140, 2023.

[152] Xiao Dingcheng, Deng Yipan, Wang Zhenyao, Li Tianhao, and Liu Yinshui. Viscosity evolution of water glycol in deep-sea environment at high pressure and low temperature. *Journal of Molecular Liquids*, 387:122387, 2023.

[153] JP Ewen, DM Heyes, and Daniele Dini. Advances in nonequilibrium molecular dynamics simulations of lubricants and additives. *Friction*, 6:349–386, 2018.

[154] Michael Rubinsten. *Polymer physics*. United States of America, 2003.

[155] Mads Nygaard, Birthe B Kragelund, Elena Papaleo, and Kresten Lindorff-Larsen. An efficient method for estimating the hydrodynamic radius of disordered protein conformations. *Biophysical Journal*, 113(3):550–557, 2017.

[156] Hsiao-Ping Hsu, Wolfgang Paul, and Kurt Binder. Standard definitions of persistence length do not describe the local “intrinsic” stiffness of real polymer chains. *Macromolecules*, 43(6):3094–3102, 2010.

[157] Jing-Zi Zhang, Xiang-Yao Peng, Shan Liu, Bang-Ping Jiang, Shi-Chen Ji, and Xing-Can Shen. The persistence length of semiflexible polymers in lattice monte carlo simulations. *Polymers*, 11(2):295, 2019.

[158] Matthias Brosz, Nicholas Michelarakis, Uwe HF Bunz, Camilo Aponte-Santamaría, and Frauke Gräter. Martini 3 coarse-grained force field for poly (para-phenylene ethynylene) s. *Physical Chemistry Chemical Physics*, 24(17):9998–10010, 2022.

[159] Helman Amaya-Espinosa, Alfredo Alexander-Katz, and Camilo Aponte-Santamaría. The interplay between adsorption and aggregation of von willebrand factor chains in shear flows. *Biophysical Journal*, 122(19):3831–3842, 2023.

[160] Frank Eisenhaber, Philip Lijnzaad, Patrick Argos, Chris Sander, and Michael Scharf. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *Journal of Computational Chemistry*, 16(3):273–284, 1995.

[161] Robert Fitzpatrick, Davide Michieletto, Karthik R Peddireddy, Cole Hauer, Carl Kyrillos, Bekele J Gurmessa, and Rae M Robertson-Anderson. Synergistic interactions between dna and actin trigger emergent viscoelastic behavior. *Physical Review Letters*, 121(25):257801, 2018.

[162] Dongqing Wei and GN Patey. Orientational order in simple dipolar liquids: computer simulation of a ferroelectric nematic phase. *Physical Review Letters*, 68(13):2043, 1992.

[163] R Eppenga and Daan Frenkel. Monte carlo study of the isotropic and nematic phases of infinitely thin hard platelets. *Molecular Physics*, 52(6):1303–1334, 1984.

[164] Hongli Liu, Haiyang Zhong, Huanxiang Liu, and Xiaojun Yao. Molecular dynamics simulations reveal the disruption mechanism of a 2, 4-thiazolidinedione derivative c30 against tau hexapeptide (phf6) oligomer. *Proteins: Structure, Function, and Bioinformatics*, 90(1):142–154, 2022.

[165] Stefano Piana, Alexander G Donchev, Paul Robustelli, and David E Shaw. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *The Journal of Physical Chemistry B*, 119(16):5113–5123, 2015.

[166] David A Swann, Henry S Slayter, and Fred H Silver. The molecular structure of lubricating glycoprotein-i, the boundary lubricant for articular cartilage. *Journal of Biological Chemistry*, 256(11):5921–5925, 1981.

[167] Isabel M Martin, Camilo Aponte-Santamaría, Lisa Schmidt, Marius Hedtfeld, Adel Iusupov, Andrea Musacchio, and Frauke Gräter. Phosphorylation tunes elongation propensity and cohesiveness of incenp's intrinsically disordered region. *Journal of Molecular Biology*, 434(1):167387, 2022.

[168] Fan Jin and Frauke Gräter. How multisite phosphorylation impacts the conformations of intrinsically disordered proteins. *PLoS Computational Biology*, 17(5):e1008939, 2021.

[169] Austen Bernardi, Roland Faller, Dirk Reith, and Karl N Kirschner. Acpype update for nonuniform 1–4 scale factors: Conversion of the glycam06 force field from amber to gromacs. *SoftwareX*, 10:100241, 2019.

[170] Alan W Sousa da Silva and Wim F Vranken. Acpype-antechamber python parser interface. *BMC Research Notes*, 5:1–8, 2012.

[171] Marcus D Hanwell, Donald E Curtis, David C Lonie, Tim Vandermeersch, Eva Zurek, and Geoffrey R Hutchison. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, 4(1):1–17, 2012.

[172] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C Smith, Berk Hess, and Erik Lindahl. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1:19–25, 2015.

[173] Berk Hess. P-lincs: A parallel linear constraint solver for molecular simulation. *Journal of Chemical Theory and Computation*, 4(1):116–122, 2008.

[174] Jonathan E Kohn, Ian S Millett, Jaby Jacob, Bojan Zagrovic, Thomas M Dillon, Nikolina Cingel, Robin S Dothager, Soenke Seifert, P Thiyagarajan, Tobin R Sosnick, et al. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proceedings of the National Academy of Sciences*, 101(34):12491–12496, 2004.

[175] Alexandra Brito, Dhwanit Dave, Ayala Lampel, Vânia IB Castro, Daniela Kroiss, Rui L Reis, Tell Tuttle, Rein V Ulijn, Ricardo A Pires, and Iva Pashkuleva. Expanding the conformational landscape of minimalistic tripeptides by their o-glycosylation. *Journal of the American Chemical Society*, 143(47):19703–19710, 2021.

[176] Estella A Newcombe, Elise Delaforge, Rasmus Hartmann-Petersen, Karen Skriver, and Birthe B Kragelund. How phosphorylation impacts intrinsically disordered proteins and their function. *Essays in Biochemistry*, 66(7):901–913, 2022.

[177] Izzy Owen and Frank Shewmaker. The role of post-translational modifications in the phase transitions of intrinsically disordered proteins. *International Journal of Molecular Sciences*, 20(21):5501, 2019.

[178] Gul H Zerze and Jeetain Mittal. Effect of o-linked glycosylation on the equilibrium structural ensemble of intrinsically disordered polypeptides. *The Journal of Physical Chemistry B*, 119(51):15583–15592, 2015.

[179] Paulo CT Souza, Riccardo Alessandri, Jonathan Barnoud, Sebastian Thallmair, Ignacio Faustino, Fabian Grünewald, Ilias Patmanidis, Haleh Abdizadeh, Bart MH Bruininks, Tsjerk A Wassenaar, et al. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nature methods*, 18(4):382–388, 2021.

[180] Giulio Tesei, Thea K Schulze, Ramon Crehuet, and Kresten Lindorff-Larsen. Accurate model of liquid–liquid phase behavior of intrinsically disordered proteins from optimization of single-chain properties. *Proceedings of the National Academy of Sciences*, 118(44):e2111696118, 2021.

[181] Fan Cao, Sören von Bülow, Giulio Tesei, and Kresten Lindorff-Larsen. A coarse-grained model for disordered and multi-domain proteins. *BioRxiv*, pages 2024–02, 2024.

[182] S Balasubramanian, Christopher J Mundy, and Michael L Klein. Shear viscosity of polar fluids: Molecular dynamics calculations of water. *The Journal of Chemical Physics*, 105(24):11190–11195, 1996.

[183] A Ya Malkin. Non-newtonian viscosity in steady-state shear flows. *Journal of Non-Newtonian Fluid Mechanics*, 192:48–65, 2013.

[184] Pedram Tazraei, Alireza Riasi, and Behrouz Takabi. The influence of the non-newtonian properties of blood on blood-hammer through the posterior cerebral artery. *Mathematical Biosciences*, 264:119–127, 2015.

[185] Clementine Pradal, Gleb E Yakubov, Martin AK Williams, Michael A McGuckin, and Jason R Stokes. Lubrication by biomacromolecules: mechanisms and biomimetic strategies. *Bioinspiration & Biomimetics*, 14(5):051001, 2019.

[186] Zhixiang Cai, Hongbin Zhang, Yue Wei, Min Wu, and Ailing Fu. Shear-thinning hyaluronan-based fluid hydrogels to modulate viscoelastic properties of osteoarthritis synovial fluids. *Biomaterials Science*, 7(8):3143–3157, 2019.

[187] Monika Huber, Siegfried Trattnig, and Felix Lintner. Anatomy, biochemistry, and physiology of articular cartilage. *Investigative Radiology*, 35(10):573–580, 2000.

[188] Abhijit M Bhosale and James B Richardson. Articular cartilage: structure, injuries and review of management. *British Medical Bulletin*, 87(1):77–95, 2008.

[189] DR Eyre. The collagens of articular cartilage. In *Seminars in arthritis and rheumatism*, volume 21, pages 2–11. Elsevier, 1991.

[190] Yosadara Ruiz-Morales and Ascención Romero-Martínez. Coarse-grain molecular dynamics simulations to investigate the bulk viscosity and critical micelle concentration of the ionic surfactant sodium dodecyl sulfate (sds) in aqueous solution. *The Journal of Physical Chemistry B*, 122(14):3931–3943, 2018.

[191] Lingnan Lin and Mark A Kedzierski. Density and viscosity of a polyol ester lubricant: Measurement and molecular dynamics simulation. *International Journal of Refrigeration*, 118:188–201, 2020.

[192] Billy D Todd and Peter J Daivis. *Nonequilibrium molecular dynamics: theory, algorithms and applications*. Cambridge University Press, 2017.

[193] Dimitrios Mathas, Walter Holweger, Marcus Wolf, Christof Bohnert, Vasilios Bakolas, Joanna Procelewska, Ling Wang, Scott Bair, and Chris-Kriton Skylaris. Evaluation of methods for viscosity simulations of lubricants at different temperatures and pressures: a case study on pao-2. *Tribology Transactions*, 64(6):1138–1148, 2021.

[194] Kenneth Alambra. Water viscosity calculator.

[195] Tatiana I Morozova, Nicolás A García, and Jean-Louis Barrat. Temperature dependence of thermodynamic, dynamical, and dielectric properties of water models. *The Journal of Chemical Physics*, 156(12), 2022.

[196] Pierre J. Carreau. Rheological equations from molecular network theories. *Transactions of the Society of Rheology*, 16(1):99–127, March 1972.

[197] Malcolm M. Cross. Rheology of non-Newtonian fluids: A new flow equation for pseudoplastic systems. *Journal of Colloid Science*, 20(5):417–437, June 1965.

[198] Joseph Kestin, Mordechai Sokolov, and William A Wakeham. Viscosity of liquid water in the range- 8 c to 150 c. *Journal of Physical and Chemical Reference Data*, 7(3):941–948, 1978.

[199] Robert Zwanzig. Time-correlation functions and transport coefficients in statistical mechanics. *Annual Review of Physical Chemistry*, 16(1):67–102, 1965.

[200] Taryn E Ludwig, Mary K Cowman, Gregory D Jay, and Tannin A Schmidt. Effects of concentration and structure on proteoglycan 4 rheology and interaction with hyaluronan. *Biorheology*, 51(6):409–422, 2014.

[201] Samuel K Lai, Ying-Ying Wang, Denis Wirtz, and Justin Hanes. Micro-and macrorheology of mucus. *Advanced Drug Delivery Reviews*, 61(2):86–100, 2009.

[202] Preethi L Chandran and Ferenc Horkay. Aggrecan, an unusual polyelectrolyte: review of solution behavior and physiological implications. *Acta Biomaterialia*, 8(1):3–12, 2012.