ABSTRACT

Carbohydrates; in the form of oligosaccharides, polysaccharides and glycoconjugates like glycopeptides, glycoproteins, glycolipids, and peptidoglycans; are amongst the abundant cell components. Glycans or oligosaccharides are an amazingly diverse class of macromolecules, varying in length, sequence order, linkages, substitution, and branching. Carbohydrates interact with a wide range of protein families, including antibodies, lectins, chemokines, and sugar transporters. Consequently, protein-carbohydrate interactions mediate a wide range of cellular processes. Owing to their structural and functional diversity, carbohydrates represent both therapeutic targets and tools. For example, glycans decorating the eukaryotic cell surface are often the hallmark of cancer and inflammation, while aberrant O-glycosylation may result in neurodegenerative disorders like Alzheimer's, Parkinson's and Huntington's disease.

Thus, a plethora of biological processes are directly or indirectly regulated by carbohydrates or their glycoconjugates, but the exact mechanism of their involvement in most of the biological events are still not clear. Therefore, understanding the physical and conformational properties of carbohydrates alone and in combination with their molecular scaffolds is crucial to elucidate the mechanistic role of these macromolecules in biological processes. It will also help to exploit these biomolecules for their utilization as therapeutic agents. Recent advances in experimental techniques allow the synthesis of naturally occurring oligosaccharides, polysaccharides and glycoproteins. However, the inherent structural complexity and variable-length complicate the experimental structural characterization of these molecules. For example, the flexible ring conformations and the flexible glycosidic linkage between carbohydrates generally lead to several possible low energy conformers, resulting in less well-ordered electron density for the entire carbohydrate unit in X-ray crystallography. Therefore the experimental studies are generally augmented with theoretical methods like Quantum mechanical (QM) calculations or molecular dynamics (MD) simulations. This thesis contributes to the understanding of carbohydrates interactions with water or proteins. The thesis is divided into seven chapters.

In chapter 1, we provide a brief introduction to carbohydrates and a comprehensive review of earlier experimental and theoretical studies on carbohydrates. The theoretical foundation of the computational techniques used in the subsequent chapters are discussed in detail in chapter 2.

In chapter 3, we use molecular dynamics simulations to gain atomistic insight into carbohydrate-water interactions and to specifically highlight the differences between additive (non-polarizable) and polarizable simulations. A total of six monosaccharide systems, α and β anomers of glucose, galactose, and mannose, have been studied using additive and polarizable CHARMM carbohydrate force fields. The solvent was modeled using three additive water models TIP3P, TIP4P, and TIP5P in the additive simulations and the polarizable water model SWM4 in the polarizable

simulations. The presence of carbohydrate has a significant effect on the microscopic water structure with the effects being pronounced for the proximal water molecules. Notably, a disruption of the tetrahedral arrangement of proximal water molecules was observed due to the formation of strong carbohydrate-water hydrogen bonds in both the additive and polarizable simulations. However, the inclusion of polarization resulted in significant bridge water occupancies, improved ordered water structure (tetrahedral order parameter) and longer carbohydrate-water H-bond correlations as compared to the additive simulations. Additionally, the polarizable simulations also allowed the calculation of the power spectra from the dipole-dipole autocorrelation function, which corresponds to the IR spectra. From the power spectra we could identify spectral signatures differentiating the proximal and bulk water structures which could not be captured from additive simulations.

In chapter 4, we report the development of Drude polarizable force-field parameters for the carboxylate and N-acetyl amine monosaccharide derivatives, extending the functionality of existing Drude polarizable carbohydrate force field. The force field parameters are developed in a hierarchical manner, reproducing the quantum mechanical (QM) gas-phase properties of small model compounds representing the key functional group in the carbohydrate derivatives. The optimized parameters were then used to generate the models for carboxylate and N-acetyl amine carbohydrate derivatives. The transferred parameters were further tested and optimized to reproduce crystal geometries and J-coupling data from NMR experiments. The parameter development resulted in the incorporation of D-glucuronate, L-iduronate, N-acetyl-D-glucosamine (GlcNAc) and N-acetyl-D-galactosamine (GalNAc) sugars into the Drude polarizable force field. The parameters developed in this study were then applied to study the conformational properties of glycosaminoglycan polymer hyaluronan, composed of D-glucuronate and N-acetyl-D-glucosamine, in aqueous solution. Upon comparing the results from the additive and polarizable simulations, it was found that the inclusion of polarization improved the description of the electrostatic interactions observed in hyaluronan resulting in enhanced conformational flexibility. The parameters developed in this study enabled us to investigate the influence of polarization on the structural properties of water around carboxylate ($\alpha/\beta-$D-glucuronate and L-iduronate) and N-acetyl amine ($\alpha/\beta-$N-acetyl-D-glucosamine and N-acetyl-D-galactosamine) monosaccharide derivatives. Unlike the unsubstituted monosaccharide-water system, a clear retardation in H-bond dynamics was observed for 1^{st} hydration shell in polarizable systems, which is indistinguishable in additive simulations. We observed that the retardation dynamics also extended beyond the 1^{st} hydration shell. This effect was due to the presence of the additional H-bond donating and acceptor groups present in carboxylate (COO^-) and N-acetyl-amine (-$NHCOCH_3$) side chains. The influence of these side chains is also captured by the fluctuations in tetrahedral order parameter.

In chapter 5, we carried out MD simulation studies to probe the microscopic properties of water brought about by structural variations of antifreeze glycoproteins (AFGPs). AFGPs are a distinctively riveting class of bio-macromolecules, which endows the survival of organisms inhabiting polar and subpolar regions. These proteins are supposed to hinder the microscopic freezing by interacting with the embryonic ice crystals and precluding their further growth. The underlying molecular mechanism of AFGP binding to ice has remained elusive due to insufficient structural characterization, with conflicting hypothesis on the possible binding mode of AFGPs; either via the hydrophobic peptide backbone or via the hydrophilic carbohydrate side chains; when interacting with ice. Chemical synthesis has allowed researchers to access synthetic variants of natural AFGPs. These studies revealed that AFGPs exhibit huge variations in the thermal hysteresis and

ice shaping behavior on slight structural variations, especially to the carbohydrate side chains. Four key structural motifs were identified as crucial to AFGP activity; the presence of threonine γ-methyl group, α-glycosidic carbohydrate-protein linkage, acetylamide group (-NHCOCH3) at the C2 position of the carbohydrate linked to protein and the presence of carbohydrate hydroxyl groups. In this study, we use Molecular dynamics (MD) simulations to probe the microscopic properties of water along these structural variations of AFGPs. We find that these variations primarily influence the conformation space of AFGPs and also crucially control their hydration dynamics. Owing to the disordered nature of AFGPs, we use Markov-State modeling to identify conformational preferences for AFGPs. The simulations reveal the importance of steric bulk, intra-molecular carbohydrate-protein H-bonds, and conformational preferences ($\alpha-$ vs $\beta-$ linkages) in controlling the spatial segregation of hydrophilic and hydrophobic regions of AFGPs. We hypothesize that the hydrophobic component of AFGPs is crucial to its binding to ice, which determines the ice shaping ability of AFGPs. However, the hydrophilic carbohydrate hydroxyl groups and their ability to form bridge waters control the subsequent hydration dynamics, which is key to the antifreeze properties. Investigating the tetrahedral order parameter of water molecules around the carbohydrates revealed a competition between solute and bulk influenced the solvent structure, with maximum restructuring being observed in the interfacial region 2.5 Å – 4.5 Å away from the AFGPs.

In chapter 6, we report the development of CHARMM additive force-field parameters to enable the MD simulation studies of nucleotide-sugar. We used a hierarchical approach to develop the force field parameters. Initial parameters were adopted from the CHARMM nucleic acid and CHARMM carbohydrate force fields. Parameters were initially optimized using target data generated for small model compounds which were selected based on fragmentation approach. The optimized parameters were then used to generate the models for nucleotide sugars. The transferred parameters were tested and optimized to reproduce crystalline cell parameters and intra-molecular geometries. To ensure the compatibility of the developed parameters with the existing CHARMM additive force field, the parameters were also used to study the conformational properties of a family of nucleotide-sugar protein complexes. The developed force-field parameters in conjunction with the remainder of the CHARMM additive force-field parameters can be used for the future studies involving nucleotide-sugars in biological systems.

In chapter 7, we provide a overall summary of the work carried out in the thesis.