Abstract

Glycosaminoglycans (GAGs) are linear anionic periodic polysaccharides that play a crucial role in various biologically relevant functions within the extracellular matrix. Through their interactions with proteins, GAGs mediate processes such as cell proliferation, cancer development, inflammation and the onset of neurodegenerative diseases. Experimental approaches suffer from difficulties in the investigations of protein-GAG systems because of the complex nature of the GAGs. Computational studies proved to be helpful in addressing some of the challenges faced by experimental approaches. Nonetheless, GAGs have not received as much attention from the computational community as other biomolecule classes, leading to a lack of modeling tools specifically designed for their theoretical analysis. As a result, researchers, for example, must rely on existing docking software developed mainly for small drug molecules that differ significantly from GAGs in terms of their basic physico-chemical properties. It would be of great importance to develop and implement new tools that allow computational GAG community to study these biologically relevant molecules with the ease, precision and accuracy similar to the ones in computational studies of other groups of biomolecules. If the mentioned issues become solved theoretical approaches can not only complement the experimental studies but also successfully investigate areas that are not yet accessible for the experimental research.

The main goals of this PhD thesis were to develop new computational tools for GAG-containing molecular systems and examination of GAG interactions using computational approaches. To achieve these goals a set of theoretical approaches were designed and applied to specific biologically relevant systems involving GAG molecules.

First, the analysis and revision of currently available molecular docking tools was performed. This allowed for developing new approaches targeting analysis of GAG interactions. One of them is a coarse-grained model representing GAG monosaccharide units that has been developed to make the analysis more accessible for less experienced researchers in the computational field. In this approach, a ready-to-use script was provided for a user to elongate a GAG molecule in the analyzed complex. Then, a novel Molecular Dynamics (MD) based docking tool named RS-REMD (Repulsive Scaling Replica Exchange Molecular Dynamics) has been implemented for studies of protein-GAG systems. In this method, van der Waals radii are being scaled in each consecutive replica which allows for a faster sampling of the system. Next, improvements regarding the use of explicit water model have been implemented further enhancing accuracy of the RS-REMD technique. In the second part of the PhD thesis, interactions of GAG with proteins have been

investigated. In order to accomplish this, a representative nonredundant dataset of protein-GAG complexes has been analyzed. The effect of length of the GAGs on binding to protein has been studied complemented by comprehensive analysis of technical computational aspects regarding the performance of the MD simulations. Then, in the investigation of GAG influence on the APRIL (A proliferation-inducing ligand) protein and its receptors - TACI (Transmembrane activator and CAML interactor) and BCMA (B-cell maturation antigen) - new molecular mechanism of APRIL-receptor complex forming facilitated by GAG binding was proposed. At last, the role of water in GAG MD simulations has been studied. The influence of particular solvent models on highly sulfated GAGs has been described. It was shown that TIP5P and OPC water models performed essentially better than widely the used TIP3P model.

This PhD thesis presents data that expand the general understanding of GAG-related systems. Additionally, novel computational techniques were designed for GAG molecular docking approaches, in particular, and GAG system computational analysis strategies, in general. The results show the importance and high potential of theoretical approaches as powerful tools in studying protein-GAG interactions.