**Interactive exploration of ligand transportation through protein structures**

Katarína Furmanová, Miroslava Jarešová, Jan Byška, Adam Jurčík, Július Parulek, Helwig Hauser, and Barbora Kozlíková

Protein structures play a crucial role in all living organisms. Therefore, detailed understanding of their function and reactivity with other molecules have been in the scope of researchers for decades. The interaction between a protein and other molecule can undergo on the protein surface or deeply inside the protein structure. In the latter case the small molecule, called ligand, is transported to a specific buried site, called an active site. In this active site the ligand can react with the protein and the product of such reaction can form a basis of new chemical matters, e.g., drugs.

In this paper we focus on the transportation path of the ligand towards the protein active site. The path is derived from the simulation of molecular dynamics consisting of thousands of time steps. In consequence, the resulting path is very complex and the movement of the ligand is scattered. Such amount of information prevents the biochemist to reveal the significant movements of the ligand and important properties influencing the ligand behavior. Therefore, we propose a method for simplification of the ligand trajectory which aims to remove the scattered behavior of the ligand and convey only its significant shifts.

Additionally, we propose an exploration system for further exploration of the trajectory consisting of three main parts. First part contains the spatial representation of the ligand trajectory where the user can immediately see the influence of our simplification method and can further manually adjust the level of simplification on selected parts of the trajectory. The second view helps to understand different properties of the ligand along the trajectory, such as its speed, distance wrt. the active site, surrounding amino acids, or so called “stuckness” which defines the parts of the trajectory where the ligand got stuck for a significant portion of the simulation. This is reached by a combination of 2D bar charts representing the aggregated information along the trajectory and line representation showing the influence of individual amino acids on the ligand. These two views are further equipped with a scatterplot matrix where the user can plot different combinations of ligand properties and by interactive selections explore only the most interesting parts of the trajectory. These linked views provide the biochemist with fast navigation and better understanding of the ligand behavior along its trajectory.

The proposed system was designed in tight cooperation with the domain experts. Its usability is demonstrated on the exploration workflow supported by our system.