**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 another 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 study 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, in addition, the movement of the ligand is scattered containing small scale jittering motion as well. The shear amount of time steps accompanied by the complex motion prevent biochemists to reveal significant movement patterns of the ligand and important properties influencing the ligand behaviour. Therefore, we propose a method for simplification of the ligand trajectory, which aims to remove the scattered behaviour of the ligand and convey only its significant shifts.

Additionally, we propose an interactive visual analysis system for further exploration of the trajectory, which consists of three main parts. First part contains the spatial representation of the ligand trajectory where the user can directly examine the simplified trajectory 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” that 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 accompanied with a 2D scatterplot view where the user can plot different combinations of ligand properties and interactively select parameters that are linked with remaining views. This allows users to focus on exploration of the most interesting parts of the trajectory. These linked views provide the biochemist with fast navigation and better understanding of the ligand behaviour 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.