**Comparative Visualization of Protein Secondary Structures**

Lucia Kocincová, Miroslava Jarešová, Jan Byška, Július Parulek, Helwig Hauser, and Barbora Kozlíková

Understanding of protein structure helps the biochemists and biologists to reveal the behavior and function of these proteins. In consequence, this can lead to designing new chemical matters, including drugs. The structure of the protein is determined by the polypeptide (popypeptidic chain = 7 500 na Google, “popypeptide chain“ = 700k) chain consisting of a sequence of amino acids. This chain is folded into a three dimensional shape which is maintained by different types of weak chemical bonds. The behavior and function of the protein is highly influenced not only by the constitution of the chain but also by its 3D shape and changes over time within molecular dynamics simulations. Therefore, the biochemists and biologists are aiming to explore these properties in detail. Moreover, in their workflow also the comparison between different proteins plays a crucial role. By comparing the protein structures the domain experts can reveal structures belonging to the same protein family or even analyze the evolution of same protein in different species. The comparison process includes the algorithms for structural or sequential alignment of protein chains.

Currently there are several existing solutions for conveying the information about the protein chain, spanning from 1D sequential list of all amino acids in the chain to different spatial representations of the folded protein. In the latter case one of the most popular representations of the protein is a so called cartoon model (spíš bych použil “ribbon diagrams” bo když jsem psal disertaci, tak to tak většinou označovali) which conveys in a highly abstracted manner the information about so called protein secondary structures. These secondary structures, namely alpha helices and beta sheets, inform the user about the hydrogen bonding inside the protein. However, when comparing more proteins, the 1D representation lacks of the important information about mutual spatial position of the proteins. The 3D cartoon (<<pokud se rozhodneš změnit, tak tady taky) representation provides the biochemists with such information. However, it is not suitable for comparing more structures at once – the visual clutter and occlusion starts to prevent the understanding of the mutual position.

To overcome these problems we propose a novel representation on the boundary between the 1D and 3D representations which uses the best properties of both of them. The representation is based on the abstract view onto the protein chain which consists of individual secondary structures but is straightened to a sequential 1D chain. When comparing more proteins, this straightened representation aims to prevent the mutual position of corresponding secondary structures by maintaining the angle between them. In consequence, the novel representation helps to understand the structural and spatial information about mutual positions of aligned secondary structures which is traditionally present only in the 3D view.

Our proposed method was tested by the domain experts and the usability is presented on a case study performed on proteins from the cytochrome P450 family.