Comparative Visualization and Analysis of Secondary Structures

Pavol Ulbrich Advisor doc. RNDr. Barbora Kozlíková, Ph.D.

MASARYK UNIVERSITY

Motivation and Problem Statement

To better understand the constitution and spatial arrangement of protein sequences, L. Kocincová et al. proposed a novel method of comparative visualization, based on traditionally used 1D and 3D views. Its main contribution is the ability to observe the spatial differences between the proteins without any occlusion problems, commonly present in 3D view. This thesis aims to create a web application for comparative visualization of protein secondary structures, which will benefit from the qualities of this method. In addition, the tool will provide both sequential and structural protein alignment based on modern methods, which rely on evolutionarily-related residues and ensure a precise alignment of secondary structures.

Rendering and Alignment process

The newly proposed application is able to handle the uploaded Protein Data Bank (PDB) files, align them, and calculate the comparative visualization views upon them. The new modern approach of sequential and structural alignment calculation, **DeepAlign** for the pair-wise alignment and **3DCOMB** for the multiple alignment, was used in order to compute the alignment in real time. The application was built upon *Express.js* framework for *Node.js*. The pipeline on the server side can be seen on the image on the right. The front-end pipeline, written in *JavaScript* library *d3.js* can by observed above. The views display protein secondary structures (alpha-helixes and beta-sheets) as cartoon ribbons (spirals and arrows).
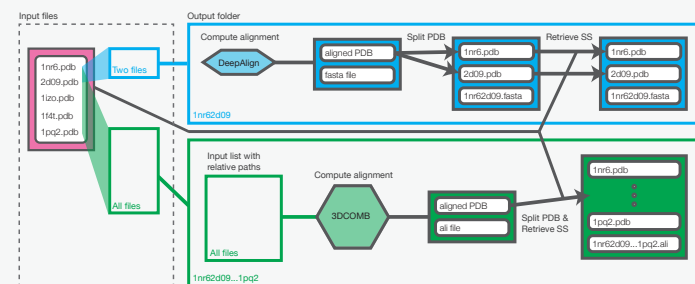
Comparative visualisations

To visualize the structure alignment and 3D models of compared proteins the **PV viewer** is used (left), demonstrating the structural alignment (spatial distance between proteins is minimised). Sequential alignment (inserting gaps to the protein chains in order to keep similar segments close together) can be seen on top right. Their combination is shown in the the **Juxtaposition** (right bottom) and the **Superimposition** (bottom) views, that demonstrate both gaps and distance between compared secondary structures. The outcome is an immediate observation of related and distant segments of proteins.
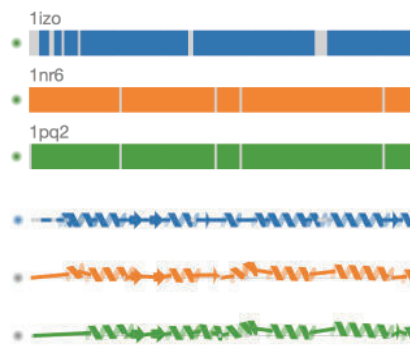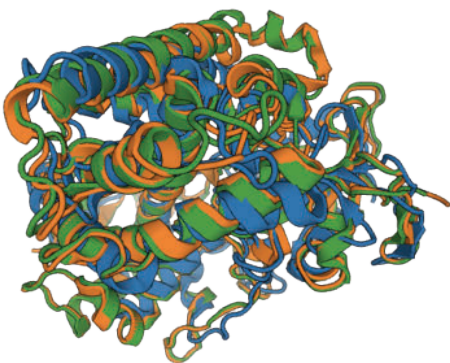
Interaction

The option to select a chosen secondary structure in any view and highlight it in all the views is the fundamental form of interaction. The selections can be easily deselected by clicking on the highlighted structure one more time. Another ways to interact with the application are removing/reinserting a protein into the PV viewer, changing the reference sequence for the comparative sequence alignment views, and adjusting the views. All the views can be zoomed, the PV viewer adds the possibility to rotate the model. The Juxtaposition view offers an innovative approach by defining the range and proteins that are to be zoomed.

Footnote

The application is available on the addres *http://decibel.fi.muni.cz/~xulbrich*. However, it is still in active development, which mostly consists of perfecting the features and user experience with biochemists of Lochsmidt laboratories of Faculty of Science, MU. The results are planned to be published in the *Nucleid Acid Research* journal.