

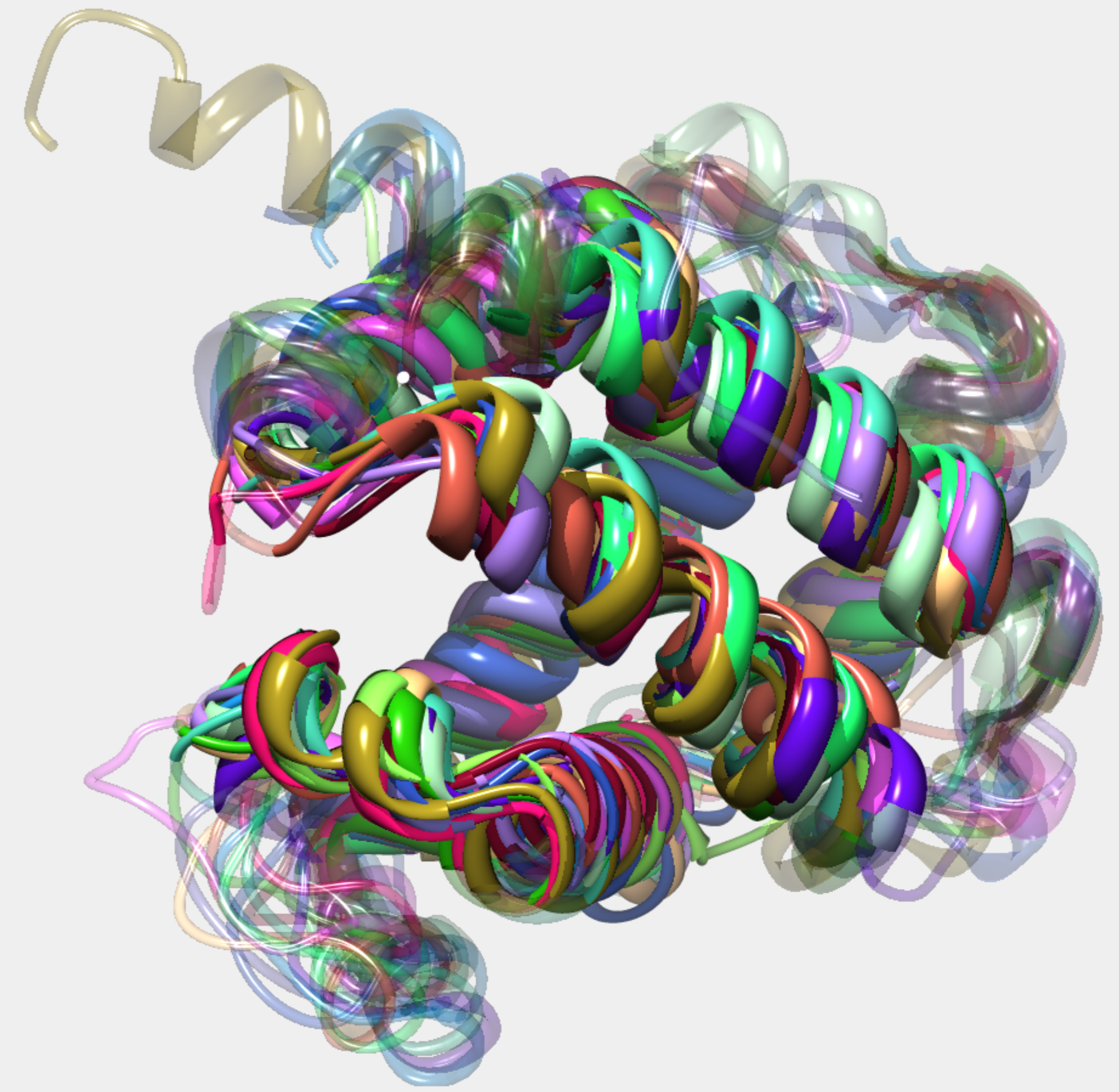


Background

Multiple alignments of protein structures are frequently used to investigate properties of protein families when sequence based methods are not applicable due to low sequence similarity. The resulting alignments are useful e.g. for inferring evolutionary processes, supporting drug design or depicting epitopes responsible for allergic cross reactions. In contrast to sequence alignments there are no such simple algorithms available to calculate optimal MStAs. MStA methods are rather based on heuristics and different measures of structural similarity. An evaluation of different MStA programs has shown that the produced alignments may differ considerably and that one method may fail to produce an alignment while another one can easily perform the task[1].

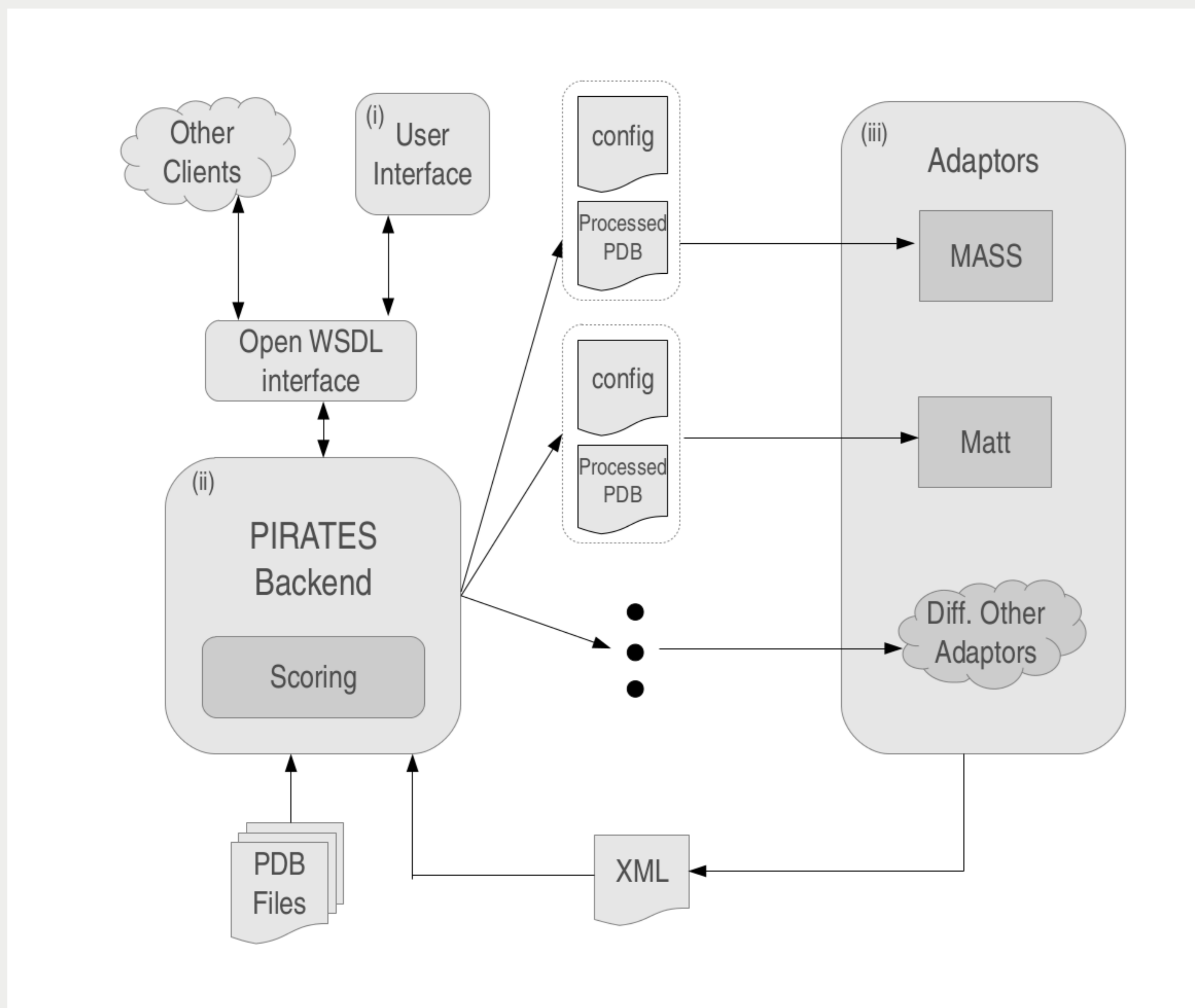
Example multiple structure alignment

MStA of a subset of the globin family. Although the sequence similarity is in the twilight zone, the structural core is well conserved (opaque). The proteins mainly differ in the loop regions (transparent). Multiple sequence alignments fail to provide a correct alignment in this case.



Implementation & Results

In order to easily access the different MStA methods and to compare and rank the individual results we implemented PIRATES in the course of an educational student project. PIRATES is a meta-server which sends preprocessed structure data to different MStA standalone programs and web servers, collects the resulting alignments and converts them into a convenient XML based format for further processing. The current version implements interfaces to three MStA methods (MASS[2], Matt[3], MUSTANG[4]), the job scheduling logics and the graphical user interface. PIRATES is implemented in C++. The server runs on a Debian Linux platform.



PIRATES consists of three main components: (i) the user interface, (ii) the PIRATES backend, and (iii) adapters to third party MStA software or webservices. The communication between the user interface and backbone is solved via WSDL standard. This enables access to the backend for other frontends or scripts.

Perspective

In the final version of PIRATES we will provide additional interfaces to MStA methods, a consensus alignment derived from the different method MStAs, and a 3D molecular graphics representation thereof in the GUI. The entire PIRATES source code will be available under an open source license and a freely accessible web application will be provided.

References

- [1] **Accuracy analysis of multiple structure alignments.** Beralk C *et al.* Protein Sci. 2009 Oct;18(10):2027-35. <http://biwww.che.sbg.ac.at/RSA/>
- [2] **MASS: Multiple structural alignment by secondary structures.** Dror O *et al.* Bioinformatics, 19 Suppl. 1: i95-i104, 2003. <http://bioinfo3d.cs.tau.ac.il/MASS/>
- [3] **Matt: Local Flexibility Aids Protein Multiple Structure Alignment.** Menke M *et al.* PLoS Comput Biol 4(1): e10. doi:10.1371/journal.pcbi.0040010 <http://groups.csail.mit.edu/cb/matt/>
- [4] **MUSTANG: A multiple structural alignment algorithm.** Arun S *et al.* Proteins: Structure, Function, and Bioinformatics. 64(3):559-574, August 2006. <http://www.csse.monash.edu.au/~karun/Site/mustang.html>

User Interface

PIRATES provides two User Interfaces:

- ▶ A graphical web interface where the settings for an alignment can be entered
- ▶ A Web Service Definition Language (WSDL) interface

Backend

- ▶ Preprocesses input data from the GUI and passes parameters and structure data to the adapters.
- ▶ Implements a scheduler for handling user requests in parallel.
- ▶ Evaluates the methods results and presents them to the user.

Adapters

PIRATES implements an adapter class for each alignment service. As a result the scheduler can easily start all programs with the same command but with individual settings.