

Work Package 7, Milestone 7.4: Workflow and input/output of HADDOCK web service

Authors

Mikael Trellet, Jörg Schaarschmidt and Alexandre M.J.J. Bonvin, Utrecht University.
October 22nd, 2016

Introduction

We present here the current workflow and input/output of the HADDOCK web server. The main workflow is presented as a diagram consisting of two interconnected parts. One focuses on the general workflow from the submission to the presentation of the final results whereas the second one makes a specific focus on the HADDOCK internal process itself. The hierarchy of the archive containing the input/output files are described in a distinct section.

HADDOCK workflow diagram

This diagram (see last page) reports the full process of HADDOCK process when used through its web server interface. Starting from the input data validation and parsing, then going through the HADDOCK internal docking simulations process to finally reach a complete analysis of the results. The main steps of this workflow can be monitored in the status page of a specific run and log files are accessible at any moment to the users.

Most of the errors that can intervene during the workflow are specifically handled and displayed accordingly to the users in order to easily detect the failure source.

All input and output files are kept all along the workflow to easily access to any step of HADDOCK process for further analyses and putative debugging. An archive containing all these files is available to the users when a run is finished (see more details in the following [section](#)).

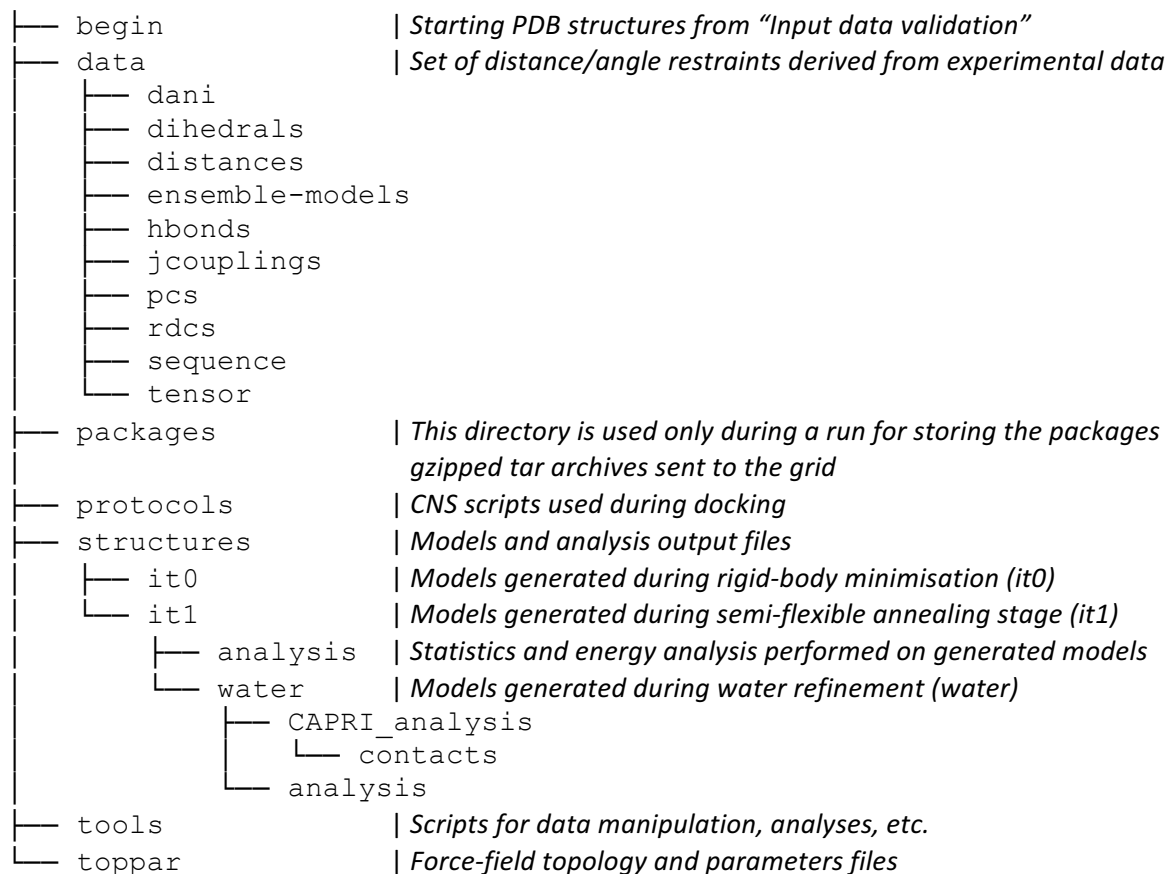
Even if the general workflow is executed mainly on the master node, the most expensive steps in terms of computing time are done remotely on either our local Linux cluster or using the EGI computing resources available worldwide.

The colour coding in the diagram (see legend) indicates which steps are performed exclusively by the web server (pink boxes) and which are common to the web server and a local, stand-alone version of HADDOCK (grey boxes). This clearly highlight the added value for end users of using the web portal rather than a local installation of HADDOCK in the current implementation.

Input and output files (with a short description of their format – all text files) are indicated in yellow in the diagram.

HADDOCK project directory structure

The project directory of HADDOCK contains all output files generated at any stage of the HADDOCK process. It also contains the whole set of input files including PDB files validated and prepared by HADDOCK for the docking and made from the users input structures.



Additional information about HADDOCK

For additional information about the HADDOCK software refer to its online pages:

- General information: <http://www.bonvinlab.org/software/haddock2.2>
- Online manual: <http://www.bonvinlab.org/software/haddock2.2/manual/>

