

HYDRA:

A WEB-BASED VISUALIZER FOR HIGH-
THROUGHPUT LIGAND DOCKING ANALYSIS

Nara Institute of Science and Technology
(NAIST)
Nara, Japan

Curtis Sera
7 August 2015

Project Overview

- This project aims to create a browser-based program that can simultaneously display many molecular interactions in a dynamically sized grid of molecular viewers.
 - Simulated interactions will be obtained from high-throughput simulation programs.
 - Yuan Zhao, a former student of Dr Haga's, previously created the framework for this program in Webix, a JavaScript library, and HTML5/CSS.
- This will enable almost any device with internet access to be used for data analysis with no end user setup.
- My specific focus will be on creating a functional graphical user interface (GUI)

Week 7 Progress

- 6 August - Met with Dr Haga & Ichikawa-sensei
 - Discussed progress and future direction
 - Final feedback for Shelby's poster and paper drafts
- Bug fix for 3Dmol viewers
 - Removed temporary patch from last week
 - Pulled and processed (with Grunt) updated code from main 3Dmol GitHub repository
 - Surfaces now behave as expected
- Webix GUI bug fixes

Week 7 Progress (cont)

- Attempted to implement Ichikawa-sensei's suggestion of synching model movements across viewers
 - Feature turned on via a checkbox
 - Code used caused indefinite looping and eventual crash
- Created a processor for raw docking output
 - Created a popup GUI opened via a button
 - The initial screen allows file uploading and provides settings for identifying relevant data
 - Upon clicking "Next" a second screen is shown for user confirmation of pre-processing results
 - GUI can currently take many files and filter them by file name
 - Relevant protein and ligand model files identified separately and shown in separate tables (see following dev. snapshots)

Week 7 Progress (cont)

The screenshot displays a software interface with a central docking simulation window. A modal dialog titled "Process files for upload" is open, showing a list of files and configuration options. The background interface includes a sidebar with "Import Compounds", "Process Docking Output", and "Upload Files" buttons, a table with columns "Col", "Row", and "File Name", and a "Grid Size" section with input fields for "Columns" and "Rows".

Process files for upload [Cancel]

Upload Raw Files

Please input the indicated information below. Protein files will be concatenated with ligand files for display together.

Use the following field encapsulated in square brackets as needed:

- **[*]** - Wild card field. Will accept any text (eg a compound unique name).
- **[ligID]** - The ID of the ligand. This is expected to be a dynamic field linking each unique protein and ligand docked pair.

Separate each field and other name components with a period `::`. Please note that your file names must be formatted to have the parts separated by `::` otherwise this parser will not work.

Protein File Nomenclature
[*],[*],[ligID].final_pose.amber.pdb

Ligand File Nomenclature
[*],[ligID].final_pose.amber.pdb

Ligand ZINC ID location

Remove [Next]

Compound List

Compound Name	Category
---------------	----------

Compound Details

Category

Compound

Residues

Bonds

Vendor List

Zinc ID : Vendor

Dev. snapshot 1: New file processing GUI with test files loaded and inputted filters.

7 August 2015

Week 7 Progress (cont)

The screenshot displays a software interface with a central dialog box titled "Process files for upload". The dialog contains two lists of files:

- Protein File Name:**
 - 1vhr_noH.3_top.1.final_pose.amber.pdb
 - 1vhr_noH.3_top.2.final_pose.amber.pdb
 - 1vhr_noH.3_top.3.final_pose.amber.pdb
 - 1vhr_noH.3_top.4.final_pose.amber.pdb
 - 1vhr_noH.3_top.5.final_pose.amber.pdb
 - 1vhr_noH.3_top.6.final_pose.amber.pdb
- Ligand File Name:**
 - 3_top.1.final_pose.amber.pdb
 - 3_top.2.final_pose.amber.pdb
 - 3_top.3.final_pose.amber.pdb
 - 3_top.4.final_pose.amber.pdb
 - 3_top.5.final_pose.amber.pdb
 - 3_top.6.final_pose.amber.pdb

Text in the dialog: "The filtered protein files to be used for processing are displayed on the top left; the filtered ligand files to be used for processing are displayed on the bottom left. Data from the identified ligand files will be added to the end of the protein files. Any files not listed here will be discarded after processing. Please confirm that all desired protein files are present. If some files are not present, please check your inputted filters for typos and confirm that your file names are in a compatible format."

The background interface includes a sidebar on the left with buttons for "Import Compounds", "Process Docking Output", "Upload Files", and "Grid Size" (with "Update Grid" button). A table with columns "Col", "Row", and "File Name" is visible. On the right, there are panels for "Compound List", "Compound Details" (with fields for Category, Compound, # Residues, # Bonds), and "Vendor List" (with "Zinc ID : Vendor" field). A "Recenter" button is located at the bottom right of the main window.

Dev. snapshot 2: Successfully filtered files placed and displayed for user confirmation.

7 August 2015

Project Highlights

- All essential functions for Hydra are now complete
 - Models can now be uploaded into a viewer in a grid of an arbitrary number of viewers
 - Proteins and ligands can have their display settings changed independently
 - Reduced computational load
 - Asynchronous file handling and 3Dmol's web workers (JS multi-threading)
 - Basic compound information is shown when the file is loaded

Plan for Final Weeks

- File processing
 - Concatenate filtered files from raw docking output
 - Process ligand files: replace “ATOM” tags with “HETATM” tags
 - Place ligand ZINC ID (if applicable) in final script output
- Improvements to Shelby’s work
 - Display more compound information upon double click of ZINC ID
 - Clarify in the GUI which compound is loaded into which viewer
- Note: will be staying with relatives for Obon

Exploration



Clockwise from below:

Outside Kyoto's Torokko Saga Station Museum, delicious ¥200 (\$2) shaved ice at Ikoma's Dondoko Matsuri (festival), exploring Kyoto Station, Ikoma Dondoko Matsuri lantern, my friend Hikari and Michelle at Dondoko Matsuri



Exploration



Clockwise from above:

Sunset from Kyoto's Togetsu Bridge; feeding a macaque at Arashiyama Monkey Park Iwatayama; me on Togetsu Bridge; pond by Mikami Shrine, Kyoto; in the Bamboo Groves of Arashiyama, Kyoto



Acknowledgements

- I would like to thank all those involved in making this possible:
 - Mentors
 - Dr Jason Haga & (Kohei) Ichikawa-sensei
 - UCSD Pacific Rim Experiences for Undergraduates (PRIME)
 - Dr Gabriele Wienhausen, Teri Simas, Madhvi Acharya, & Jim Galvin
 - NAIST's Software Design & Analysis Lab (SD Lab)
 - (Hajimu) Ōda-sensei, (Yasuhiro) Watashiba-sensei, & (Tomoko) Arai-sensei
 - Financial support
 - Julia Brown
 - UCSD Undergraduate Research Scholarship coordinator Dr Sophia Tsai
 - NAIST & Japan Student Services Organization (JASSO)
 - NAIST coordinator Nao Terada
 - National Science Foundation
 - Additional thanks to all the SD Lab graduate students, the PRIME alumni, Yuan Zhao, and the developers of Webix & GLmol
 - Special thanks to 3Dmol developer Dr David Koes for his great helpfulness and communicativeness