Hydra: A High-Throughput Virtual Screening Data Visualization and Analysis Tool

Shelby Matlock
UNCC
9201 University City Blvd.
Charlotte, NC 28223, USA
+1 704-687-8622
smatlock@uncc.edu

Curtis Sera
UCSD
9500 Gilman Dr.,
La Jolla, CA 92093, USA
+1 858-532-2230
csera@ucsd.edu

Kohei Ichikawa,
Yasuhiro Watashiba
NAIST 8916-5 Takayama
Ikoma, Nara 630-0192, JP
+81 074-372-5111
Ichikawa@is.naist.jp,
Watashiba@is.naist.jp

Jason H. Haga
AIST
1-1-1 Higashi Tsukuba,
Ibaraki 305-8561, Japan
+81 029-861-2395
jh.haga@aist.go.jp

ABSTRACT
One fundamental problem in computational drug discovery is that optimal docking scores do not necessarily indicate the best drug target. This results in a need for drug developers to manually examine molecular interactions between protein-ligand complexes. The Hydra virtual screening visualization tool aims to aid drug developers in this endeavor. Hydra is a scalable, lightweight, JavaScript-based [1] virtual screening tool that combines synchronous visualization with 3DMol [2] and compound information from ZINC [3] to solve this problem.

Categories
H.5.0 [INFORMATION INTERFACES AND PRESENTATION (e.g., HCI)]: General

J.3 [LIFE AND MEDICAL SCIENCES]: Biology and Genetics

General Terms
Design, Performance, and Experimentation.

Keywords
Virtual screening, visualization, data analysis, grid computing, bioinformatics.

1. INTRODUCTION
The purpose of Hydra is to create a user-friendly, high-throughput protein ligand docking visualization tool that can be used to aid in drug development and academia alike. With the ever-increasing amount of interrelated bioscience data, it is necessary to develop tools that allow researchers to access multiple data sources from a single streamlined interface. Unlike other molecular visualization tools, Hydra allows for synchronous molecular viewing in multi-dimensions, secure stand alone and online versions, and most importantly the ability to access chemical metadata and purchase compounds within the interface. Additionally, Hydra is open source, thus assuaging the financial burdens that come with expensive drug development software.

2. TECHNICAL DETAILS
Hydra's interface allows the user to upload data that are generated from virtual screening experiments. These data are files containing protein-ligand conformation data that are sent to a secure server where the data are parsed and sent to the molecular viewer while vendor and chemical compound information is culled from Hydra's database. The resulting vendor information and visualizations can be accessed and saved by the user from any device with Internet access.

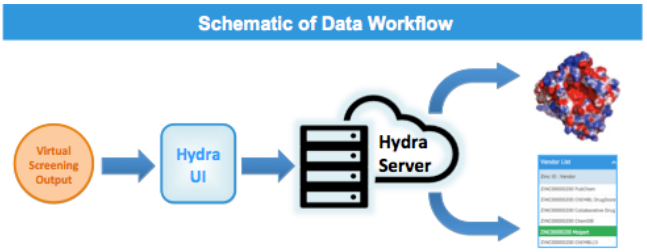

Figure 1. This schematic demonstrates how the user's data are parsed and then displayed within Hydra's interface.

2.1 Application Framework
To provide a quick and intuitive user experience, we used the Webix JavaScript/HTML5 framework [1]. This framework created the multiple viewer grid, easy-to-use buttons, and accordion menus in the interface. The user can alter the number of molecules viewed in Hydra's interface by manipulating the grid settings, thus allowing a more high-throughput analysis to be performed. This framework also allows for scalability across all visualization platforms from tile display walls to mobile devices.

2.2 Molecular Visualization
Hydra generates a 3-dimensional view of data uploaded by the user by leveraging on 3DMol [2], an object-oriented JavaScript library for viewing molecular data. 3DMol does not need browser plugins or Java, which adds to Hydra's security and allows for a more streamlined data workflow.

2.3 Chemical Compound Data
The chemical compound metadata is derived from ZINC [3], which is an open source chemical compound database developed by the Irwin and Shoichet Laboratories at UCSF. Incorporating ZINC into Hydra allows for the user to have access to over 35 million compounds commonly used in virtual screening.

2.4 User Interface
Hydra successfully combines visual, interactive, and chemical compound data in a multidimensional fashion as seen in Figure 2. The user can upload data and select visualization settings and manipulate grid settings on the left side of Hydra's interface. These controls adjust the number of 3DMol cells that are displayed within the multidimensional viewer space. The right side of Hydra's interface provides the user with access to a selected compound's chemical metadata, information about vendors who supply the compound, and the ability to purchase the compound. Incorporating this information within Hydra's

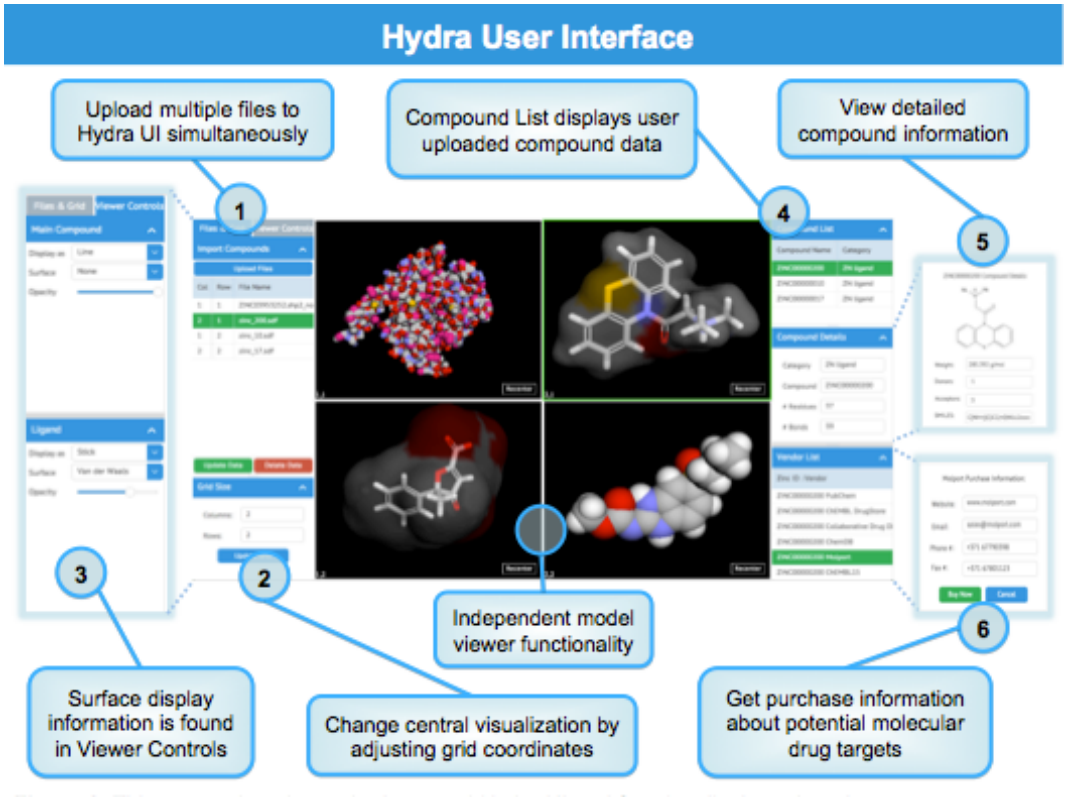

Figure 2. This screenshot shows the improved Hydra UI and functionality based on the current prototype. Data is uploaded to the Import Compounds tab (1) and the grid is adjusted in (2). Visualization settings are selected in the Viewer Controls tab (3). Compound data is viewed in Compound List (4) and additional chemical information in Compound Details (5). Vendor information is displayed by selecting a vendor from Vendor List (6)

interface provides the user with chemical data to do further research and the option of purchasing the drug target directly.

3. RESULTS

The working prototype of Hydra is the next step in virtual screening visualization tools [4, 5]. Hydra's ability to facilitate the purchase of compounds directly within the interface adds functionality that all other visualization tools lack. Hydra's user-friendly interface and flexibility provides customizable functionality for each user. Additionally, the cutting edge, multidimensional viewing space provided by Hydra simplifies manual analysis of multiple molecular interactions.

4. DISCUSSION

Hydra's working prototype succeeds in aiding in the virtual screening visualization process by providing a streamlined interface with select metadata. In the future, we hope to implement SQL functionality to our chemical databases for quick, seamless access to large scale compound and purchase information. Hosting these databases on a public cloud would give Hydra an overall increase in speed and performance when the user uploads massive data files. Additionally, cloud hosted databases would allow for succinct 3D visualization response.

5. ACKNOWLEDGMENTS

The authors would like to thank the Partnership in International Research (PIRE) program, the Open Science Data Cloud (OSDC), the National Institute for Advanced Industrial Science and Technology (AIST), and the NSF for their support.

6. REFERENCES

[1] Webix, http://www.webix.com

[2] N Rego and D Koes, 3Dmol.js: Molecular Visualization with WebGL. Bioinformatics. 2014.

[3] JJ Irwin, T Sterling, MM Mysinger, ES Bolstad and RG Coleman. ZINC: a free tool to discover chemistry for biology. J. Chem. Inf. Model. 2012.

[4] CD Lau, M Levesque, S Chien, S Date, and JH Haga. ViewDock TDW: high-throughput visualization of virtual screening results. Bioinformatics. 2010

[5] Y Zhao and JH Haga, Hydra: an HTML5-based application for high-throughput visualization of ligand docking. Poster session presented at: Supercomputing 2014; Nov. 16-21, 2014; New Orleans, LA.