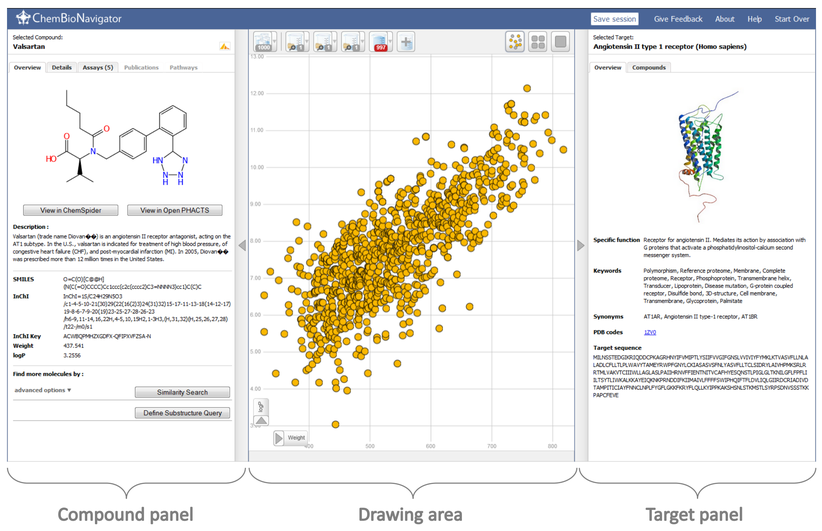
Accessing Open PHACTS: Interactive exploration of compounds and targets from the semantic web

Katrin Stierand | Tim Harder | Lothar Wissler | Christian Lemmen | Matthias Rarey

Heterogeneous data sources represent a major challenge in pharmacological research. Gathering all publicly available information for a molecule requires numerous different database requests. Hence, it increases research costs and limits throughput. Over the last two years, the Open PHACTS Discovery Platform[1] has been developed as a centralized repository, integrating pharmacological data from a variety of information resources and providing tools and services to query these integrated data. One of these services is the ChemBioNavigator[2] (CBN), a web application allowing to navigate the chem-bio space with a focus on small molecules and their targets. The CBN is available at www.chembionavigator.org

The ChemBioNavigator

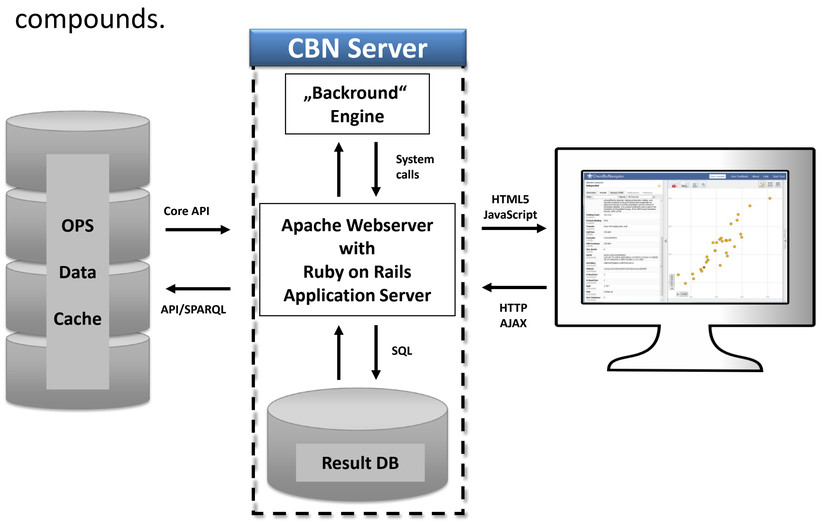


The CBN comprises of three areas:

1. The **compound panel** providing general information and physico-chemical details of a compound as well as available assay data. Similarity and substructure searches for the selected compound can be initiated from here.
2. The **drawing area** with different visualization modes: A cloud view, where the compounds are drawn as dots sorted by physicochemical properties, a table view showing the molecules structure diagrams and a cover flow view showing single structure diagrams.
3. The **target panel** gives information about a selected target. A search for compounds, which are active against the selected target.

Architecture

The architecture of the CBN is designed around a Ruby on Rails web-application server, connecting the different parts of the application. Next to the task of serving the webpages, the Rails server also handles the remote OPS queries, connects to the local database as well as manages the so-called background engine, responsible for the initialization and validation of compounds.

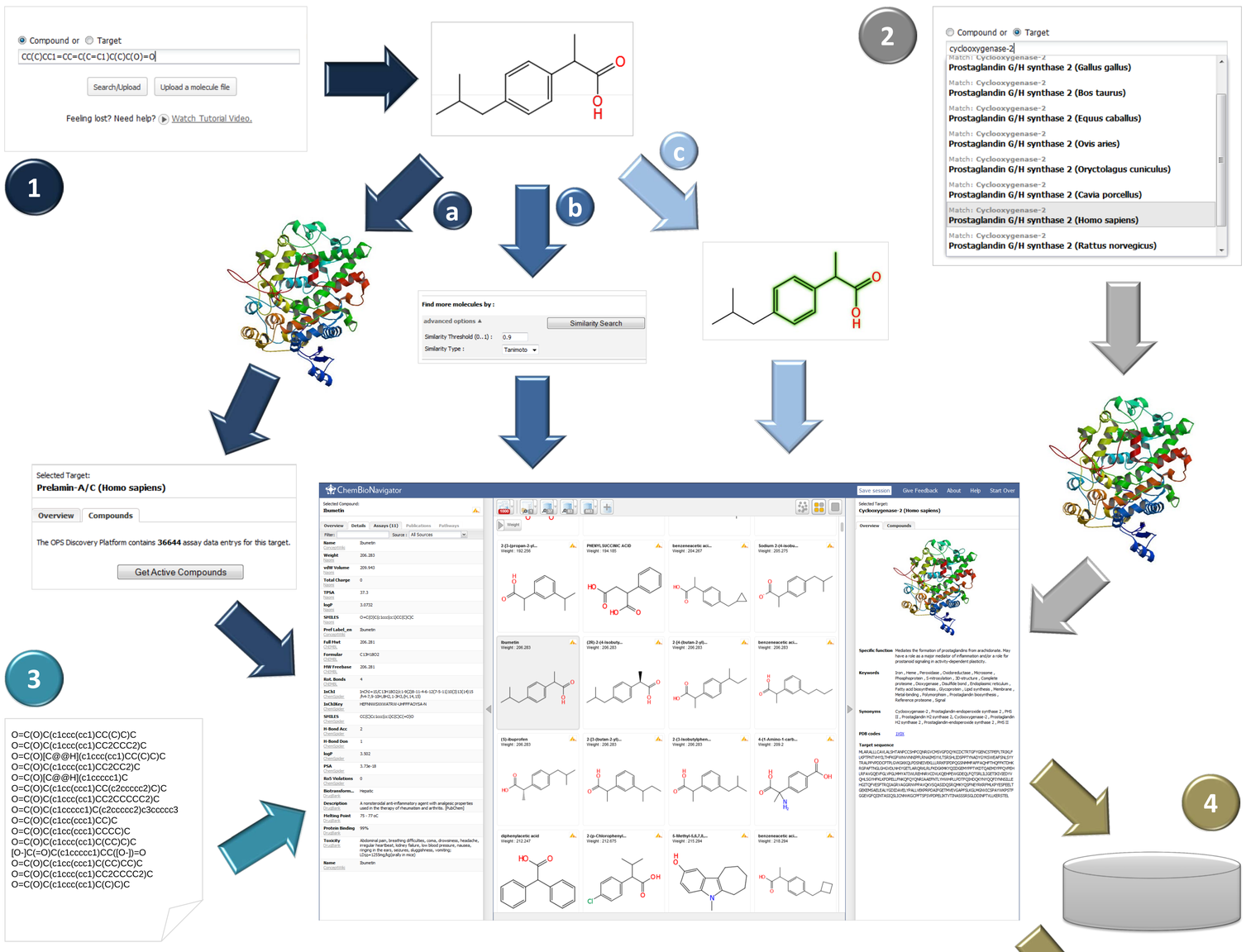


Data Sources

The OPS Data Cache (Triple Store) contains triples generated from various data sources. It is permanently updated and extended by new (public and commercial) data.

ChemSpider · UniProt · the Gene Ontology · DrugBank · swissprot · ChEBI · ChEMBL · ConceptWiki · WikiPathways

Workflows



Following an application-oriented approach, the Open PHACTS project started with the definition of prioritized research questions[3], most of which can only be answered by accessing multiple data sources in the web. The development of the platform as well as the services has been guided by these questions. In order to cover as many high prioritized questions (concerning small ligands and their targets) as possible, the CBN incorporates different workflows.

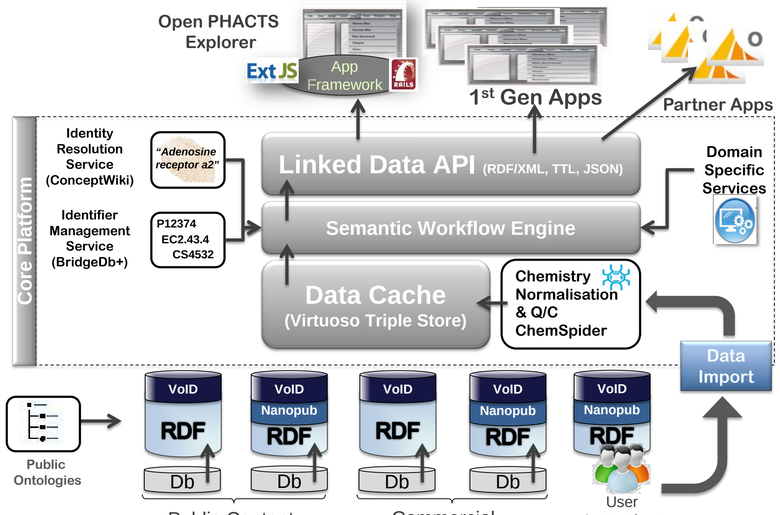
1 Start with a single compound name or a SMILES and extension of the molecule set by related compounds:
- a Find target via assay data and load all compounds from OPS, which are active against this target
- b Start a similarity search
- c Define a substructure and start a substructure search

2 Start with target name:
The CBN loads the target information from OPS. In the mean time all active compounds are loaded in the CBN.

3 Start with a set of SMILES or a SD file:
The given compounds are initialized by NAOMI and annotated with data from the OPS.

4 Storing the results:
Compound sets can be either exported in different file formats or the session can be stored .

Open PHACTS discovery platform



The platform consists of 6 classes of components:

1. Data Sources - these provide the underlying data
2. Linked Data Cache - this manages the integration and querying across multiple data sources
3. Domain Specific Functions/Services - these provide domain specific functions for comparison or analysis that can be called by the LinkedData cache
4. Identity Resolution Service - this maps a particular identifiers (either uris or terms) to other identifiers (uris or terms)
5. Mapping Providers/Curation - repositories of mappings and interfaces to manage and create those mappings
6. User Interfaces

Informations about the platform are available at www.openphacts.org

[1] Williams A., Harland L., Groth P., et al. Drug Discovery Today, June 06, 2012
[2] Stierand K., Harder T., Marek T., et al. Molecular Informatics, Volume 31, Issue 8, p. 543–546, August 2012
[3] Azzaoui K, Jacoby E, Senger S, et al. Drug Discovery Today, Volume 18, Issues 17–18, Pages 843-852 , September 2013

Katrin Stierand – stierand@zbh.uni-hamburg.de
Matthias Rarey – rarey@zbh.uni-hamburg.de
ZBH Center for Bioinformatics, Bundesstr. 43, 20146 Hamburg, Germany
www.zbh.uni-hamburg.de

Lothar Wissler – lothar.wissler@biosolveit.de
Christian Lemmen – clemmen@biosolveit.de
BioSolveIT GmbH, An der Ziegelei 79, 53757 Sankt Augustin, Germany
www.biosolveit.de

The development of the ChemBioNavigator in the context of the OpenPHACTS project is funded by the Innovative Medicines Initiative Joint Undertaking under grant agreement n° [115191], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.