

## Accessing OpenPHACTS:

Interactive exploration of compounds and targets from the semantic web



**Katrin Stierand**

ZBH Center for Bioinformatics  
Hamburg



## Outline

### ✦ OpenPHACTS

- What is OpenPHACTS and why do we need it?
- Use case driven development
- Technical structure, eApps

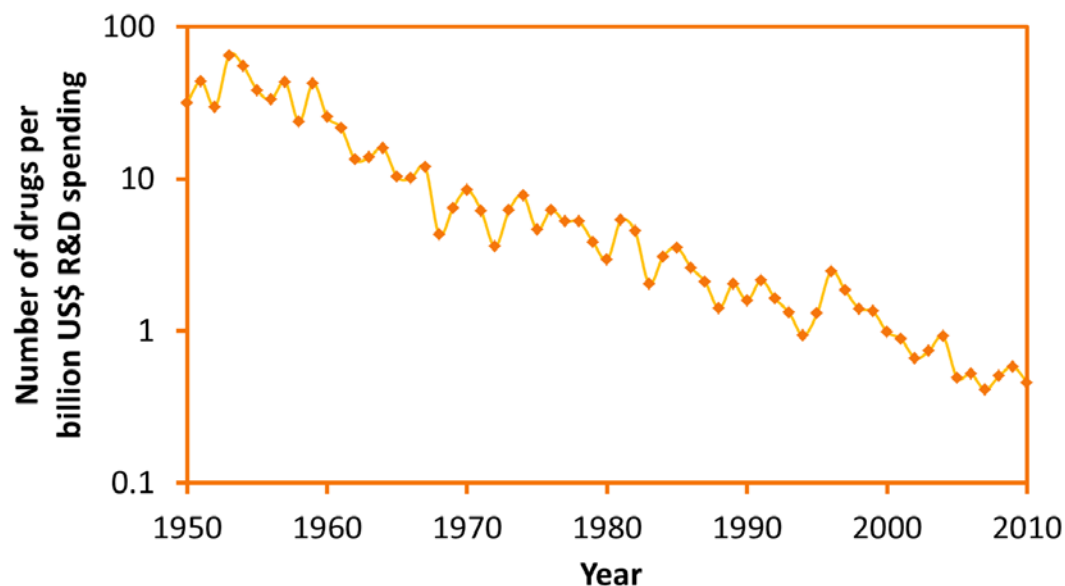
### ✦ ChemBioNavigator

- GUI and technical structure
- Workflows



## Why do we need Open PHACTS?

- ✦ Pharmaceutical companies currently expend significant effort integrating the vast amount of data publicly available into internal architectures.



Currently, pharmaceutical companies assemble their own in-house databases of pharmacological and physicochemical data.

**Drug discovery process is hindered by repetition of:**

- Data extraction
- Transformation
- Loading stage

**Overall trend in R&D efficiency, inflation-adjusted** (J. W. Scannel, A. Blanckley, H. Boldon and B. Warrington, *Nat. Rev. Drug Discov.*, 2012, **11**, 191-200, (doi:10.1038/nrd3681))



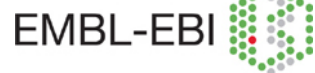
## What is OpenPHACTS?

- ✦ OpenPHACTS Discovery Platform
  - an online platform with a set of integrated publicly available pharmacological data
- ✦ Open Pharmacological Space
- ✦ Intended to facilitate improvements in drug discovery in academia and industry
- ✦ Use and enhance semantic web standards



## Open PHACTS Project Partners

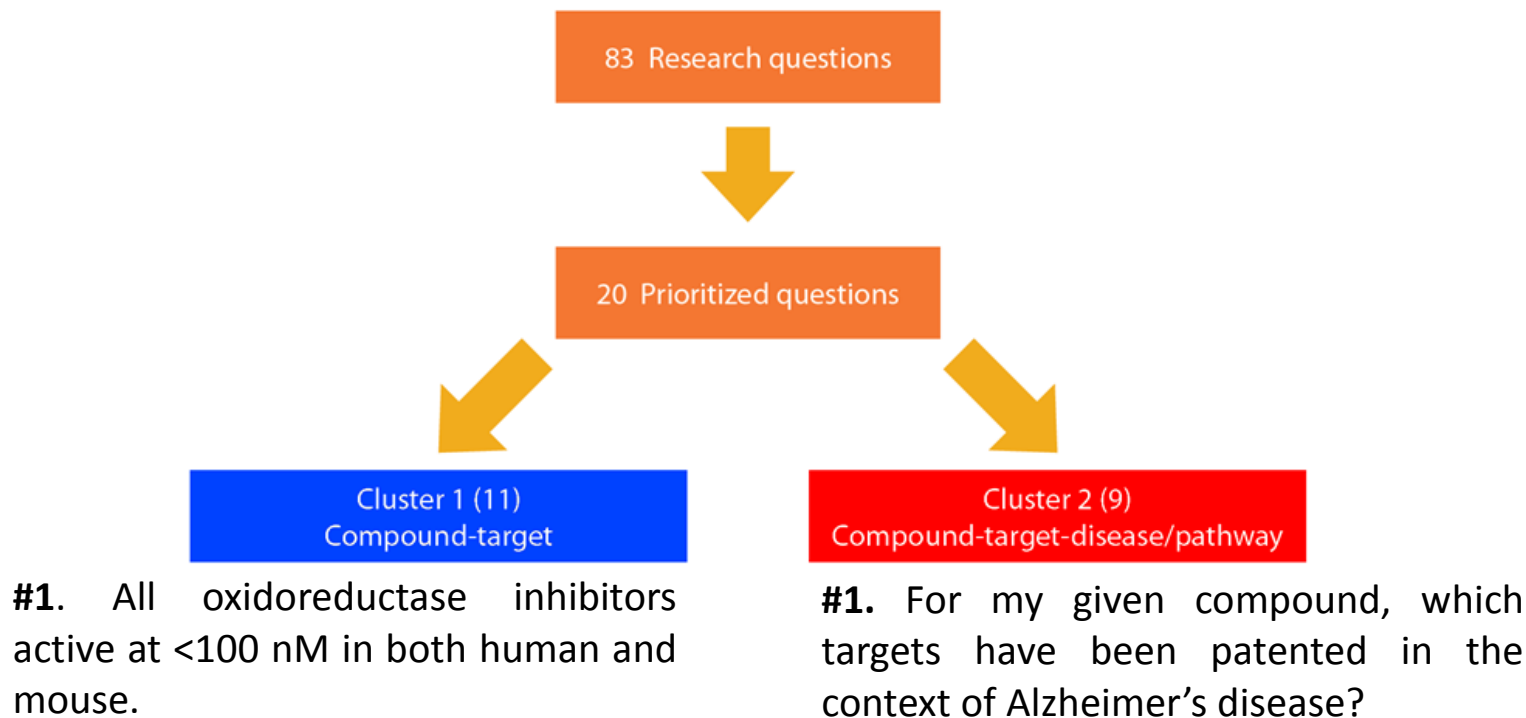
[www.openphacts.org](http://www.openphacts.org)



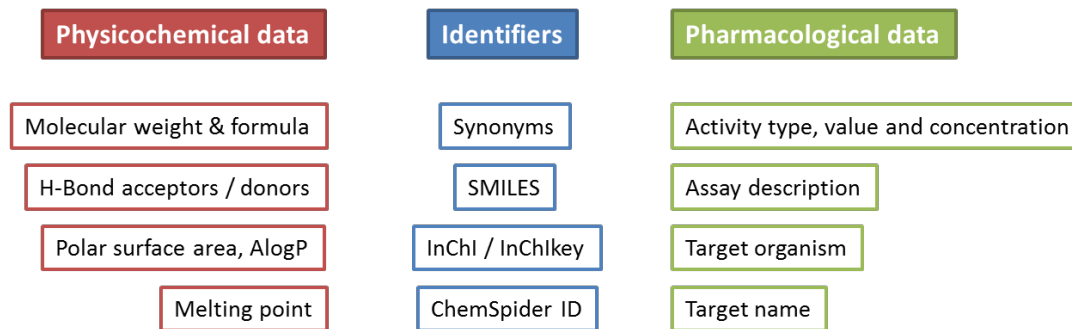
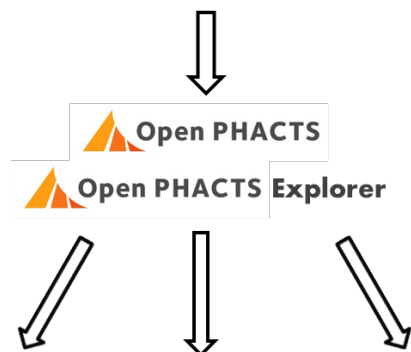


## Pharmacology within Open PHACTS

The number of pharmacological questions that could be useful to answer is large, and Open PHACTS concentrates on answering the **top 20 ranked research questions** from a list of 83 proposed by consortium members.

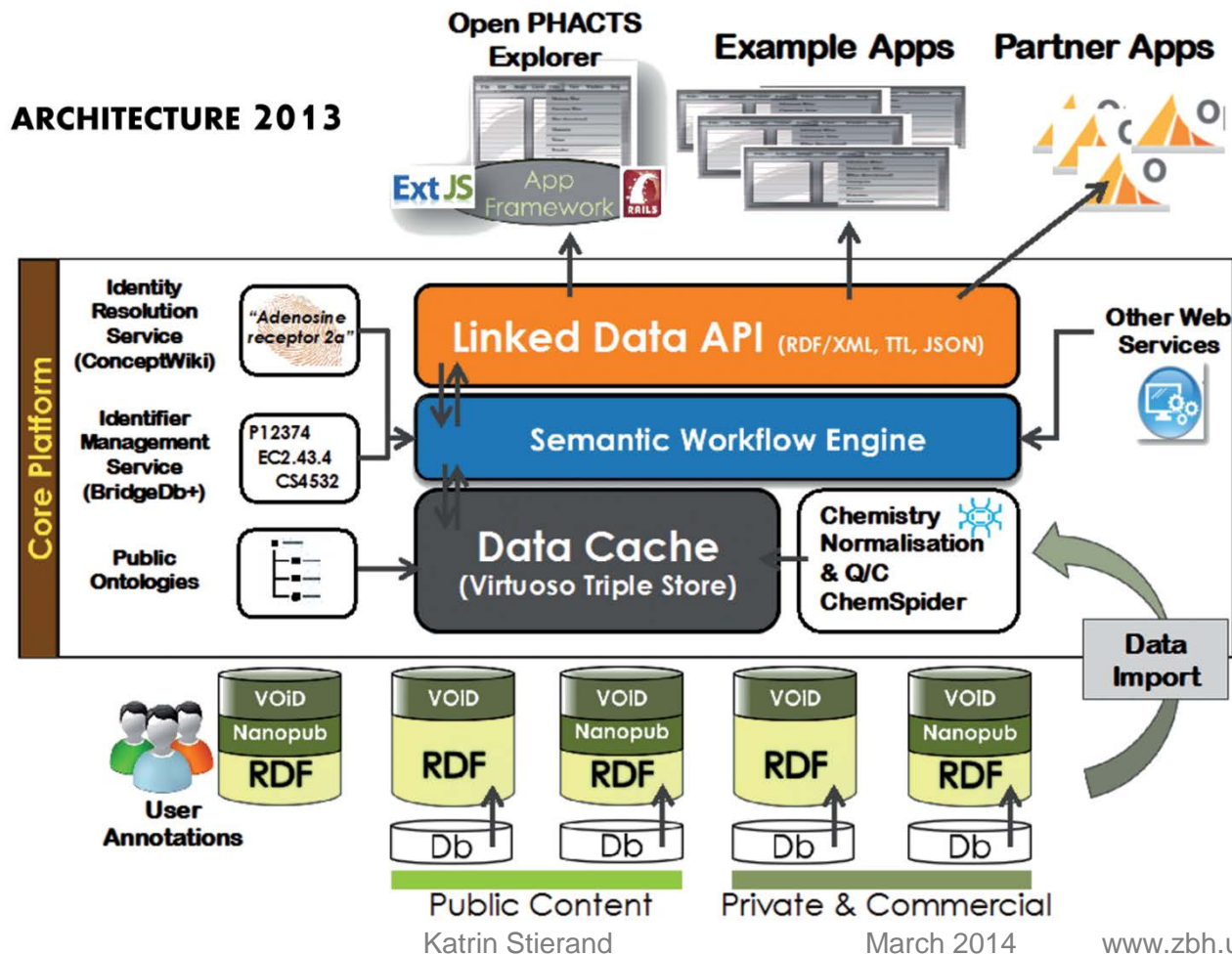








## How does Open PHACTS work?







## The ChemBioNavigator (CBN)

[www.chembionavigator.org](http://www.chembionavigator.org)

Google style: Interactive – Simplicity rules



Stierand K., Harder T., Marek T., et al. Molecular Informatics, Volume 31, Issue 8, p. 543–546, August 2012



## ChemBioNavigator GUI

**ChemBioNavigator**

Selected Compound: **Valsartan**

Overview Details Assays (5) Publications Pathways

**Description:**  
Valsartan (trade name Diovan) is an angiotensin II receptor antagonist, acting on the AT1 subtype. In the U.S., valsartan is indicated for treatment of high blood pressure, of congestive heart failure (CHF), and post-myocardial infarction (MI). In 2005, Diovan was prescribed more than 12 million times in the United States.

**SMILES**  
O=C(O)[C@H](N(C(=O)CCCC1Cc1ccc(cc1)C2=NNN3C(C)C(C)C2)C1)C(C)C

**InChI**  
InChI=1S/C24H29N5O3/c1-4-5-10-21(30)/29(22)(16(23)24(31)32)15-17-11-13-18(14-12-17)19-8-6-7-9-20(19)23-25-27-28-26-23/f6-9,11-14,16,22H,4-5,10,19H2,1-3H3,(4,31,32)(4,25,26,27,28)/t22-m/s1

**InChI Key**  
AC1WBQPMHXGDFX-QFIPXVFZSA-N

**Weight**  
437.541

**logP**  
3.2556

**Find more molecules by:**  
advanced options ▼ Similarity Search Define Substructure Query

**Selected Target:**  
**Angiotensin II type 1 receptor (Homo sapiens)**

Overview Compounds

**Specific function**  
Receptor for angiotensin II. Mediates its action by association with G proteins that activate a phosphatidylinositol-calcium second messenger system.

**Keywords**  
Polymorphism, Reference proteome, Membrane, Complete proteome, Receptor, Phosphoprotein, Transmembrane helix, Transducer, Lipoprotein, Disease mutation, G-protein coupled receptor, Disulfide bond, 3D-structure, Cell membrane, Transmembrane, Glycoprotein, Palmitate

**Synonyms**  
AT1AR, Angiotensin II type-1 receptor, AT1R

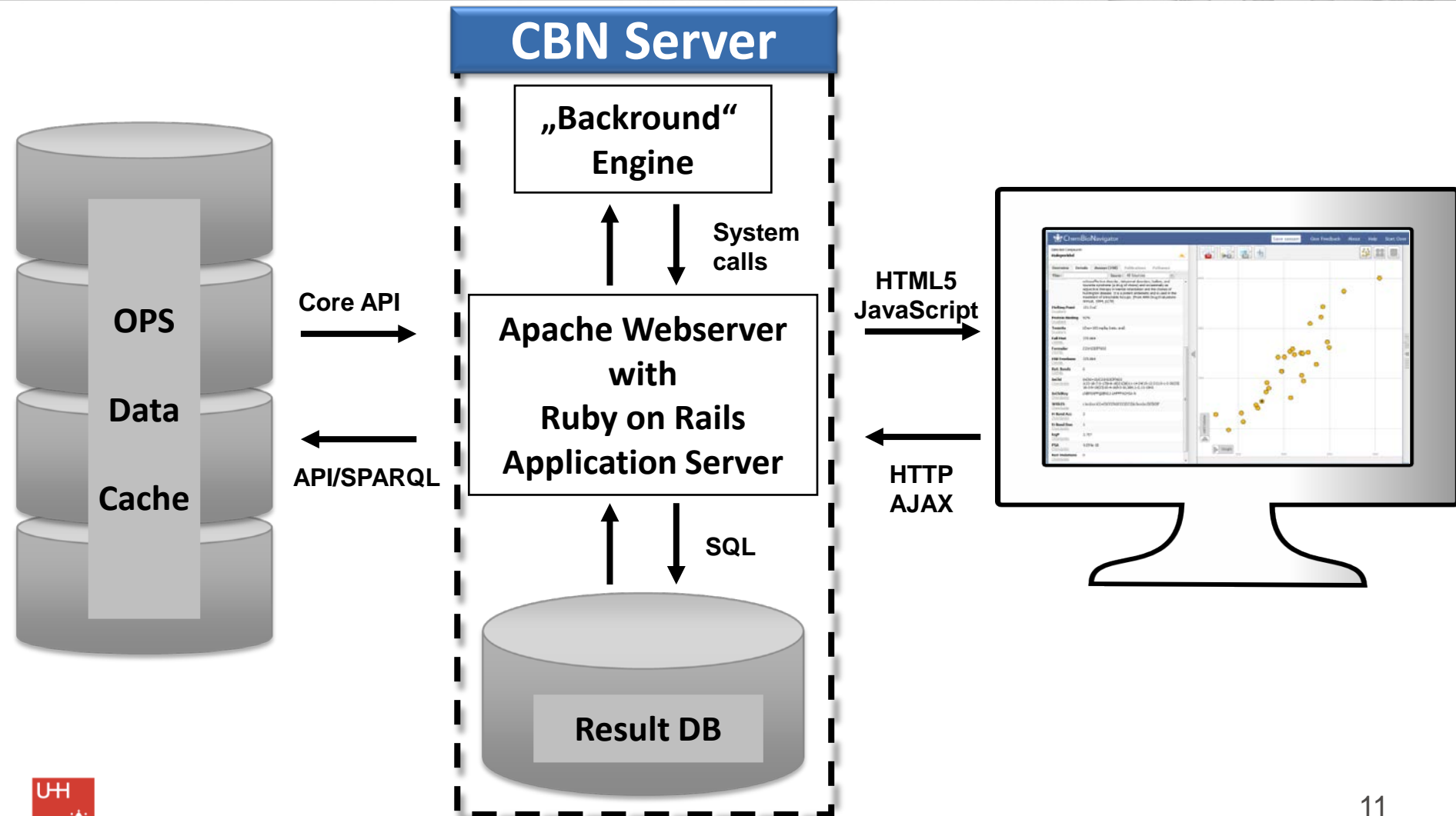
**PDB codes**  
[1ZVQ](#)

**Target sequence**  
MILNSTEDGKRWQDDCPKAGRHNYFVMPITLYSIIFVYGFGNSLVIVITFYMKLTVASVFLNLA  
LADLCPLTLPLWAYITAMETRWPPGNYLQKASASVFNLYASVFLTLQSLDRYLAIVHPKMSRLR  
RTMLVAKVYCTETLLAGLASPLADIRHVFIEINTNTVCAPHYEQNSTPLGLGTLKYNLGFPPFI  
ILTSYTLIKALKGAYEIQKQKPRNDIPKIMAINLFFFSWIPHQPTFLDWLIQGLDRCRIADIVD  
TAMPTTICIAVFNINCLNPLFYGLGKDFRYFLQLKLYIPPKASHNSLSTKMSLTSYRPSDNVSSSTKK  
PAPCFEVE

Compound panel

Drawing area

Target panel





## Features

- ✦ Provenance: all data is interlinked with the original source
- ✦ Drill-down: interactive sorting based on public and private properties
- ✦ Housekeeping: put compounds in „buckets“ to retain an overview
- ✦ Searching: based on similarity, substructure or related target
- ✦ Persistence: store a session and resume work later



## Development strategies

- ✦ Several interviews with people from pharmaceutical industry
- ✦ Knowledge of valuable features and deficiencies
- ✦ Extract trends from all interviews

=> agile and target-oriented development





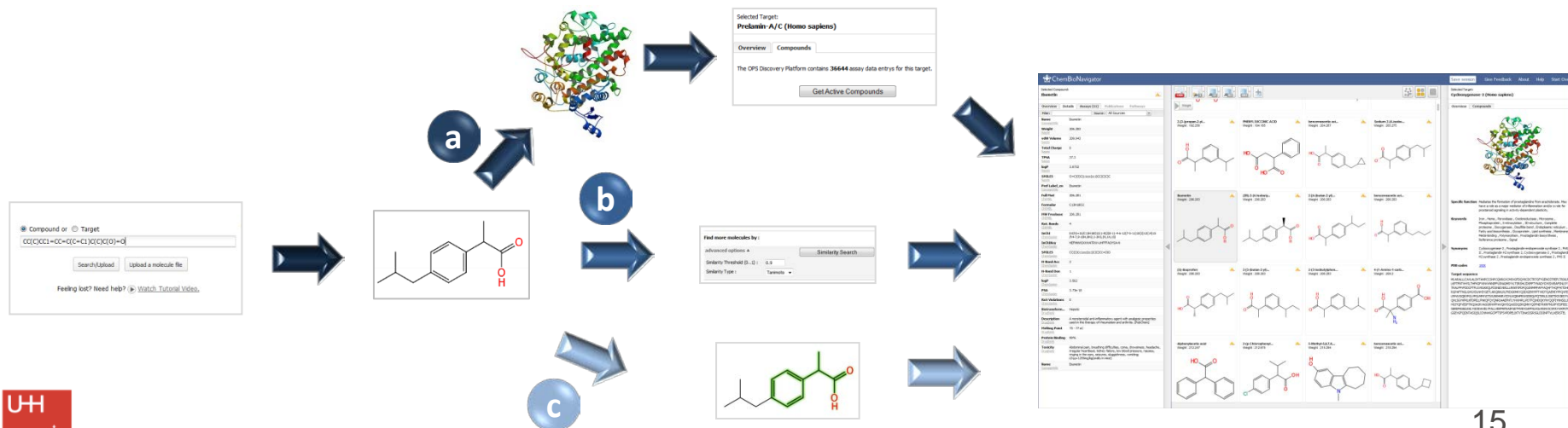
## Workflows

- ✦ Compound centric <-> target centric
- ✦ Extension and drill-down of data
- ✦ Three main workflows (following slides)
  - two compound centric
  - one target centric
- ✦ Workflows can be combined



## Workflow I

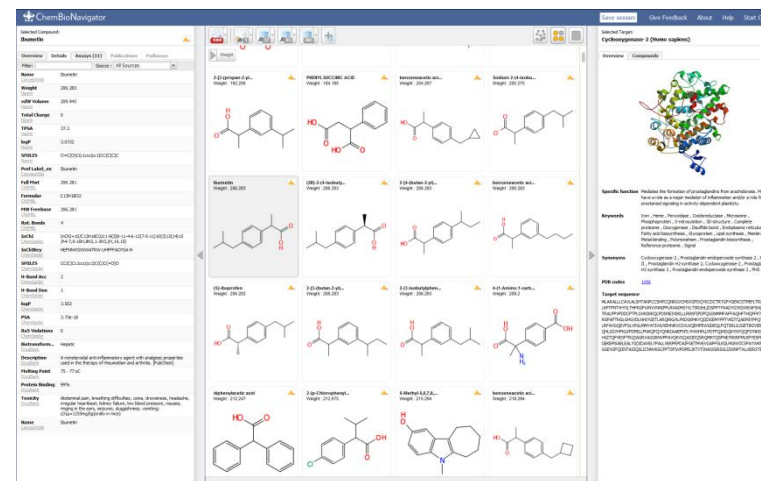
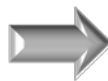
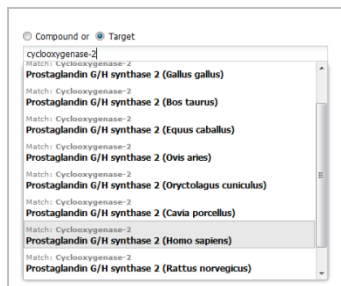
- ✦ Start with a single compound name or a SMILES and extend the molecule set by related compounds:
  - Find target via assay data and load all compounds from OPS, which are active against this target
  - Start a similarity search
  - Define a substructure and start a substructure search





## Workflow II

- ✦ Start with target name
- ✦ The CBN loads the target information from OPS.
- ✦ On mouse click, all active compounds are loaded in the CBN.





## Workflow III

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

```
O=C(O)C(c1ccc(cc1)CC(C)C)C
O=C(O)C(c1ccc(cc1)CC2CCC2)C
O=C(O)[C@@H](c1ccc(cc1)CC(C)C)C
O=C(O)C(c1ccc(cc1)CC2CC2)C
O=C(O)[C@H](c1ccccc1)C
O=C(O)C(c1cc(ccc1)CC(c2ccccc2)C)C
O=C(O)C(c1ccc(cc1)CC2CCCC2)C
O=C(O)C(c1ccc(cc1)C(c2ccccc2)c3ccccc3)C
O=C(O)C(c1cc(ccc1)CC)C
O=C(O)C(c1cc(ccc1)CCCC)C
O=C(O)C(c1ccc(cc1)C(CC)C)C
O=C(O)C(c1ccc(cc1)CC(C)C)C
O=C(O)C(c1ccc(cc1)CC2CCC2)C
O=C(O)C(c1ccc(cc1)C(C)C)C
```



The screenshot displays the ChemBioNavigator interface. On the left, a sidebar lists various chemical properties and annotations for the input compounds. The main area shows a grid of chemical structures, each with its name and molecular weight. On the right, a panel titled 'Specific functions' provides detailed information about the compounds, including their biological activities and potential targets.



## Conclusion

- ❖ OpenPHACTS integrates data from different data sources and provides it via an API
- ❖ Design and choice of sources use-case driven
- ❖ CBN exemplarily uses the OPS
- ❖ Enables navigation through the chem-bio space
- ❖ Development of workflows based on interviews with scientists of pharmaceutical industry





## Acknowledgement

- ✦ **Tim Harder** – Philips Medical Systems DMC GmbH
- ✦ **Lothar Wissler** – BioSolveIT GmbH
- ✦ Christian Lemmen – BioSolveIT GmbH
- ✦ Matthias Rarey – ZBH Center for Bioinformatics

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.

# Thank you for your attention!