



High Throughput Screening at EU-OPENSSCREEN

Edgar Specker

E-mail: edgar.specker@eu-openscreen.eu

Tel: +49 30 94793 510

European Research Infrastructure for Chemical Biology and Early Drug Discovery

EU-OPENSOURCE provides access to

- Technologies (screening platforms)
- Resources (compound collections)
- Expertise (assay development, medicinal chemistry)
- Data (compound structures, bioactivity)
- Training

Collaboration partners from academia and industry





eu:openscreen

Member Countries and Partner Sites

- EU-OPENSREEN is a distributed RI with 24 partner sites across Europe
- 3 partner site categories:
 - Screening platforms (17)
 - Chemistry groups (6)
 - Database host (1)
- Partner site accreditation is a 3-step procedure:
 - Nomination of site by ministry
 - Evaluation by external reviewers
 - Approval of individual sites by all ERIC member countries, based on evaluation reports



Screening & Chemistry Partner Sites

Denmark



Finland



Czech Republic



Latvia



Poland



Norway



Spain



Germany



in Braunschweig

Ursula Bilitewski



in Hamburg

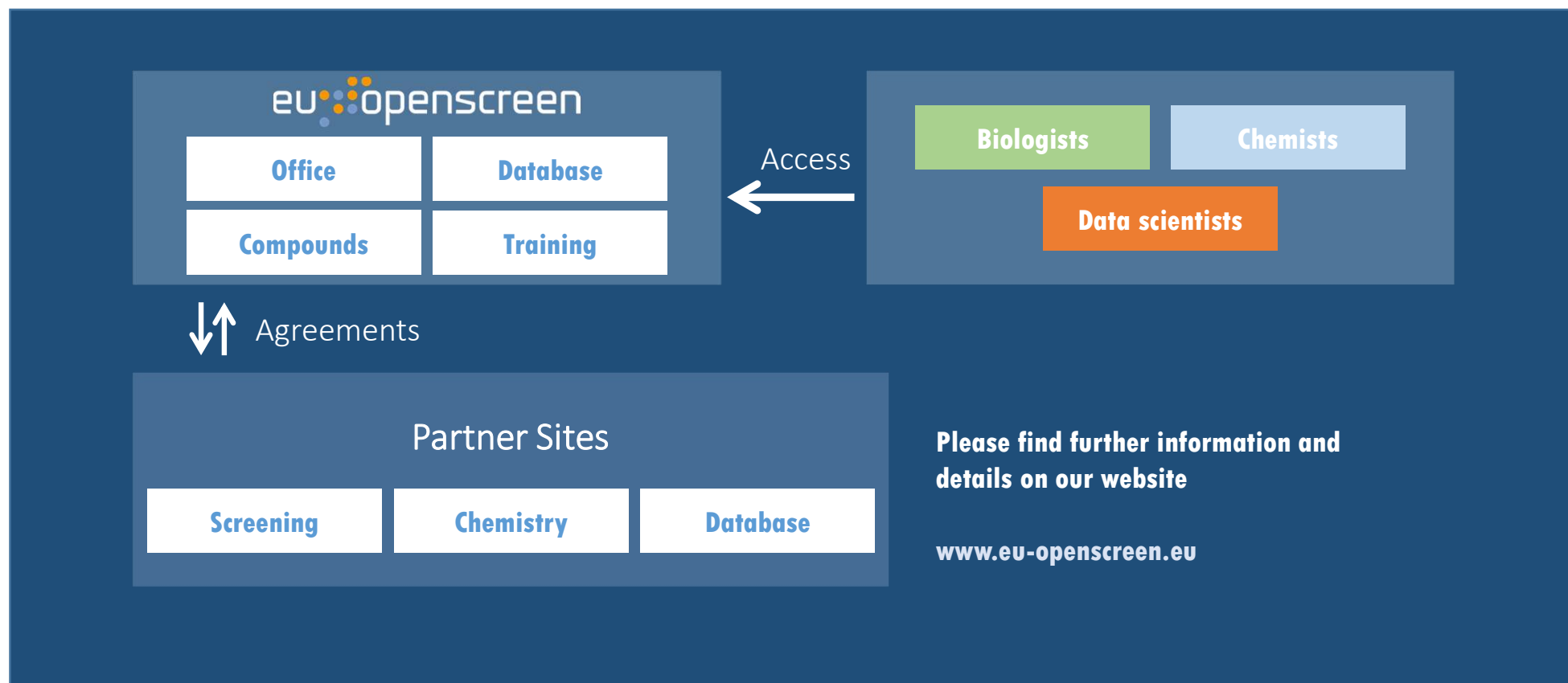
Phil Gribbon



in Berlin

Jens von Kries
Marc Nazaré

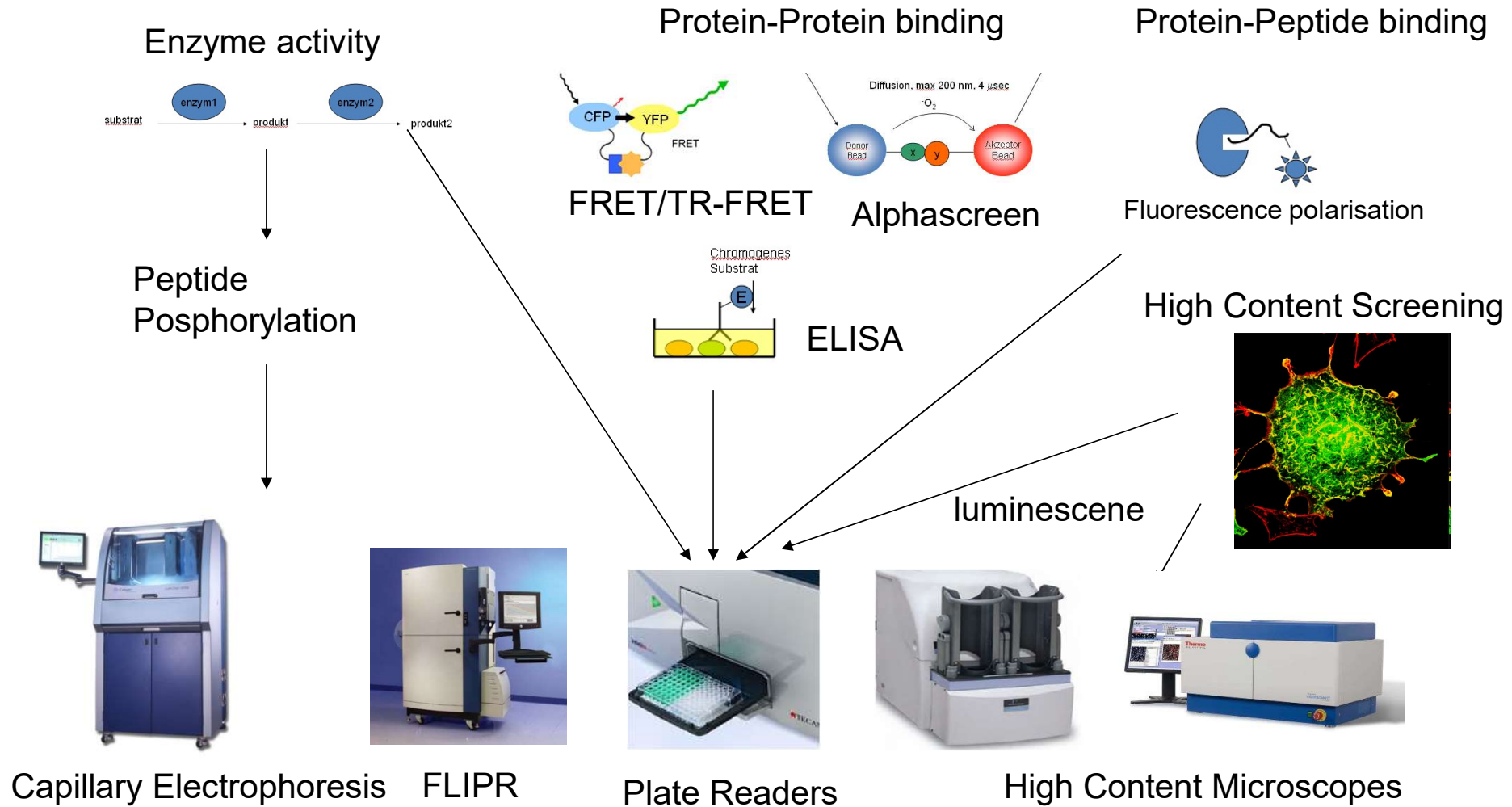
Access model



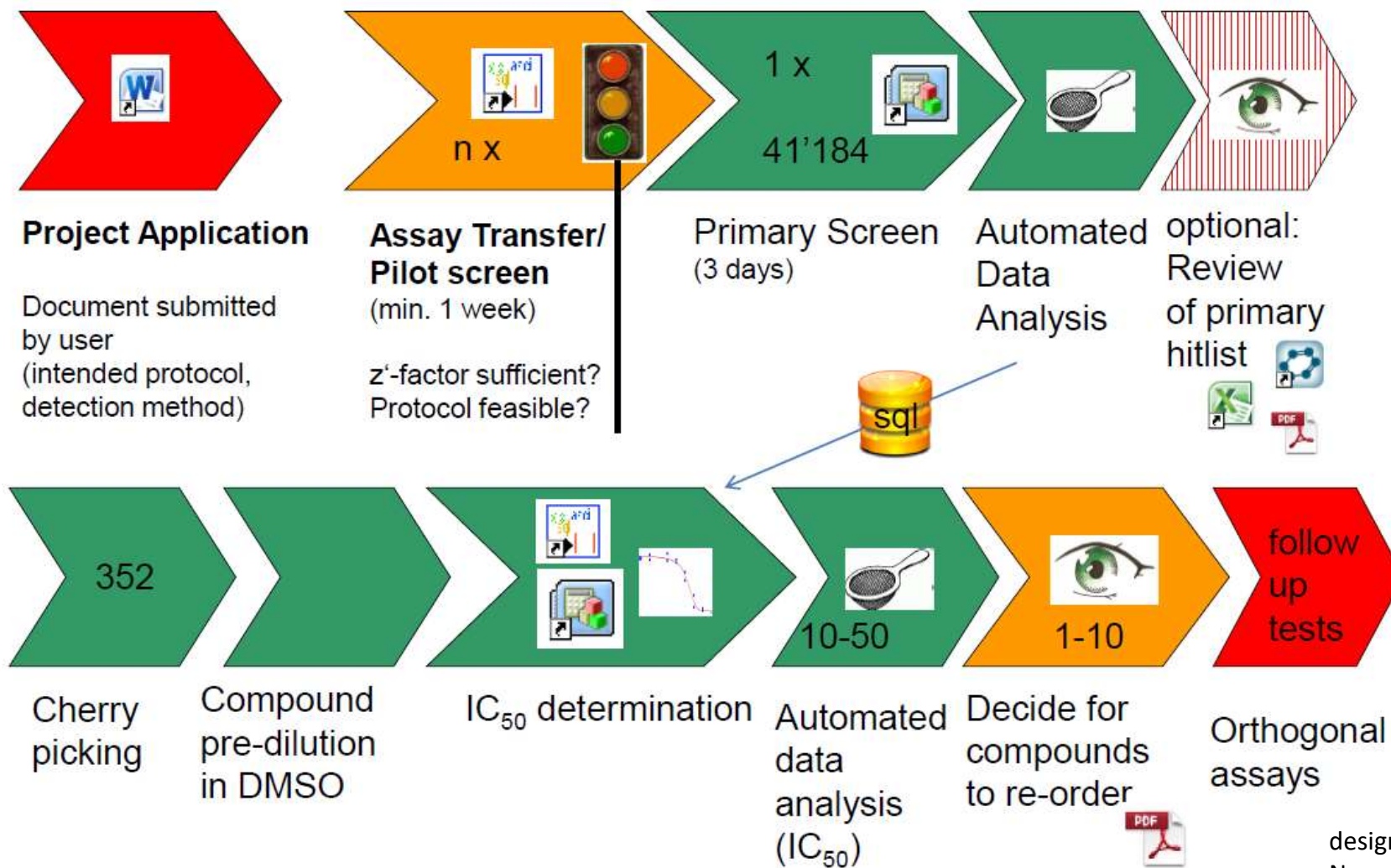
Two Libraries available for Screening at the FMP

Internal FMP Library	EU-OPENSOURCE Library
75.000 cmpds	100.000 cmpds
Acquired 2010 - 2020	Acquired in 2020-2021
Selection based on different concepts (highly diverse, spiro, sp3 rich, macrocycles, natural derived cmpds, bioactives, fragments)	Selection based on 5 different computational groups with 5 methods (highly diverse, 2464 bioactives, fragments)
25 % of structures visible	Structures visible in the ECBD
Structures of primary hits visible	Structures of primary hits visible
Comprehensive knowledge of properties	Property analysis initiated
User control of the data	User control of the data (obligation to publish after 3 years embargo in the ECBD)
8.000 academic cmdps available	Process initiated, target 40.000 cmpds
Costs and conditions individually negotiated	Cost and conditions in the EU-OPENSOURCE network (e.g. replenishment fee)
Library only accessible at the Screening Unit	Library accessible at 10 different Partner Sites
No structural overlap between the libraries	

Which Screening Technologies to Apply?

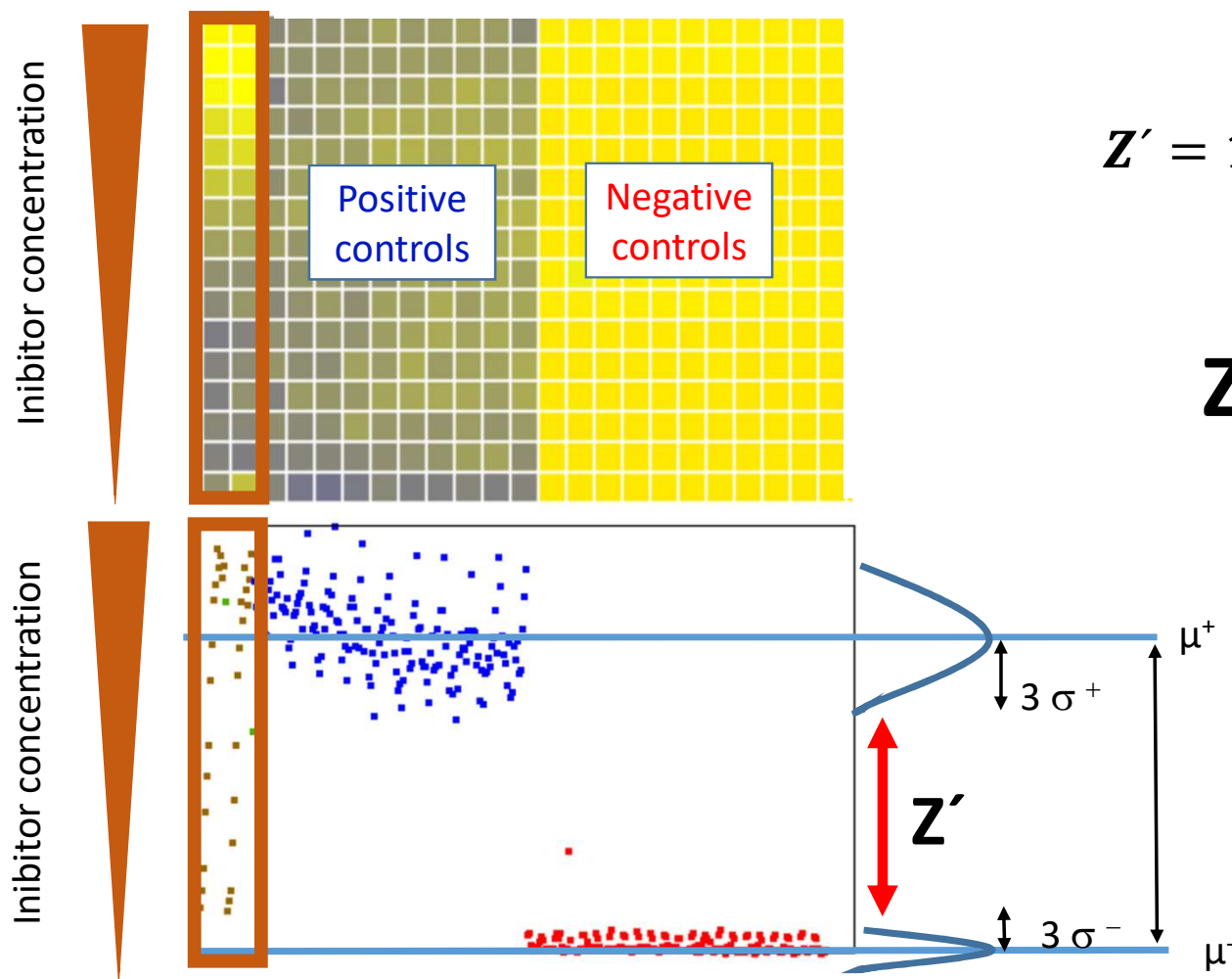


Our Workflow



designed by Martin Neuschwander
Neuschwander@fmp-berlin.de

Assay Quality, Z' Prime Factor

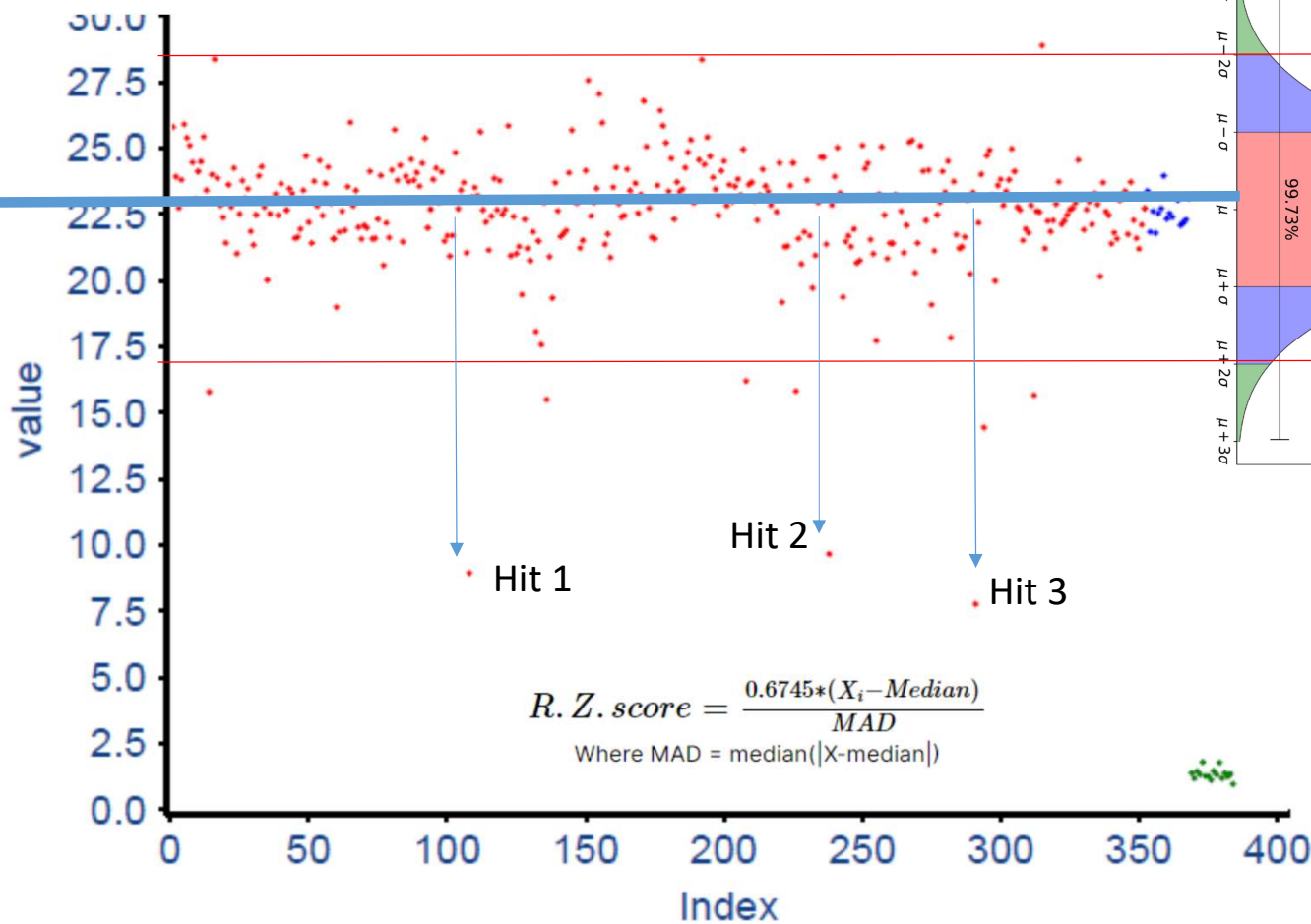
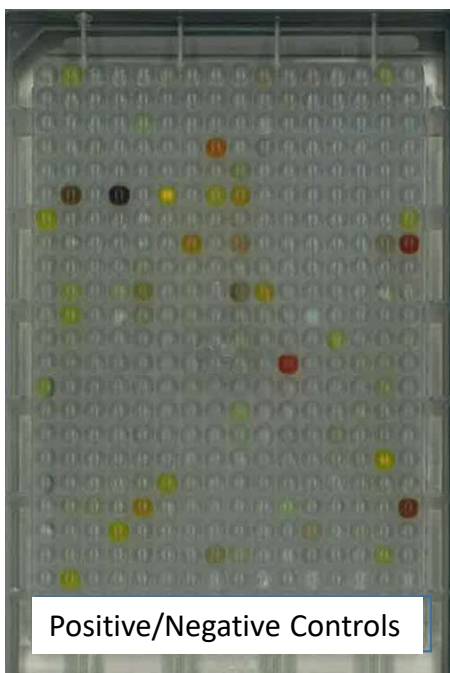


$$Z' = 1 - \frac{3(\sigma^- + \sigma^+)}{|\mu^- - \mu^+|}$$

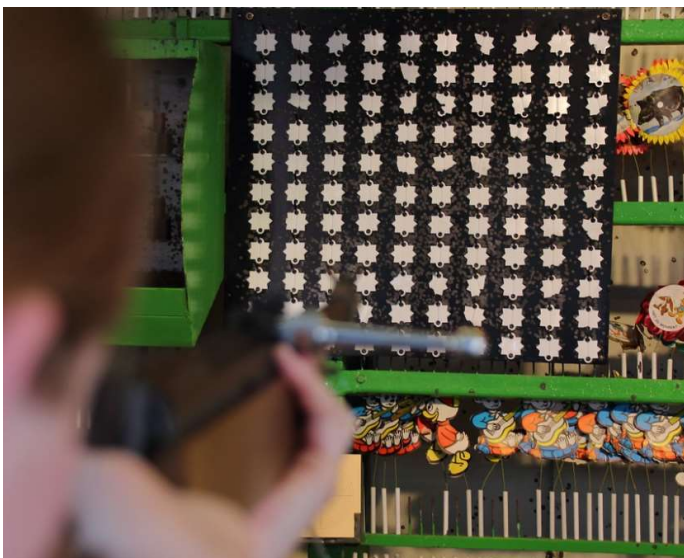
$Z' > 0.5$ 😊

Robust Statistics against Outliers: the Median

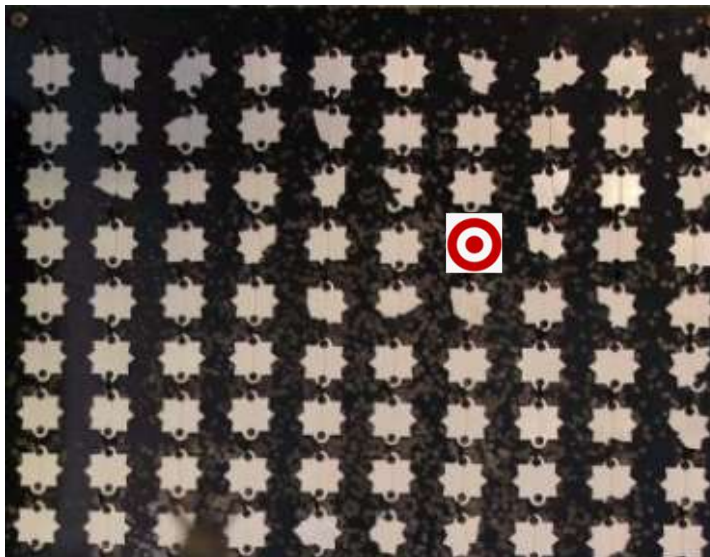
Median instead of Mean



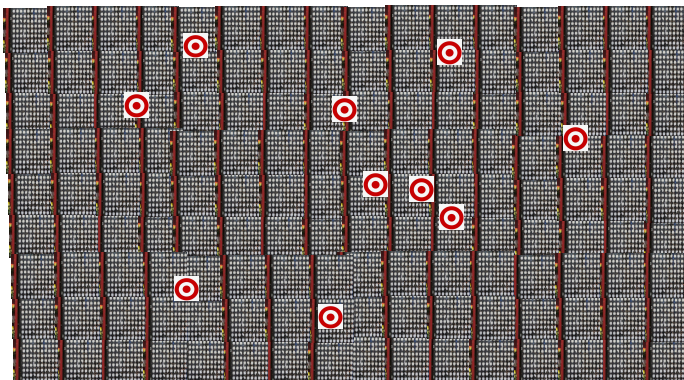
Empirical Hit Rate versus Assay Specificity



Source: Pixabay



Shooting specificity 99 %

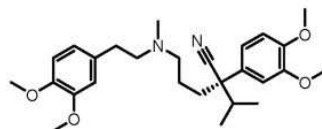


100.000 Cmpds library with an empirical hit rate of true positives 0.01 % => $100.000 \times 0.00001 = 10$ Hits

Assay/Shooting Specificity of 1%
 $100.000 \times 0.01 = 1.000$ Cmpds
=> **990** shots misleading

The large majority of hits are stochastic false positives

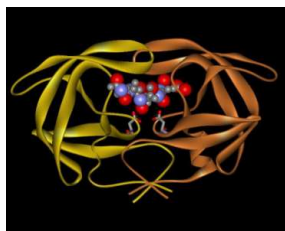
How to find Property based False Positives?



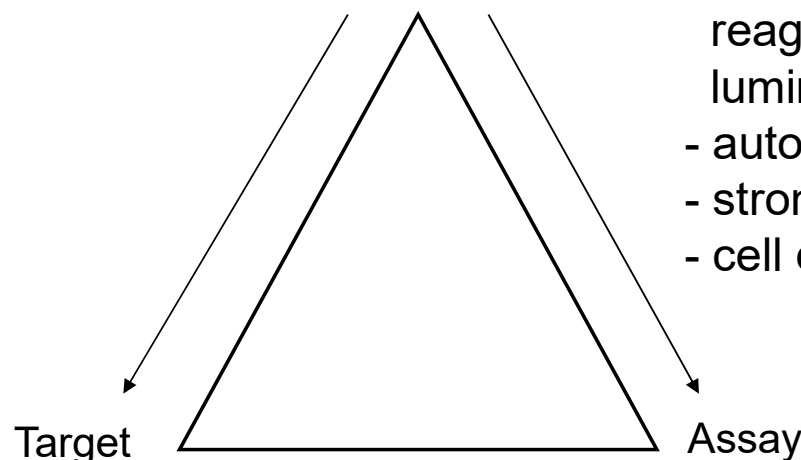
Compound

- covalent reaction with protein
- unspecific binding
- aggregation

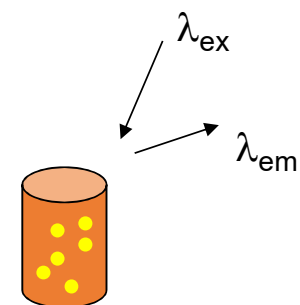
- interference with detection reagents (alphascreen, luminescence)
- autofluorescence (FRET, FP)
- strong color (luminescence)
- cell death



Target

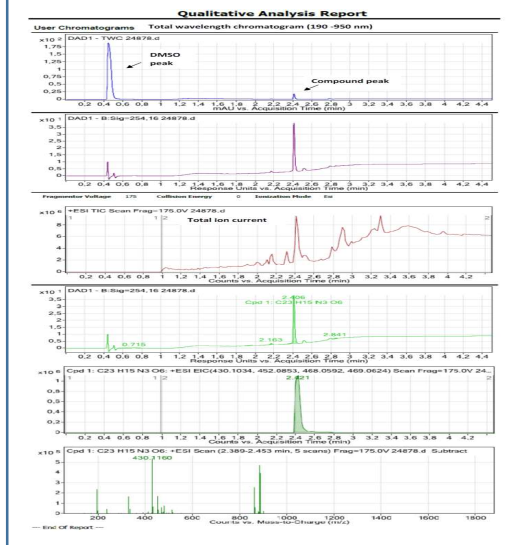


Assay

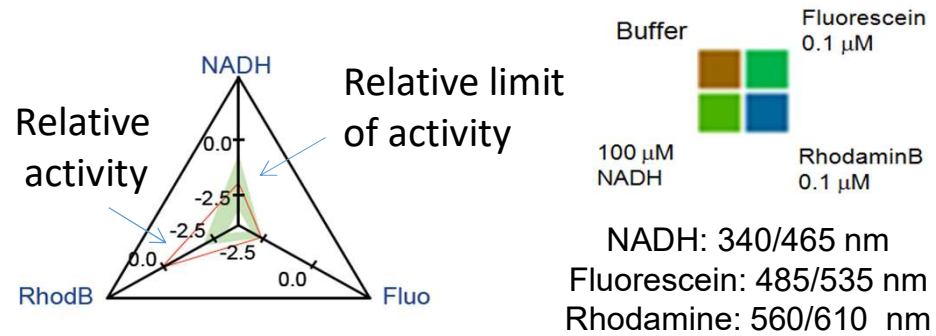


Analyse the Properties of Compounds

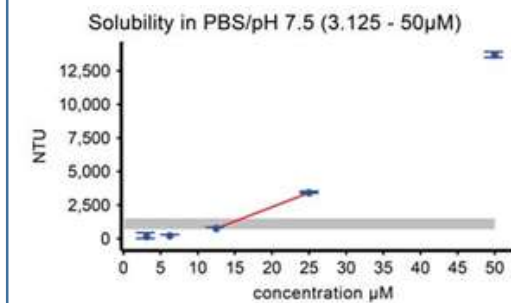
Purity of each cmpd



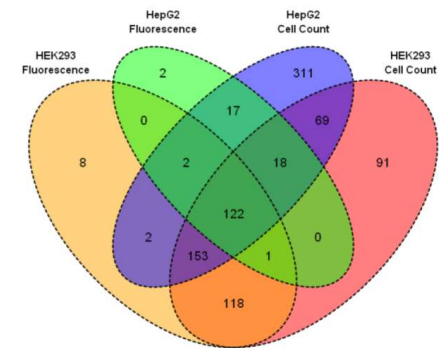
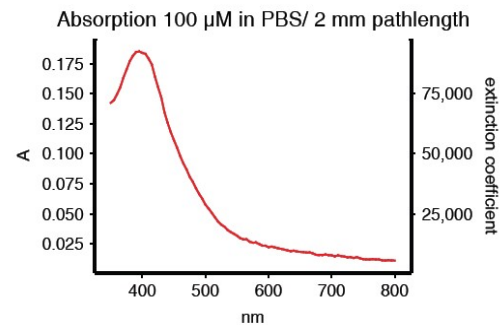
Fluorescence of each cmpd



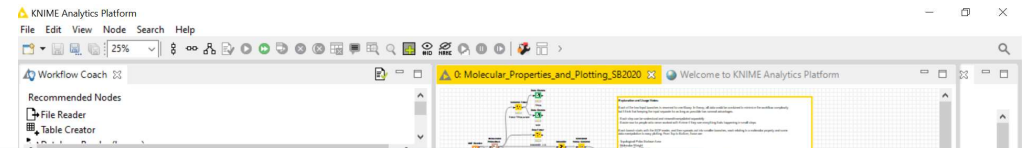
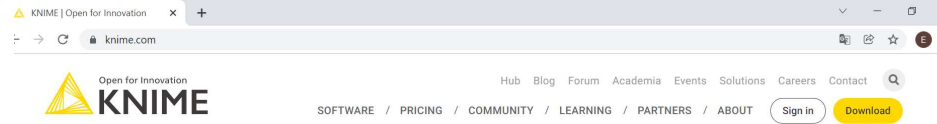
Solubility of each cmpd



Absorbance and cytotoxicity of each cmpd



Konstanz Information Miner (www.knime.org)



End to End Data Science

At KNIME, we build software to create and productionize data science using an intuitive environment, enabling stakeholders in the data science process to focus on what they do best.

Cheat Sheet: Building a KNIME Workflow for Beginners

GETTING STARTED

- Read through the installation guide at www.knime.org/installation
- Check out the [KNIME Learning Center](https://www.knime.org/learning) at www.knime.org/learning for more information on how to get started with KNIME.
- Take the [KNIME Learning Center](https://www.knime.org/learning) at www.knime.org/learning for more information on how to get started with KNIME.

READ

- CSV Reader:** Reads CSV files. It has an auto-detect function to automatically guess the file structure. As for some special cases, you can also specify the column names and data types.
- Excel Reader:** Reads content from sheets in Excel files (XLS, XLSX). Sheet and table names can be defined in the configuration window.
- Table Reader:** Reads content from a table in a database. The table name and database name can be defined in the configuration window.
- JSON Reader:** Reads content from JSON files. The file path and table name can be defined in the configuration window.

TRANSFORM

- GroupBy:** Groups the rows of a table by the unique values in selected columns and calculates the average, sum, min, max, and other statistics for each group.
- Math Formula:** Implements a number of math functions on selected columns. It also allows for conditional logic and the use of variables.
- String Manipulation:** Performs operations on string values in columns, such as concatenating, splitting, and replacing.
- Joiner:** Joins rows from two data tables based on common keys to create a new table with all columns.
- Sorter:** Sorts the table by ascending or descending order based on the value of a column. It also allows for sorting by multiple columns.
- Missing Value:** Defines a strategy to deal with missing values in the input data table, such as replacing them with a specific value or removing the rows.

DEPLOY

- Export to CSV:** Writes the input data table to a CSV file or to a remote location (e.g., SFTP).
- Export to Excel:** Writes the input data table to an Excel file (XLS or XLSX).
- Export to Database:** Writes the input data table to a database.
- Export to Table:** Writes the input data table to a table in a database.

Converts a raw data file to assay report

Workflow diagram showing the process: SDF Reader (Node 1) → Row Filter (Node 2) → Column Rename (Node 3) → Assay Report (Node 4).

Row ID	SMILES	Name	Library	Batch	PLATE	WELL	EC50	volume	source	concentration	composition
Row0	<chem>CC1=CC=C(C=C1)C2=CC=CC=C2</chem>	EC50103521	Commercial	2007	C1011L2007.01	A1	EC50103521	30	LV2002621761	10	641617
Row1	<chem>CC1=CC=C(C=C1)C2=CC=CC=C2</chem>	EC50103522	Commercial	2007	C1011L2007.01	C1	EC50103522	30	LV2002621773	10	960375
Row2	<chem>CC1=CC=C(C=C1)C2=CC=CC=C2</chem>	EC50103523	Commercial	2007	C1011L2007.01	E1	EC50103523	30	LV2002621785	10	611217
Row3	<chem>CC1=CC=C(C=C1)C2=CC=CC=C2</chem>	EC50103524	Commercial	2007	C1011L2007.01	G1	EC50103524	30	LV2002621801	10	653847

216407

FMP-216407
ChemDiv N038-0218
Structurally similar: [56 activities found in 100-95% similar molecules](#)

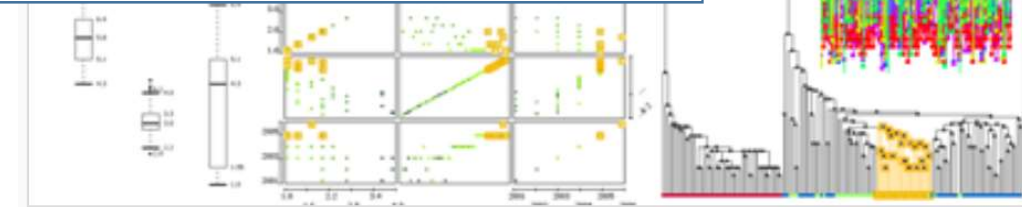
Fluorescence increase

Optical properties:
Absorbance @100 uM in PBS, 2 mm path length:
10 mM in DMSO, scan:

PAINS analysis:
CHEMBL bioactivities: [56 activities found in 100-95% similar molecules](#)

Specificity (primary screening):
Absorbance @100 uM in PBS, 2 mm path length:
10 mM in DMSO, scan:

LC-MS analysis:
PurityValue (%): 100.0
Redox cycling assay:
Cytotox assay: HEK293, HepG2

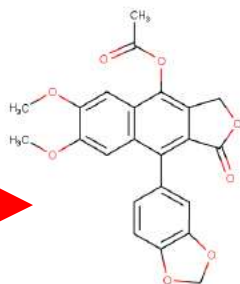


<http://informationandvisualization.de/blog/knime-interactive-views#box>

Example of Assay Report

Cmpd structure

216407



FMP-216407

ChemDiv N038-0218

Structurally similar:

PAINS analysis:

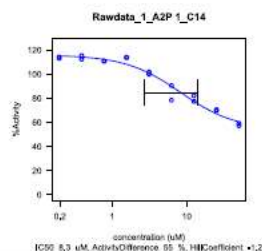
ChEMBL bioactivities:
[56 activities found in 100-95% similar molecules](#)

No PAINS structure

56 activities in ChEMBL

IC₅₀-Evaluation

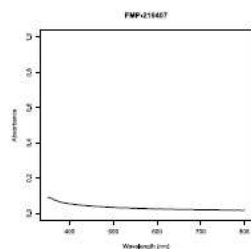
Fluorescence increase



Wavelength scan

Optical properties:

Absorbance @100 uM in PBS, 2 mm path length:



Well picture

10 mM in DMSO, scan:



Specificity (primary screening):
(inhibitors/activators) of all_measured:

(0/0) of 7 in alphascreen
(0/1) of 4 in HTRF
(0/0) of 4 in elisa
(2/1) of 7 in firefly
(0/1) of 4 in general_cell
(0/0) of 6 in LabChip
(0/0) of 2 in FluoPol
(1/0) of 11 in biochemical
(0/0) of 1 in cAMP
(0/0) of 2 in HCS

(3/3) of 48 in total

LC-MS analysis:
PurityValue (%): 100.0

Redox cycling assay:

Cytotox assay:
HEK293, HepG2

LC-MS Purity

Cytotoxic in two cell types

Active in 3 assays

designed by Martin Neuenchwander
Neuenchwander@fmp-berlin.de

EU-OPENSOURCE Libraries

European Chemical Biology Library (ECBL)

Diversity library

- 99.096 structurally highly diverse compounds
- Average MW=350 g/mol
- 0.05 % of PAINS

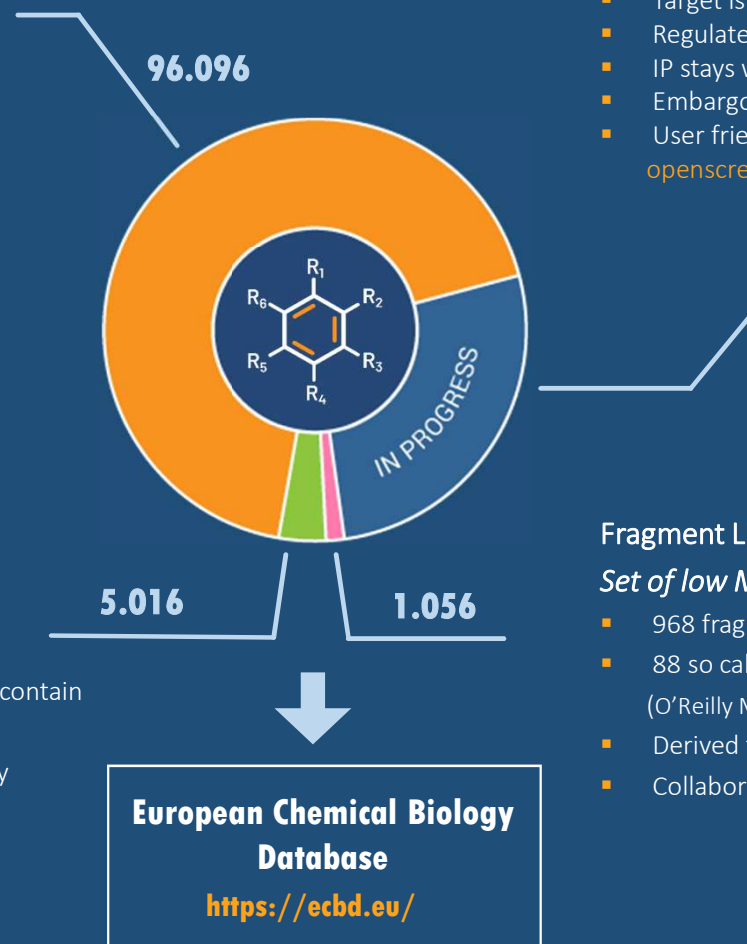
Horvath D. *et al.*, *ChemMedChem* 2014, 9, 2309



European Chemical Biology Library (ECBL)

Pilot library

- 2.464 bioactives: active against 1039 different targets, contain 654 approved drugs and 368 highly selective probes
- 2.464 representative compounds of the diversity library
- 88 assay interference compounds in 4 dilutions



The European Academic Compound Library (EACL)
Novel compounds sourced from chemists worldwide

- Target is 40.000 compounds
- Regulated and confidential access (e.g. MTA)
- IP stays with the chemist
- Embargo period up to 3 years
- User friendly online submission: <http://www.eu-openscreen-cmpds-donation.eu/login.php>

Fragment Library **NEW in 2020!**

Set of low MW and ultra-low MW fragments

- 968 fragments with HAC > 8 in DMSO-*d*₆
- 88 so called "minifrag" with HAC < 8 in DMSO-*d*₆ (O'Reilly M. *et al.*, *Drug Discov. Today* 2019, 24, 1081)
- Derived from the fragment space of the ECBL
- Collaboration with INSTRUCT/ iNEXT-Discovery sites

Library Quality and Questionnaire for Screening

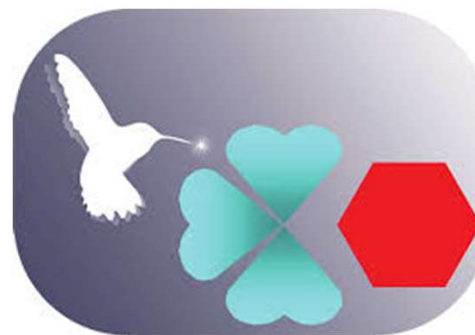
- **Highly diverse** with a plethora of different selection concepts
- **Small clusters** of similar compounds for early SAR
- Lower numbers in property distribution compared to Rule of 5 (Lipsinski Rules) e.g Rule of 3
- How unique are the structures? What is known about similar compounds in the literature?
- Comparison with **ChEMBL** (publications: <https://www.ebi.ac.uk/chembl/>)
- Comparison with **SureChEMBL** (patents: <https://www.surechembl.org/>)
- Cmpds are **synthesizable, easily to derivatize and scalable?**
- Open access to **negative screening results?**
- **Transparency about the selection** process of the hit list?
- **User involvement in the selection** process of the hit list?
- Access to specific libraries for specific targets and applications
- Cmpds property knowledge to filter for false positives.

Comparison to External Libraries



National Institutes
of Health

NIH Library (USA)
**882.649 commercial and academic
compounds (PubChem)**



Chimiothèque Nationale CN (France)
62.824 academic compounds (2018)



Broad Institute (USA, Boston)
5.681 bio-annotated compounds

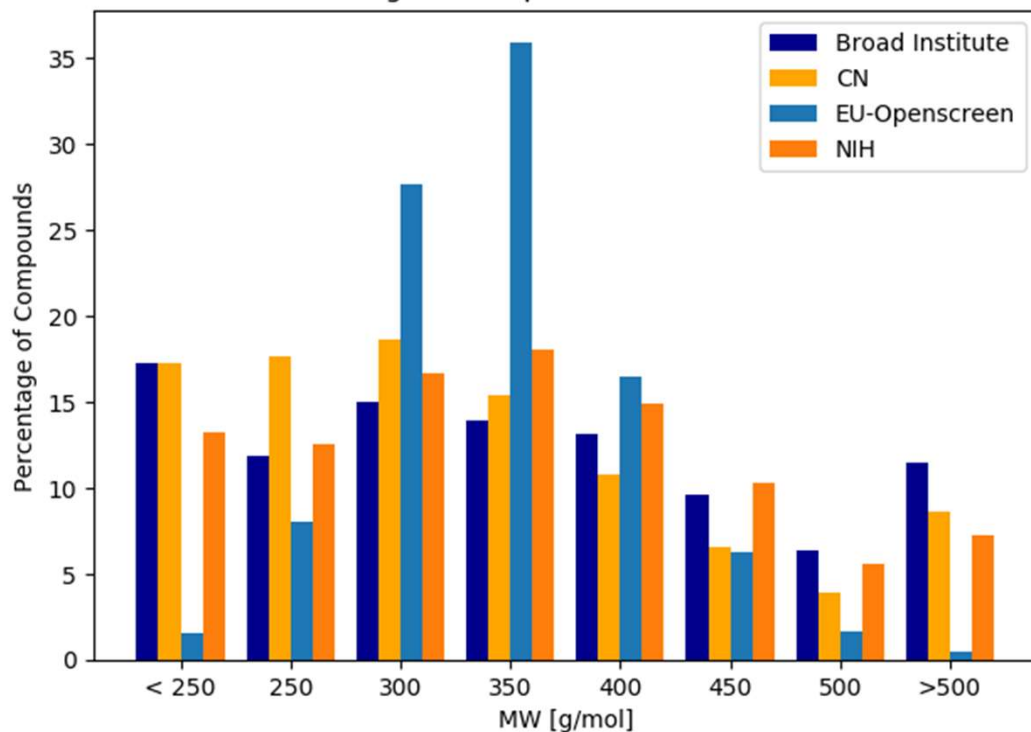
All Libraries suitable for Screening

NIH, Broad and CN with broader portfolio of compound properties

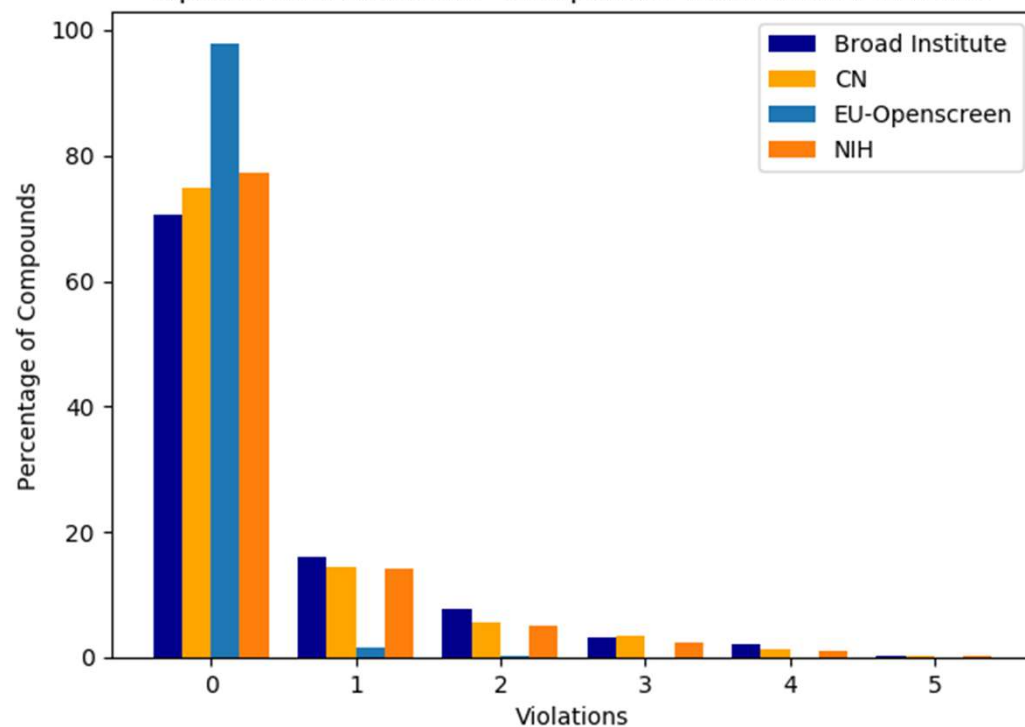
EU-OPENSREEN with lower number of Lipinski violations: 25 year old concept

(Ro5: MW < 500 Da, ClogP < 5, HBD < 5, HBA < 10)

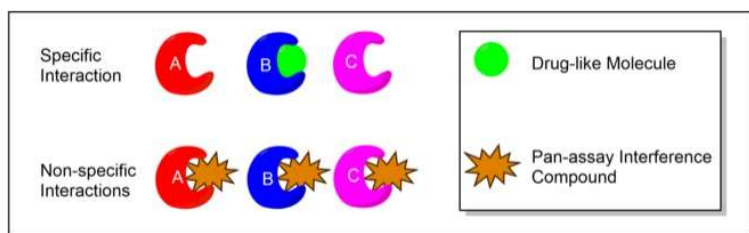
Molecular Weight - Comparison with external libraries



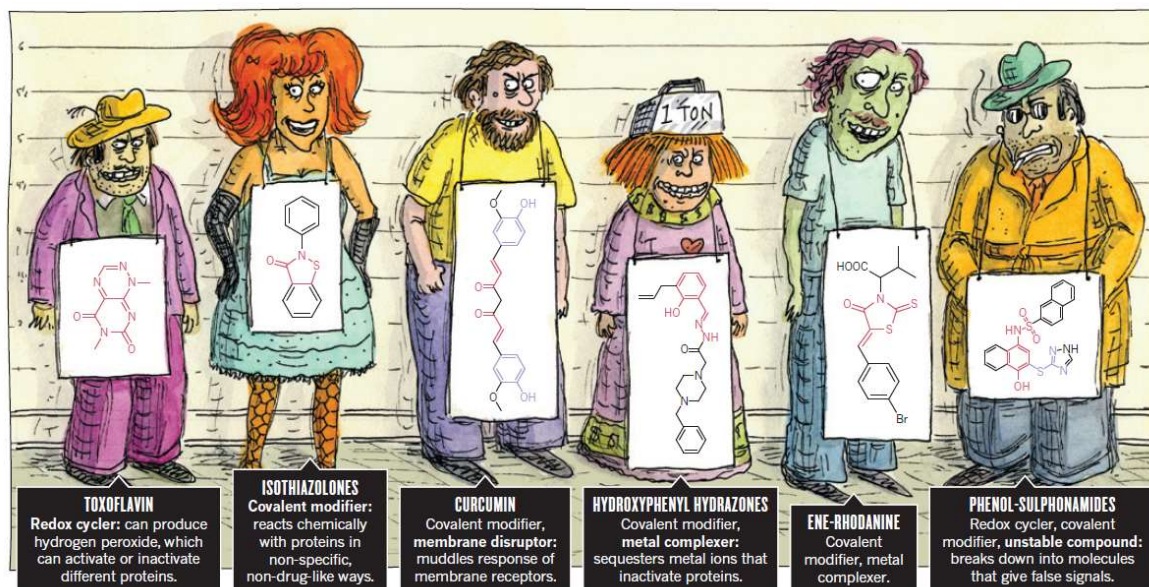
Lipinski Rule Violations - Comparison with external libraries



PAINS (Pan Assay INterference compounds)



Baell JB, Holloway GA. *New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays*, Journal of Medicinal Chemistry, 2010, 53, (7): 2719–40.



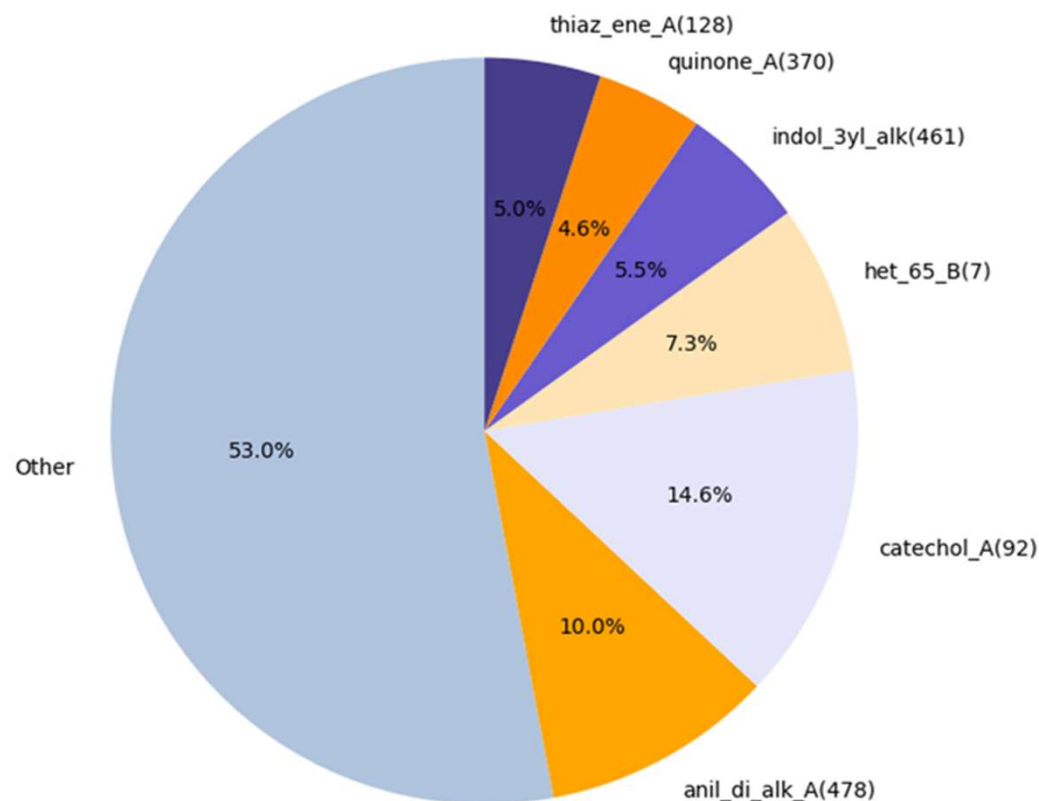
Jonathan Baell and Michael A. Walters. *Chemistry: Chemical con artists foil drug discovery* Nature, 2014, 513 (7519): 481–483.

PAINS Compounds

**219 out of 101.276 compounds
are classified as PAINS → 0.2 %**

**without bioactives library
classified as PAINS → 0.05 %**

Distribution of PAINS in ECBL



Delivery of Compounds in 384 Well Plates

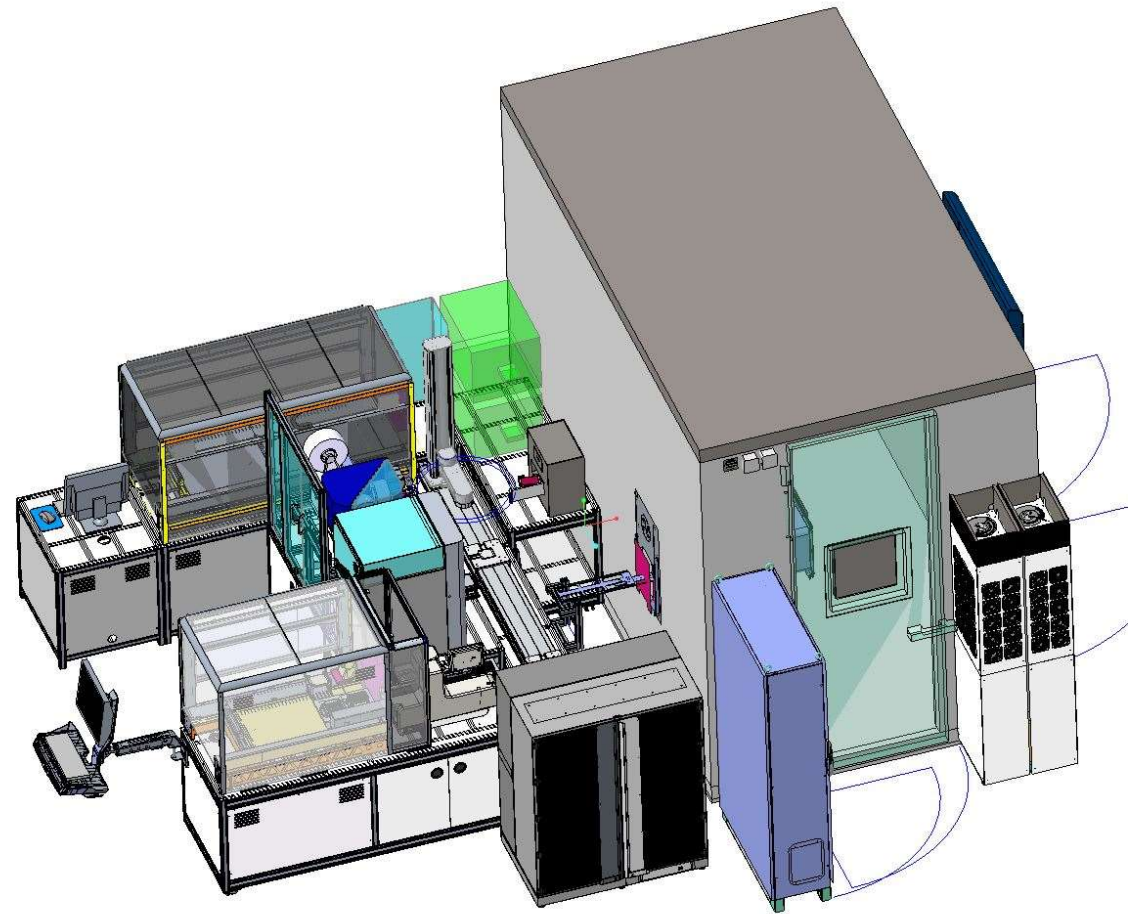
**290 plates with 352 compounds to 10 Partner Sites in 1-3 copies
=> 1.400.000 Compounds**



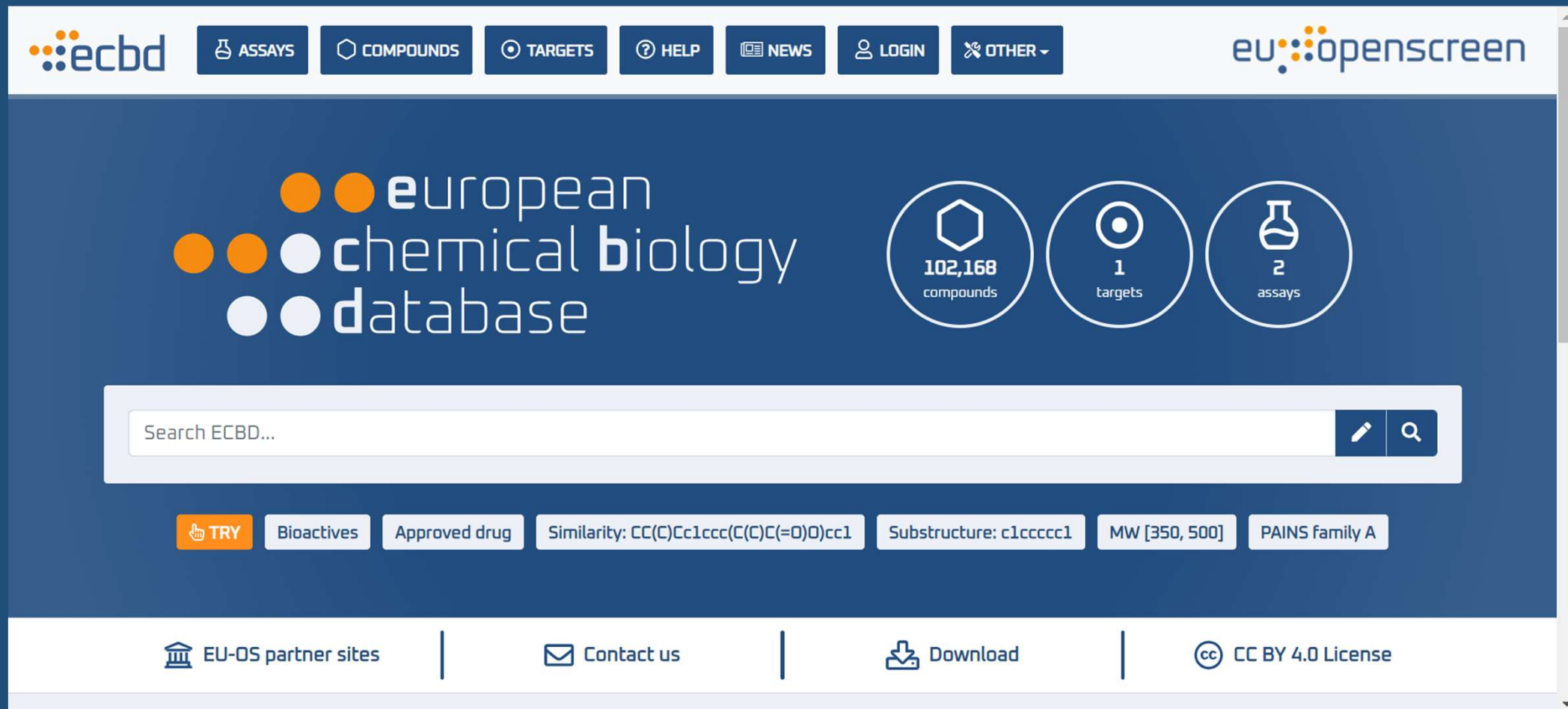
Hamilton Verso Store with Hamilton Star Robot



Capacity of about 300.000 Tubes



<https://ecbd.eu>



The screenshot shows the homepage of the European Chemical Biology Database (ECBD). The header features the ECBD logo on the left and the eu:openscreen logo on the right. A navigation bar contains buttons for ASSAYS, COMPOUNDS, TARGETS, HELP, NEWS, LOGIN, and OTHER. The main content area displays the ECBD logo and three circular statistics: 102,168 compounds, 1 target, and 2 assays. Below this is a search bar with the placeholder text "Search ECBD...". A filter bar includes a "TRY" button and several filter options: Bioactives, Approved drug, Similarity: CC(C)Cc1ccc(C(C)C(=O)O)cc1, Substructure: c1ccccc1, MW [350, 500], and PAINS family A. The footer contains links for EU-OS partner sites, Contact us, Download, and CC BY 4.0 License.

<https://ecbd.eu>; Institute of Molecular Genetics of the ASCR (CZ-OPENSREEN: National Infrastructure of Chemical Biology)

European Chemical Biology Database (ECBD)

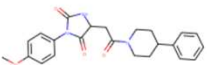
ecbd ASSAYS COMPOUNDS TARGETS HELP NEWS LOGIN OTHER - Search ECBD...

Compounds 1 ... 10214 > 102135

Search in Compounds

- PROPERTIES 7
- LIBRARIES 7
- COMPOUND TYPE 2
- STRUCTURAL ALERT 3

E051



Compound

Name: E051

Library: European Chemical Biology Library
ECBL Pilot Compounds
Representative set of the diversity library

Tags:

InChIkey: BDSKOWUCRDT5BA-UHFFFAOYSA-N

ROS: MW 407.47, HBD 1, HBA 7, RB 5, cLogP 2.92

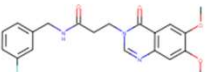
Assays 0

stage	✓	x	?	...
primary	0	0	0	0
secondary	0	0	0	0
confirmatory	0	0	0	0
other	0	0	0	0

Links 3

- PubChem: [91820963](#)
- MolPort: [MolPort-039-017-466](#)
- MolPort: [MolPort-042-671-876](#)

E052



Compound

Name: E052

Library: European Chemical Biology Library
ECBL Pilot Compounds
Representative set of the diversity library

Tags:

InChIkey: BBHNGOSSYMSMBL-UHFFFAOYSA-N

ROS: MW 385.39, HBD 1, HBA 7, RB 7, cLogP 2.26

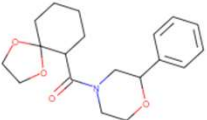
Assays 0

stage	✓	x	?	...
primary	0	0	0	0
secondary	0	0	0	0
confirmatory	0	0	0	0
other	0	0	0	0

Links 5

- PubChem: [39351211](#)
- ZINC: [ZINC000032103955](#)
- Mcule: [MCULE-2036913492](#)
- MolPort: [MolPort-006-810-137](#)
- eMolecules: [27458340](#)

E053



Compound

Name: E053

Library: European Chemical Biology Library
ECBL Pilot Compounds
Representative set of the diversity library

Tags:

InChIkey: FDNIQBQHPFKADZ-UHFFFAOYSA-N

Assays 0

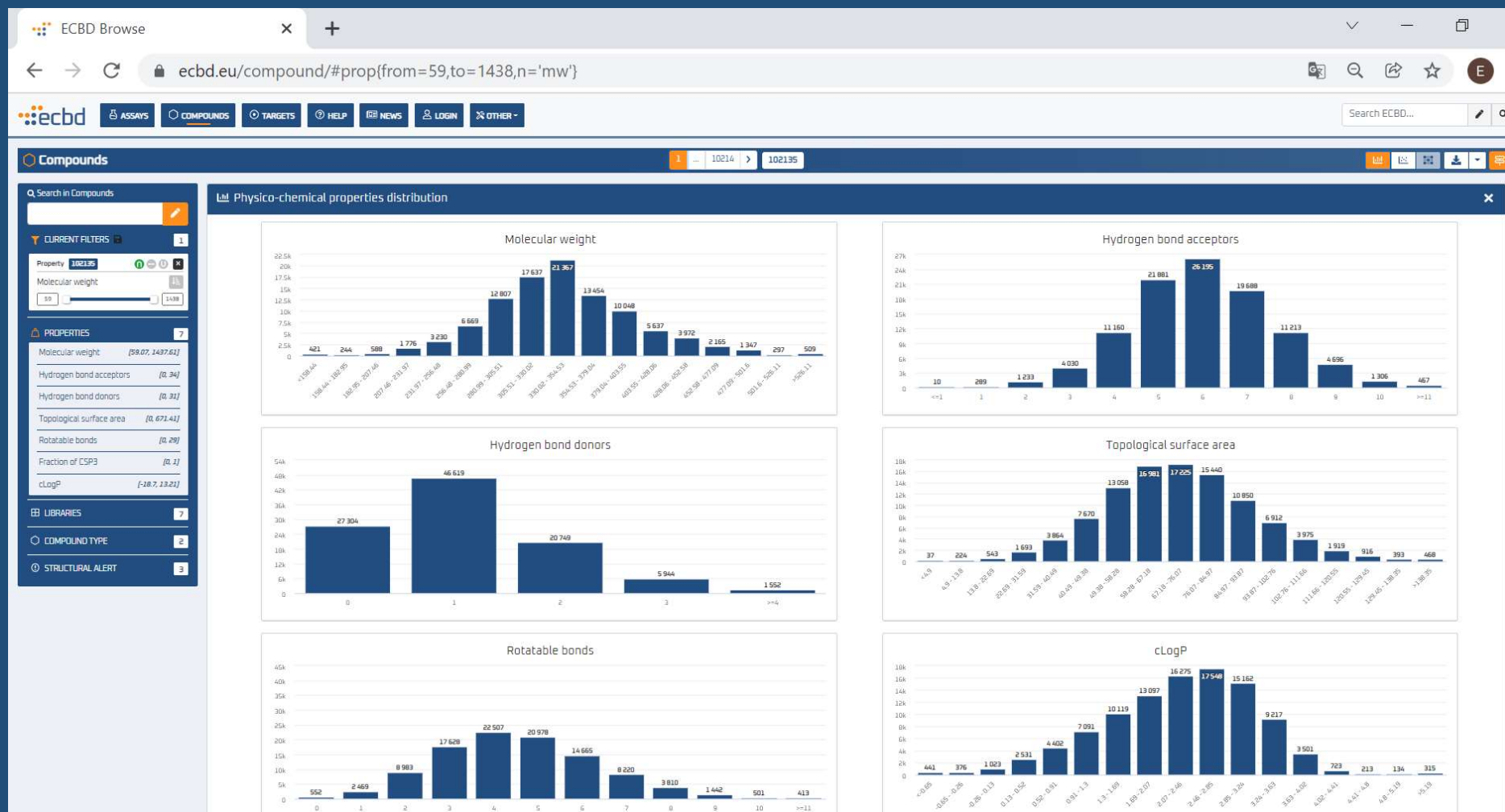
stage	✓	x	?	...
primary	0	0	0	0
secondary	0	0	0	0
confirmatory	0	0	0	0
other	0	0	0	0

Links 2

- PubChem: [118986097](#)
- MolPort: [MolPort-039-347-465](#)

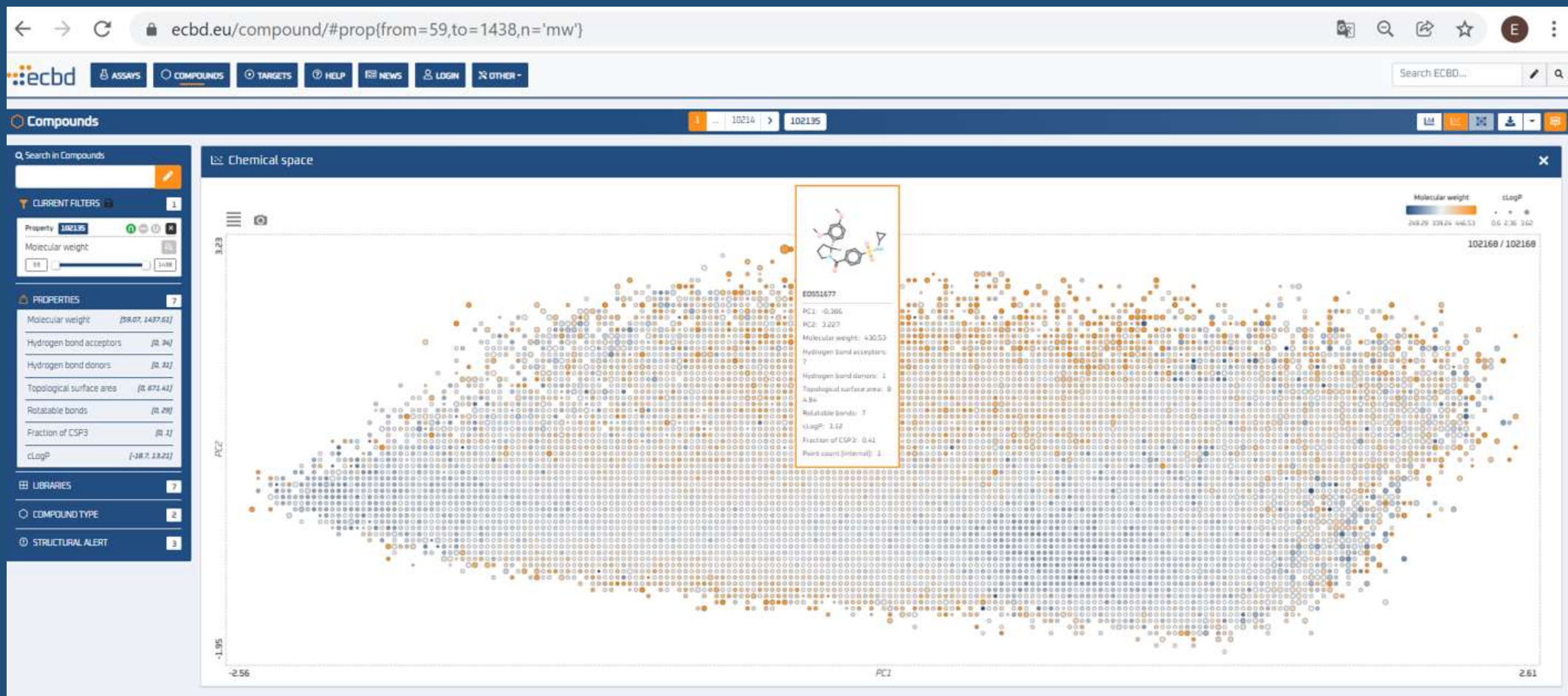
<https://ecbd.eu>; Institute of Molecular Genetics of the ASCR (CZ-OPENSREEN: National Infrastructure of Chemical Biology)

Compound Property Distribution



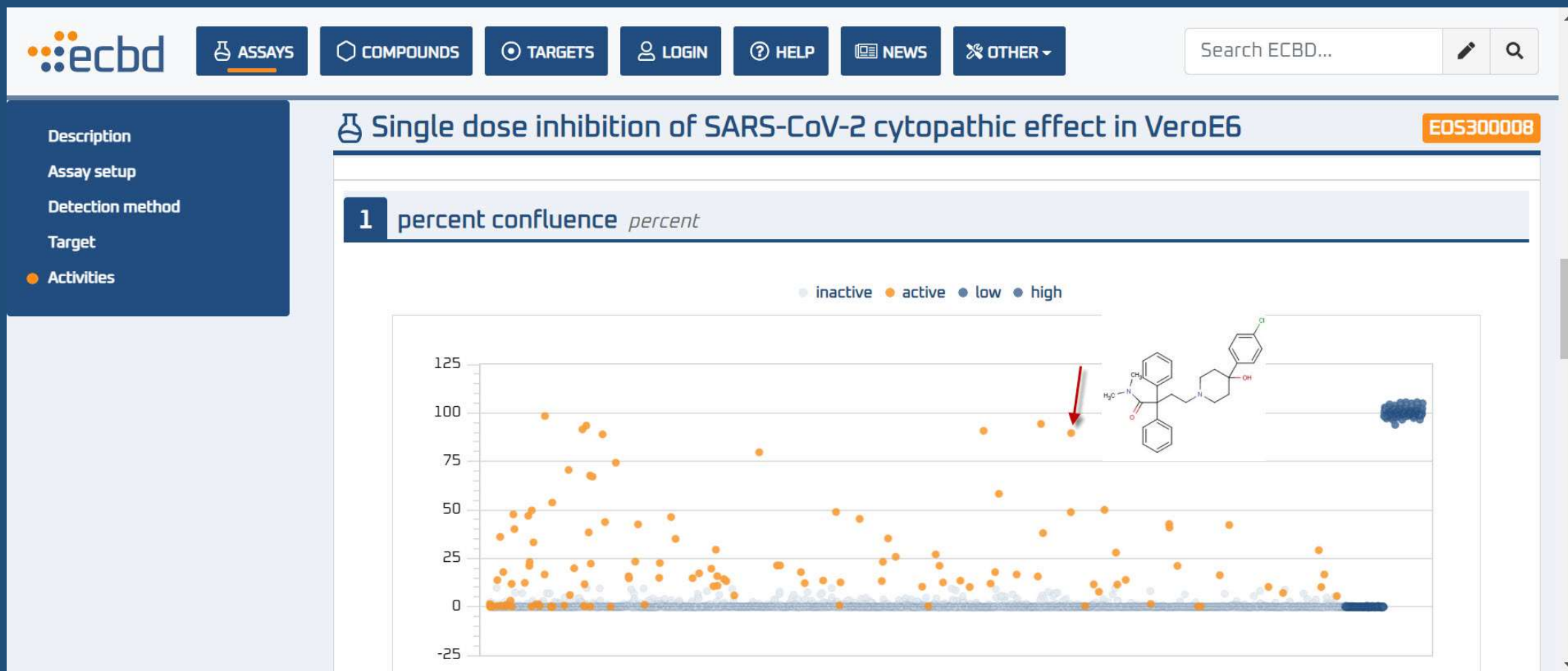
<https://ecbd.eu>; Institute of Molecular Genetics of the ASCR (CZ-OPENSREEN: National Infrastructure of Chemical Biology)

Coverage of Chemical Space



<https://ecbd.eu>; Institute of Molecular Genetics of the ASCR (CZ-OPENSREEN: National Infrastructure of Chemical Biology)

First Screening Assay Published on the ECBD Website





Chemical Biology Platform of the FMP

Structural Chemistry & Computational Biophysics

Han Sun
Michael Lisurek
Bernd Rupp
Raed Al-Yamori
Songhwan Hwang
Tillmann Utesch
David Bushiri Pwesombo
Haoran Liu
Florian Heiser
Xiaolu Li
Johann Biedermann
Berke Türkaydin
David Zierke

Screening Unit

Jens von Kries
Andreas Oder
Carola Seyffarth
Silke Radetzki
Martin Neuenschwander
Katina Lazarow
Sabrina Kleissle
Felix Hansen
Christopher Wolff
Astrid Mühl



Medicinal Chemistry

Marc Nazaré
Peter Lindemann
Lioudmila Perepelittchenko
Keven Mallow
Jerome Paul
Sandra Mischke
Malgorzata Wasinska-Kalwa
Carolina Vinagreiro
Leonard Mach
Axel Hentsch
Monica Guberman
Davide Cirillo

Collaboration Partners & Funding



Stefan Krauss (Univ. Oslo)

Michael Bader (MDC, Berlin)

Walter Birchmeier (MDC, Berlin)

Udo Heinemann (MDC, Berlin)

The EU-OPENSOURCE Team



Director
General

Wolfgang Fecke

Phone: +49 (0)30 9489 2420



Head of
Operations &
Scientific
Strategy

Bahne Stechmann

Phone: +49 (0)30 9489 2421



European
Relations,
Grant and
Scientific
Project
Manager

Dr. Katja Herzog

Phone: +49 (0)30 9489 2422



Scientific
Project &
Industry
Liaison
Manager

Robert Harmel

Phone: +49 (0)30 9489 2429



Coordination
Manager
Scientific
Services

Kathy Skopelitou

Phone: +49 (0)30 9489 2417



Interim
Compound
Manager

Edgar Specker

Phone: +49 30 9489 2426



Finance &
Process
Manager

Christian Gujjarro Reznicek

Phone: +49 (0)30 9489 2414



Communi-
cations & HR
Manager

Maren Kappe

Phone: +49 (0)30 9489 2424



Scientific
Project
Manager

Tanja Miletic

Phone: +49 (0)30 9489 2428



Laboratory
Automation
Manager

Sophie Masin
(formerly Brusseau)

Phone: +49 (0)30 9489 2413



Analytical
Chemist

Federica Rossella

Phone: +49 (0)30 9489 2418



Analytical
Chemist

Victoria Mora

Phone: tba



Accountant

Annegret Haden

Phone: (+49) 030 9489 2417



Senior
Advisor

Ronald Frank

Interested in HTS with us? Please contact us!

Please get an overview about the Partner Sites:

<https://www.eu-openscreen.eu/services/screening.html> for Screening contacts

<https://www.eu-openscreen.eu/services/medicinal-chemistry.html> for MedChem contacts

<https://www.eu-openscreen.eu/services/database.html> for the Database contact

or everyone in the EU-OPENSREEN Team:

<https://www.eu-openscreen.eu/about/contact-team.html>

Please also visit the overview at ECBD

<https://ecbd.eu/organization/>

Any information of the FMP library can be found here:

<https://www.leibniz-fmp.de/compound-management> or <https://www.leibniz-fmp.de/the-screening-unit/downloads>