

EU Life CONCERT REACH Gateway - course

09/06/2023, QSAR2023

Nelly Giuseppa Raitano



THE PROJECT



September

June

2018

2019

2020

2021

2022

2023



Associated Beneficiaries



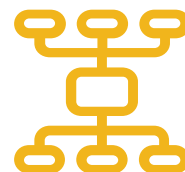
Istituto di Ricerche Farmacologiche Mario Negri, Italy
Research institute, **Coordinator**



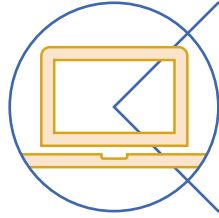
Support

Evaluate the **potential impact** of CS in the
EU by *exp + in silico*

A big network of systems offering non-
testing methods (NTM) useful both for
authorities and industries.



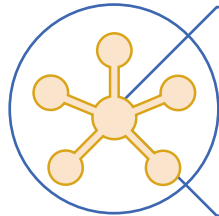
Results



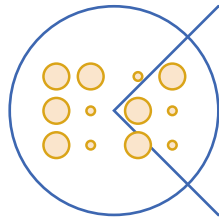
The network offers **more than 450 *in silico* tools**



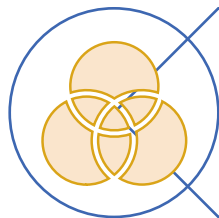
42 new *in silico* models



A new tool for grouping

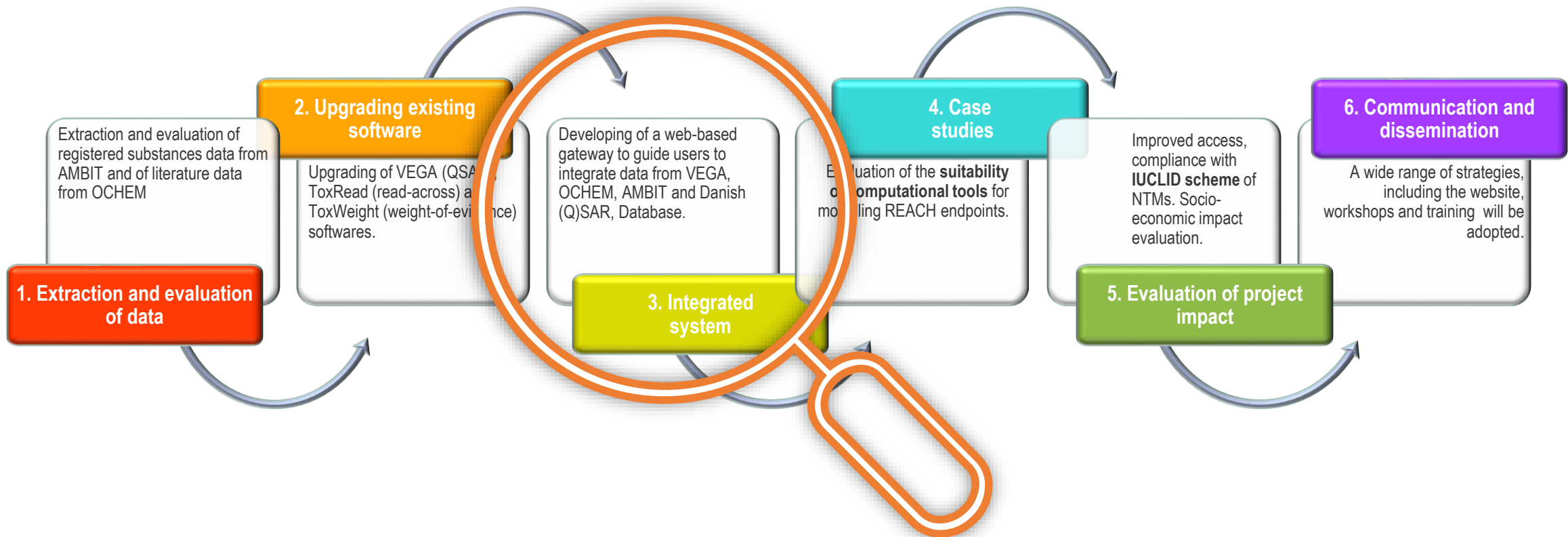


Updated version of **read-across tools**, freely available and user-friendly



New strategy integrating the results of read-across and *in silico* models (**weight-of-evidence**).

PROJECT ACTIVITIES

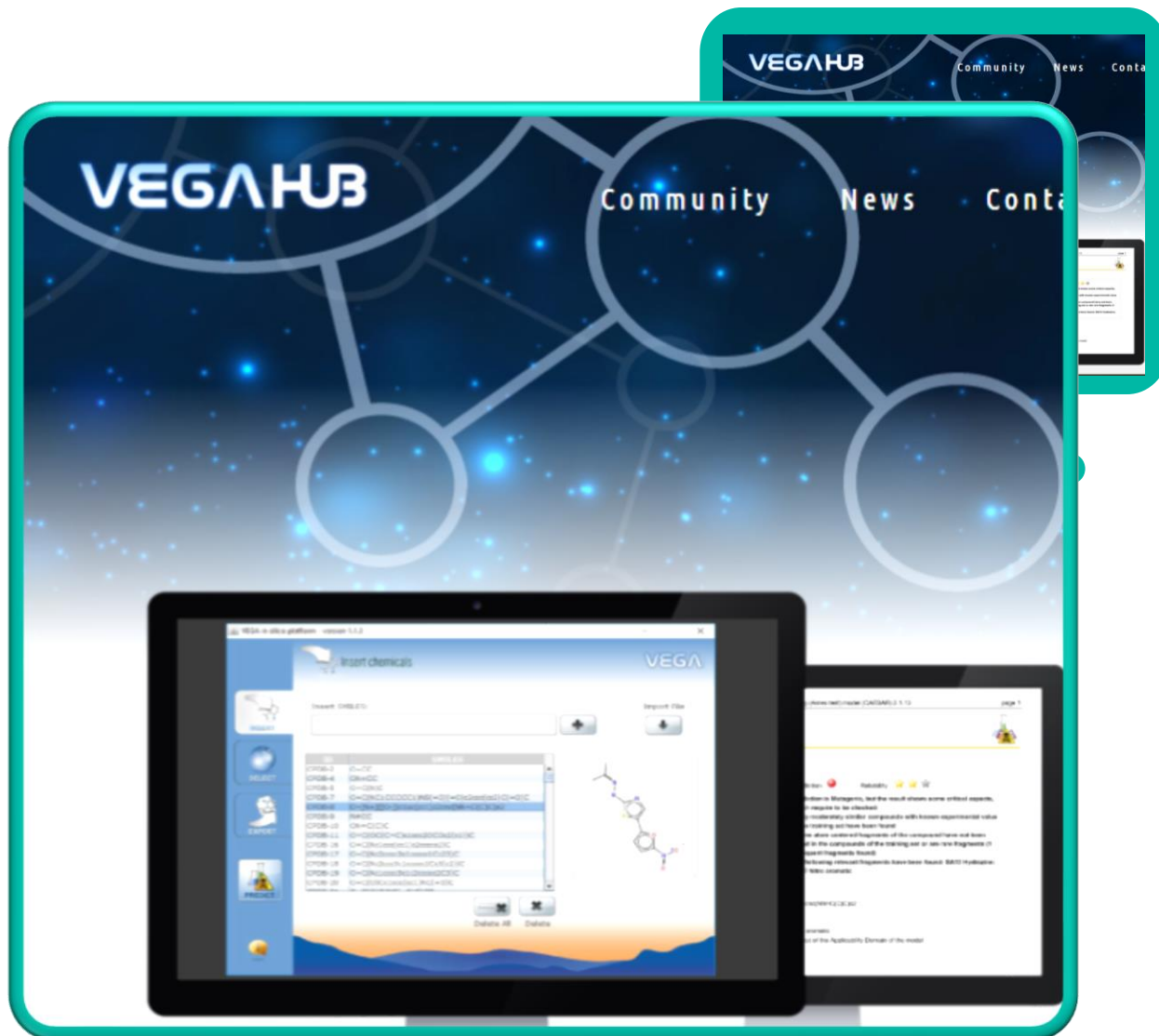


THE TOOLS



The network is composed of **VEGA**, the **Danish (Q)SAR Database**, **OCHEM** for the *in-silico models*, and for the read across workflow and data from the registered substances, of **ToxRead** and **AMBIT**.

THE TOOLS



110 (Q)SAR freely available models for regulatory purposes.

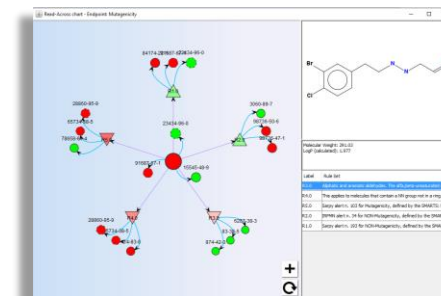
Different areas:

- Human toxicity
- Eco-toxicity
- Environmental
- Physico-chemical
- Toxicokinetics



Reproducible read-across evaluation for 25 endpoints showing similar compounds and SAs in common between chemicals.

TOXREAD



THE TOOLS



LIFE17 GIE/IT/000461

The OCHEM package offers a database of molecules and their ADMET properties.

Online chemical database
with modeling environment

Home Database Models

Welcome to OCHEM! Your possible actions

Check out the properties available on OCHEM

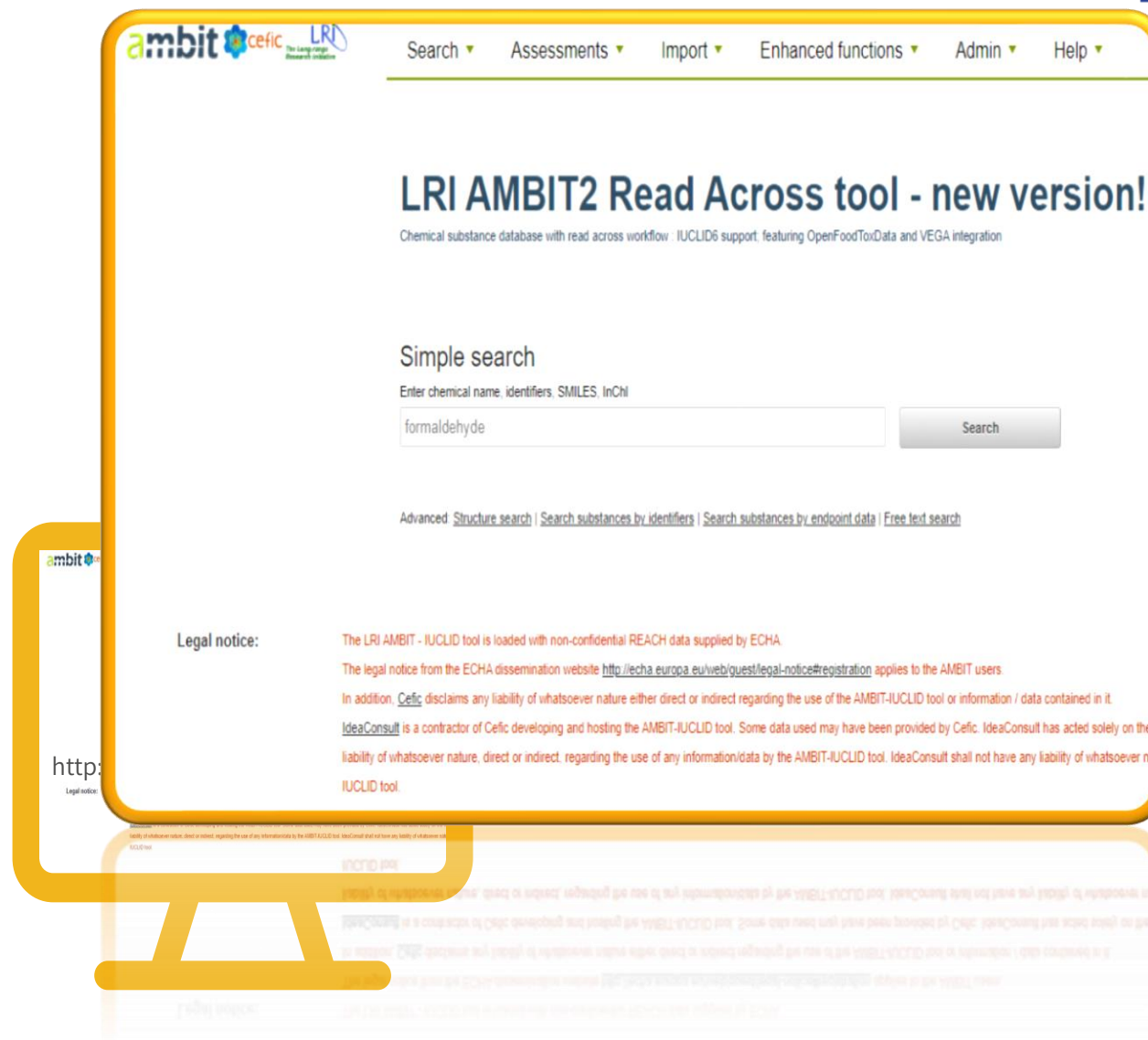
OCHEM contains 3345610 records for 689 properties (with at least 50 records) collected from 15083 sources

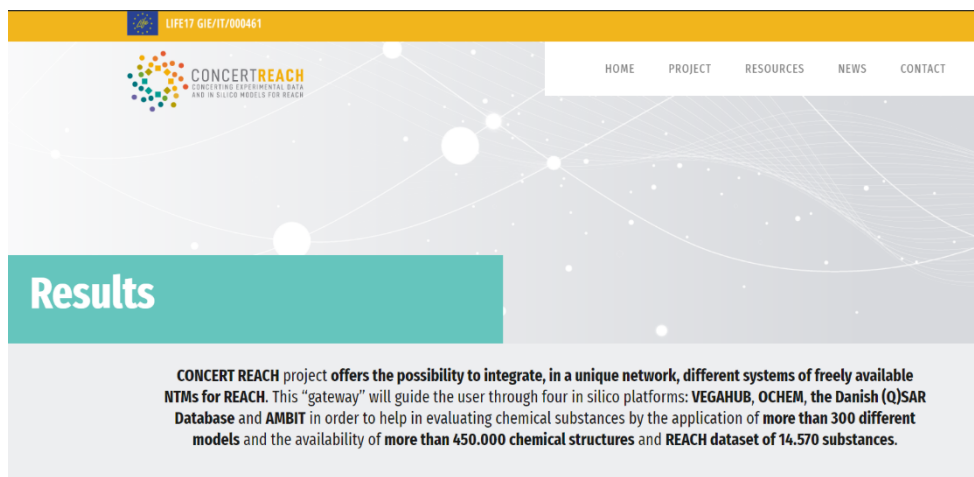
Melting Point logPow logBB LogL(water) LogD logPI(+)
Water solubility LogL(blood) LogL(oil) ER Cbrain/Cplasma IC50 Papp(Caco-2)
Papp(MDCK) Oral absorption LIC 50 Papp ratio(Caco-2)
Plasma protein binding Papp ratio(MDCK-mdr1) pIC50 %Human FA Human IA
Human FA fraction unbound (fu) fraction ionized (f) pKa VDss LogIC50 LogPI
BBB permeability (qualitative) LogKoa LogRBA CYP450 modulation
CYP450 reaction Vapor Pressure EC50 aquatic NOEC aquatic
LOEC aquatic IC50 aquatic LC50 aquatic log(IGC50-1) LEL
Henry's law constant EC50 EROD induction LC 50 Boiling Point LD50 dermal
LD50 oral LC50 terrestrial AMES LD50 Biodistribution
Water solubility Kinetic Papp(PAMPA) IC50 CYP450 Inhibition Ki CYP450
logK' hsa Dissipation half-life DT50 Freundlich coefficient KT BMF
Atmospheric OH Rate Constant Ki TDLo LDLo Carcinogen Anti-inflammatory activity
Methanol solubility LogLD50 MIC Retention Time Surface tension Cblood/Cair(Human)
Cfat/Cair(Rat) Cbrain/Cair(Rat) Cliver/Cair(Rat) Cmuscle/Cair(Rat) IC50 PDE4 % inhibition PDE4
IC50 inhibition Density pKa (smiles as ob. cond.) DMSO Solubility
log Kb logk'0 logLOAEL hERG K+ Channel Blocking (IC50) 5-HT2B (Ki) LogKoc
BCF CHSEL % inhibition hERG, K+ Channel Blocking hERG K+ Channel Blocking (Ki)
logP Chloroform/Water 5-HT2C (Ki) 5-HT2b (Kb) Pgp substrate 5-HT2A (Ki) D2R (Ki) o1 adrenergic receptor (Ki)

OCHEM contains more than 1 million experimental records for about 499 properties collected from 12428 sources

THE TOOLS

The AMBIT system consists of a database including more than **450.000 chemical** structures and REACH data on **14.570 substances**



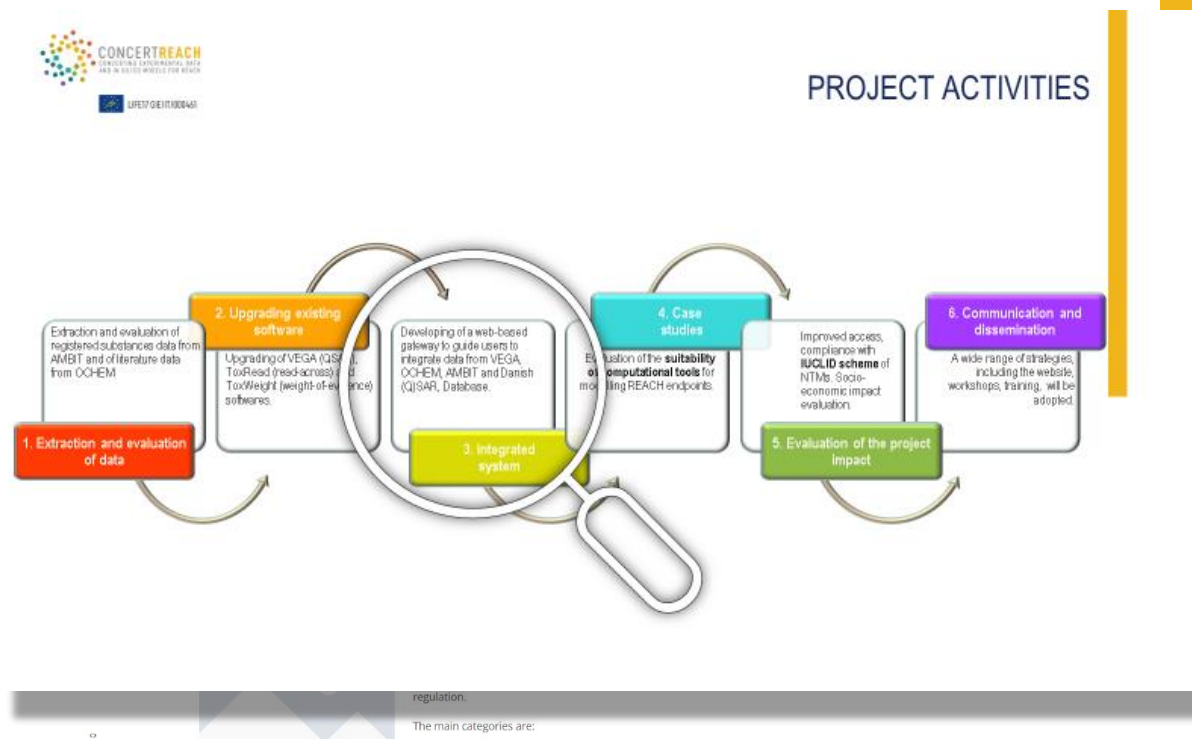


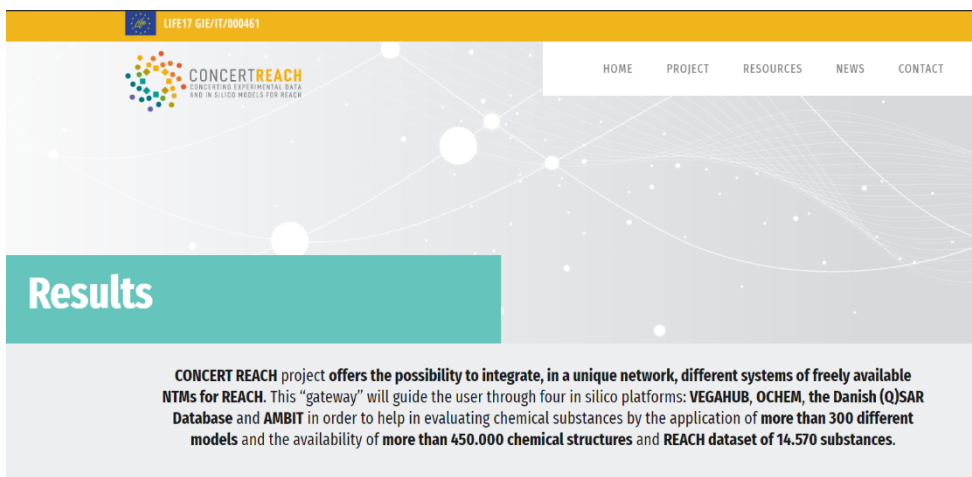
Results

CONCERT REACH project offers the possibility to integrate, in a unique network, different systems of freely available NTMs for REACH. This "gateway" will guide the user through four in silico platforms: VEGA HUB, OCHEM, the Danish (Q)SAR Database and AMBIT in order to help in evaluating chemical substances by the application of more than 300 different models and the availability of more than 450.000 chemical structures and REACH dataset of 14.570 substances.

The "gateway" reports all the predictive software available in the four platforms relative to REACH endpoints.

However, please consider that we cannot guarantee that they are correct and usable for the REACH legislation. Additionally, if industry wants to use the result from a certain model, it has to verify if this is legally legitimate. For certain very specific endpoints we have reported models that may have been developed using more general data. These models may not perfectly adhere to the endpoint.





LIFE17 GIE/IT/000461

HOME PROJECT RESOURCES NEWS CONTACT

Results

CONCERT REACH project offers the possibility to integrate, in a unique network, different systems of freely available NTMs for REACH. This "gateway" will guide the user through four in silico platforms: VEGAHUB, OCHEM, the Danish (Q)SAR Database and AMBIT in order to help in evaluating chemical substances by the application of more than 300 different models and the availability of more than 450.000 chemical structures and REACH dataset of 14.570 substances.

The "gateway" reports all the predictive software available in the four platforms relative to REACH endpoints.

However, please consider that we cannot guarantee that they are correct and usable for the REACH legislation. Additionally, if industry wants to use the result from a certain model, it has to verify if this is legally legitimate. For certain very specific endpoints we have reported models that may have been developed using more general data. These models may not perfectly adhere to the endpoint.



LIFE17 GIE/IT/000461

HOME PROJECT RESOURCES NEWS CONTACT

Gateway User Guide

This Guide will guide the user through four in silico platforms: **VEGAHUB, OCHEM, the Danish (Q)SAR Database and AMBIT** in order to help how the platforms works.

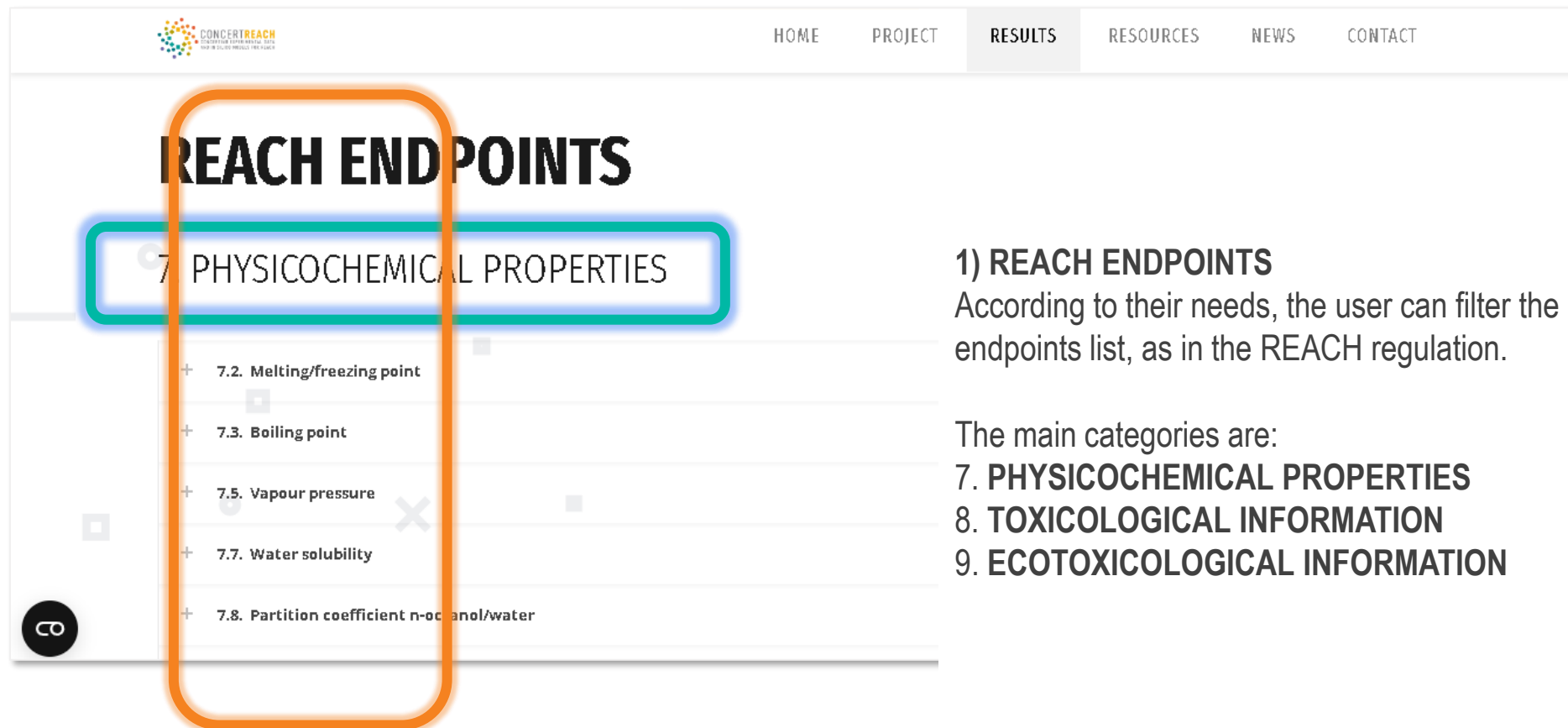
GUIDE TO THE USE OF THE GATEWAY



1. REACH ENDPOINTS

According to his/her needs, the user can filter the models by the endpoints list, as in the REACH regulation.

The main categories are:



HOME PROJECT **RESULTS** RESOURCES NEWS CONTACT

REACH ENDPOINTS

7. PHYSICOCHEMICAL PROPERTIES

- + 7.2. Melting/freezing point
- + 7.3. Boiling point
- + 7.5. Vapour pressure
- + 7.7. Water solubility
- + 7.8. Partition coefficient n-octanol/water

1) REACH ENDPOINTS

According to their needs, the user can filter the models by the endpoints list, as in the REACH regulation.

The main categories are:

- 7. PHYSICOCHEMICAL PROPERTIES
- 8. TOXICOLOGICAL INFORMATION
- 9. ECOTOXICOLOGICAL INFORMATION






2) SELECTION OF THE SUITABLE MODEL

<https://www.life-concertreach.eu/results/>

+ 7.5. Vapour pressure

- 7.7. Water solubility

All VEGA AND ToxRead DANISH QSAR DATABASE AMBIT OCHEM

End Point	Model	Type	Dataset size	Training set size	Test set size	Platform	Remarks
P-CHEM 4.8. Water solubility	Dataset		18126			AMBIT	
P-CHEM, 4.8 water solubility, OECD 105	Water solubility model (IRFMN)	continuous	5018	4014	1004	VEGA	
P-CHEM, 4.8 water solubility, OECD 105	Water solubility from Kow (mg/L) (EPI)	continuous				DanishQSARDatabase	
P-CHEM, 4.8 water solubility, OECD 105	Water solubility from Fragments (mg/L) (EPI)	continuous				DanishQSARDatabase	
Water solubility	ASNN	continuous		1311		OCHEM	 

3) PREDICTING

Once selected the model of interest, click on the link present in the “model” column; you will be redirected to the access page of the models.

3) PREDICTING

Once selected the model of interest, click on the link present in the “model” column; you will be redirected to the access page of the models.



4) INTERPRETATION OF THE RESULTS

The user can consult all the available documentation of the *in silico* tools in the dedicated section.



Environment International

Volume 131, October 2019, 105060



Review article

Integrating *in silico* models and read-across methods for predicting toxicity of chemicals: A step-wise strategy

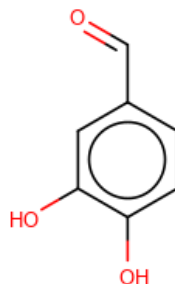
Emilio Benfenati ^a  , Qasim Chaudhry ^b, Giuseppina Gini ^c, Jean Lou Dorne ^d

 Show more

<https://doi.org/10.1016/j.envint.2019.105060>

[Cite this article](#)

Target molecule Tools Assessment



3,4-Dihydroxybenzaldehyde
O=Cc1ccc(O)c(O)c1
Genetic toxicity in vitro.
Ames test



- Checking the availability of the tools for the endpoint in the gateway
- Running the models






OUTPUTS
INTERPRETATION

Checking the availability of the tools for in vitro gene mutation in bacteria in the CONCERT REACH gateway

HOME	PROJECT	RESULTS	RESOURCES	NEWS	CONTACT
GATEWAY USER GUIDE					
GATEWAY					

8.4.1. In vitro gene mutation study in bacteria

All VEGA AND ToxRead DANISH QSAR DATABASE AMBIT OCHEM

End Point	Model	Type	Dataset size	Training set size	Test set size	Cross-validation procedure	Platform	Remarks
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) CONSENSUS model	classification	0	0	0		VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) model (CAESAR)	classification	4204	3367	837		VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) model (KNN-Read-Across)	classification	5770	5770	0		VEGA	

Currently **23 Models** from **VEGA**, **Danish (Q)SAR Database** and **OCHEM**

(Statistical&knowledge-based)

1 module of **ToxRead**

(read-across)

1 dataset of **AMBIT**

Direct link to the documentation:

QMRF, papers or guides

Checking the availability of the tools for in vitro gene mutation in bacteria in the CONCERT REACH gateway

8.4.1. In vitro gene mutation study in bacteria

ALL **VEGA AND ToxRead** DANISH QSAR DATABASE AMBIT OCHEM

End Point	Model	Type	Dataset size	Training set size	Test set size	Cross-validation procedure	Platform	Remarks
TOX 7.6.1. Genetic toxicity in vitro	Mutagenicity	reproducible read-across	6060				VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) CONSENSUS model	classification	0	0	0		VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) model (CAESAR)	classification	4204	3367	837		VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) model (KNN-Read-Across)	classification	5770	5770	0		VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) model (ISS)	classification	670	670	0		VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) model (SarPy-IRFMN) (version 1.0.8)	classification	4204	3367	837		VEGA	

VEGA

4 individual models + 1 consensus

- CAESAR - **Hybrid** model (statistical + knowledge-based)
- KNN-Read-Across - **read-across** model
- ISS - **knowledge-based** structural **alerts** (Benigni-Bossa rule-base)
- SarPy-IRFMN - **statistical** structural **alerts**

Checking the availability of the tools for in vitro gene mutation in bacteria in the CONCERT REACH gateway

8.4.1. In vitro gene mutation study in bacteria

ALL **VEGA AND ToxRead** DANISH QSAR DATABASE AMBIT OCHEM

End Point	Model	Type	Dataset size	Training set size	Test set size	Cross-validation procedure	Platform	Remarks
TOX 7.6.1. Genetic toxicity in vitro	Mutagenicity	reproducible read-across	6060				VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) CONSENSUS model	classification	0	0	0		VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) model (CAESAR)	classification	4204	3367	837		VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) model (KNN-Read-Across)	classification	5770	5770	0		VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) model (ISS)	classification	670	670	0		VEGA	
TOX 7.6.1. Genetic toxicity in vitro. Ames test (OECD 471)	Mutagenicity (Ames test) model (SarPy-IRFMN) (version 1.0.8)	classification	4204	3367	837		VEGA	

ToxRead - 1 module

Dataset = 6060 substances and their public data

4 different rulesets:

- ISS (**knowledge-based structural alerts**)
- SARpy (**statistical structural alerts**)
- CSR4 (**statistical structural alerts**)
- IRFMN (**knowledge-based structural alerts**)

Checking the availability of the tools for in vitro gene mutation in bacteria in the CONCERT REACH gateway

HOME

PROJECT

RESULTS

RESOURCES

NEWS

CONTACT

8.4.1. In vitro gene mutation study in bacteria

GATEWAY USER GUIDE

GATEWAY

ALL VEGA AND ToxRead DANISH QSAR DATABASE AMBIT OCHEM

End Point	Model	Type	Dataset size	Training set size	Test set size	Cross-validation procedure	Platform	Remarks
Ames test (OECD 471)	Bacterial reverse mutation test (Ames test in <i>S. typhimurium</i> in vitro) (CASE Ultra)	classification		4102		5 times 2-fold external crossvalidation	Danish QSAR Database	
Ames test (OECD 471)	Bacterial reverse mutation test (Ames test in <i>S. typhimurium</i> in vitro) (Leadscope)	classification		4102		5 times 2-fold external crossvalidation	Danish QSAR Database	
Ames test (OECD 471)	Bacterial reverse mutation test (Ames test in <i>S. typhimurium</i> in vitro) (SciQSAR)	classification		4102		5 times 2-fold external crossvalidation	Danish QSAR Database	
Ames test (OECD 471)	Direct acting Ames mutagens (without S9) - ONLY use for Ames POS_IN (CASE Ultra)	classification		388		5 times 2-fold external crossvalidation	Danish QSAR Database	
Ames test (OECD 471)	Direct acting Ames mutagens (without S9) - ONLY use for Ames POS_IN (Leadscope)	classification		388		5 times 2-fold external crossvalidation	Danish QSAR Database	
Ames test (OECD 471)	Direct acting Ames mutagens (without S9) - ONLY use for Ames POS_IN (SciQSAR)	classification		388		5 times 2-fold external crossvalidation	Danish QSAR Database	
Ames test	Base pair Ames mutagens -					5 times 2-fold		

Danish (Q)SAR Database

15 statistical models and 2 knowledge-based alert profilers

- Bacterial reverse mutation test (Ames test in *S. typhimurium* in vitro)
- Direct acting Ames mutagens (without S9)
- Base pair Ames mutagens
- Frame shift Ames mutagens
- Potent Ames mutagens, reversions ≥ 10 times controls

Profilers (OECD QSAR Toolbox V.4.2)

- DNA alerts for AMES by OASIS, alerts in parent only
- In vitro mutagenicity (Ames test) alerts by ISS, alerts in parent only

Checking the availability of the tools for in vitro gene mutation in bacteria in the CONCERT REACH gateway



HOME PROJECT RESULTS RESOURCES NEWS CONTACT

GATEWAY USER GUIDE

GATEWAY

8.4.1. In vitro gene mutation study in bacteria

ALL VEGA AND ToxRead DANISH QSAR DATABASE AMBIT **OCHEM**

End Point	Model	Type	Dataset size	Training set size	Test set size	Cross-validation procedure	Platform	Remarks
Ames test (OECD 471)	ASNN	Classification		4361	2181		OCHEM	 

OCHEM
1 statistical model & ToxAlert match

ALL VEGA AND ToxRead DANISH QSAR DATABASE **AMBIT** OCHEM

End Point	Model	Type	Dataset size	Training set size	Test set size	Cross-validation procedure	Platform	Remarks
TOX 7.6.1. Genetic toxicity in vitro	Dataset		50366				AMBIT	

AMBIT
1 dataset

Models for in vitro gene mutation in bacteria

How to select the appropriate model(s) for my substance?



A priori selection is generally **not possible**

- However, **experience in using the models** might suggest which could give more reliable results for certain types of substances (e.g., industrial chemicals, active substances, etc.)
- Information on **compliance** of the target molecule **with the applicability domain of the model**
- **Comparison with similar molecules** with available experimental results
- It is generally required to **use multiple and different models** for evaluating the same endpoint



The Gateway makes all those requirements available

Expert analysis of the results and supporting information is needed

01 Running **VEGA** models & **ToxRead** module and results analysis

02 Using **Danish (Q)SAR Database** and results analysis

03 Running **OCHEM** model & **ToxAlerts** and results analysis

04 Using **AMBIT** database and results analysis

TABLE OF CONTENTS

01

Welcome to the VEGA HUB

Offering a family of tools to evaluate chemical hazard: VEGA, ToxRead, SWAN, VERA, ToxWeight, ToxDelta, and JANUS.

VEGA is the QSAR software with tens of models for individual properties.

ToxRead

SWAN

VERA

VEGA

Vermeer

Do you need assistance for a property prediction ?

CONTACT US

VEGA HUB - QSAR - Download

Models and Tools

Assistance



We want to en

We want to enhance the r of the in silico methods b the properties of chem substances.

<https://www.vegahub.eu/download/>

VEGA HUB - QSAR - Download

VEGA HUB

Community News Contacts

VEGA QSAR

Try out VEGA QSAR and its features

All the VEGA models are also available in a unique stand-alone application.

With the VEGA application you can easily execute all the models on your local machine without sending any information to our server. VEGA is the ideal application for batch processing large datasets. VEGA can be installed and used on any operative system supporting JAVA technology (for any doubt please visit JAVA website).

- Introduction
- Screenshots
- Interpretation
- How to quote VEGA QSAR
- VEGA QSAR for KNIME
- Download
- Previous Versions

Free download VEGA QSAR Application

Visit the link to download the application.

DOWNLOAD

ToxRead

Privacy | Terms

VEGA in silico platform - version 1.2.3

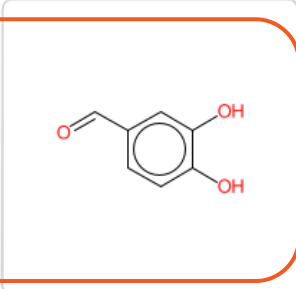
Insert chemicals

1. Add molecule(s) using SMILES notation

Insert SMILES:
O=Cc1ccc(O)c(O)c1 + Import File

ID	SMILES
Molecule 0	<chem>O=Cc1ccc(O)c(O)c1</chem>

2. Added molecules are listed and 2D structure can be visualized



DELETE All DELETE

VEGA in silico platform - version 1.2.3

Select models

Filter models: All available endpoints

Mutagenicity (Ames test)

- Select all models
- Mutagenicity (Ames test) model (CAESAR) - v. 2.1.14
- Mutagenicity (Ames test) model (ISS) - v. 1.0.3
- Mutagenicity (Ames test) model (SarPy-IRFMN) - v. 1.0.8
- Mutagenicity (Ames test) model (KNN-Read-Across) - v. 1.0.1
- Mutagenicity (Ames test) model for aromatic amines (CONCERT/IRFMN) - v. 1.0.0
- Mutagenicity (Ames test) CONSENSUS model - v. 1.0.4

Developmental toxicity

Select all models

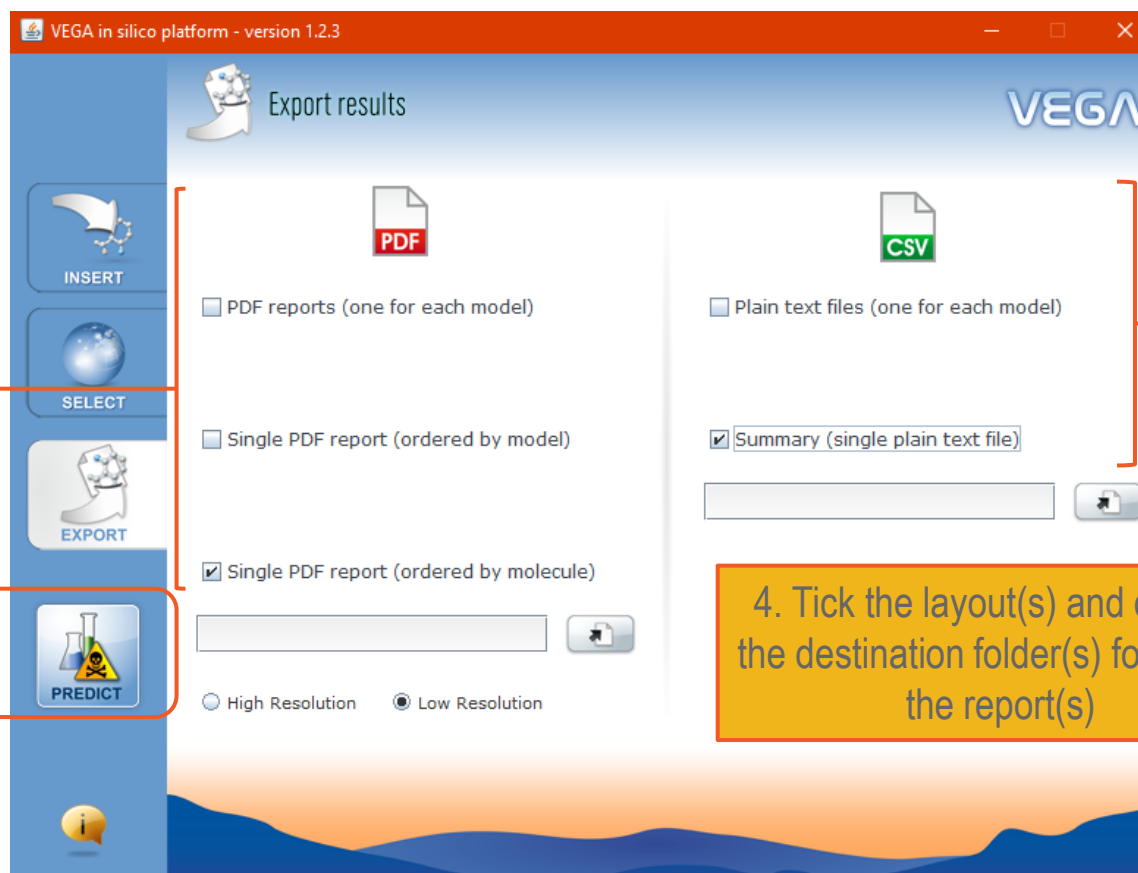
3. Select the model(s)

VEGA: running predictions

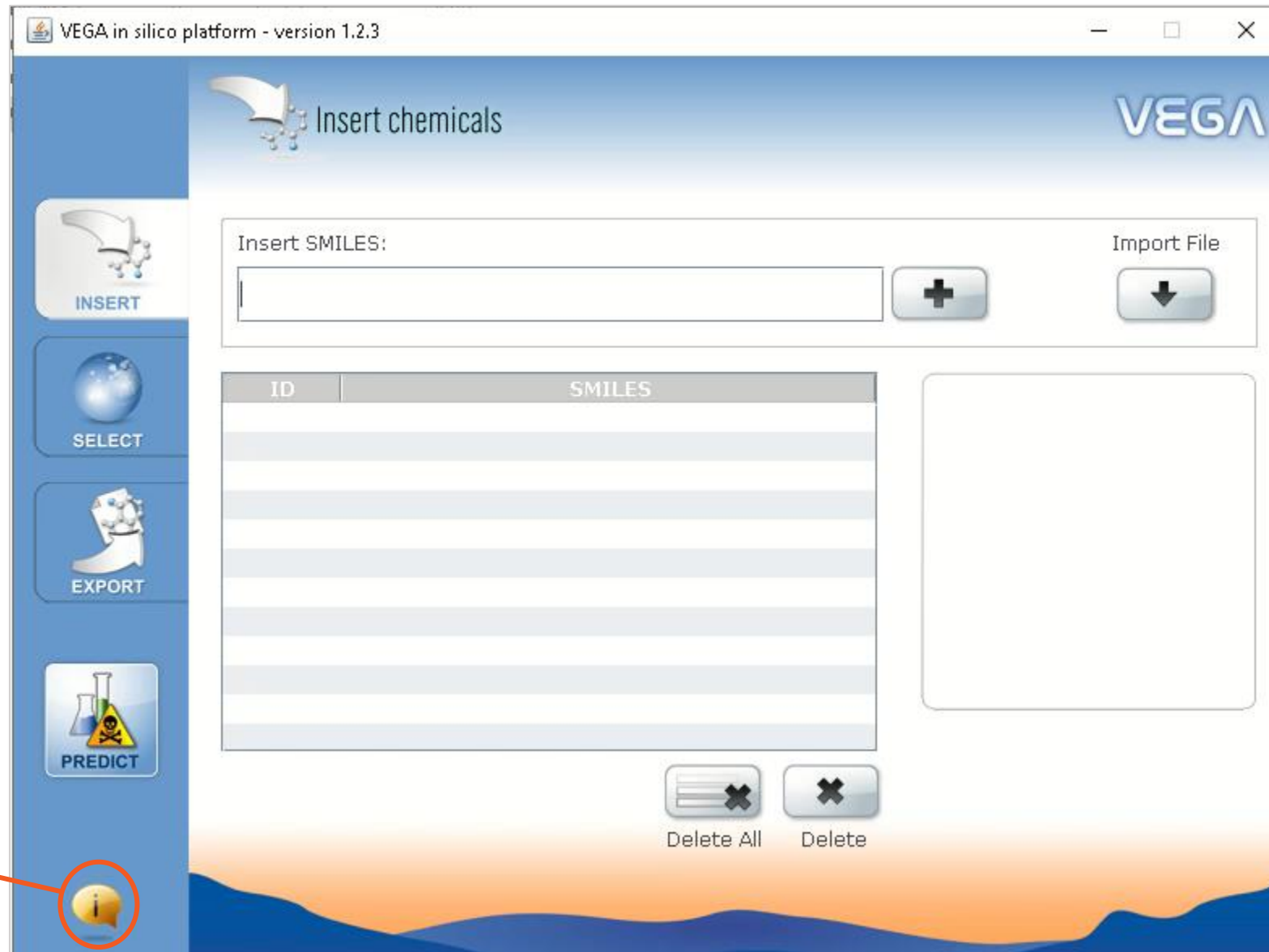
Full PDF reports:

- prediction(s) results
- applicability domain
- experimental data of the target (if any)
- most similar substances
- other supporting info (if any)

5. Click on «Predict»



Simplified text reports
(useful for excel import)



VEGA in silico platform - version 1.2.3

Insert chemicals

VEGA

Insert SMILES:

Import File

ID	SMILES

Delete All Delete


INSERT

SELECT

EXPORT

PREDICT

About VEGA

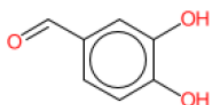
Version: 1.2.3 (build date: 13/03/2023)
Calculation core version: 1.3.18
The application is released under the GNU GPL-3 license
The user's guide is available (PDF document) 
Visit the project website: www.vega-qsar.eu


VEGA uses the following open source libraries:
Chemistry Development Kit (CDK) ver 2.3
OpenBabel ver 1.3
Vespa ver 3.5.0
PMML ver 1.5.6

MN MARIO NEGRI

kode chemoinformatics

VEGA: result analysis



Prediction: 

Prediction is NON-Mutagenic with a consensus score of 0.675, on 4 models.

Compound: Molecule 0

Compound SMILES: O=Cc1ccc(O)c(O)c1

Used models: 4

Predicted Consensus Mutagen activity: NON-Mutagenic

Mutagenic Score: 0.05

Non-Mutagenic Score: 0.675

Model Caesar assessment: NON-Mutagenic (GOOD reliability)

Model ISS assessment: Mutagenic (LOW reliability)

Model SarPy assessment: Possible NON-Mutagenic (GOOD reliability)


Model KNN assessment: NON-Mutagenic (GOOD reliability)

Remarks:

none

VEGA Mutagenicity (Ames test) model (CAESAR) 2.1.14 page 2

1. Prediction Summary

Prediction: 


VEGA Mutagenicity (Ames test) model (SarPy-IRFMN) 1.0.8 page 10

1. Prediction Summary

Prediction: 

VEGA Mutagenicity (Ames test) model (KNN-Read-Across) 1.0.1 page 14

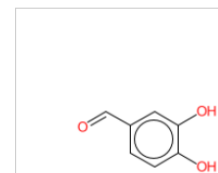
1. Prediction Summary



Prediction: 

VEGA Mutagenicity (Ames test) model (ISS) 1.0.3 page 5

1. Prediction Summary

Prediction for compound Molecule 0 -



Prediction:  Reliability: 

Prediction is Mutagenic, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:

- similar molecules found in the training set have experimental values that disagree with the predicted value
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)

The following alerts have been found: SA11 Simple aldehyde

Compound: Molecule 0
Compound SMILES: O=Cc1ccc(O)c(O)c1
Experimental value: -

Predicted Mutagen activity: Mutagenic

Structural Alerts: SA11 Simple aldehyde

Reliability: The predicted compound is outside the Applicability Domain of the model

Remarks:

none

Stars for reliability from 1 to 3 (high) primary of the prediction

The reliability of the prediction is based on an automated check of the molecule compliance with the applicability domain of the model.

VEGA: result analysis

Applicability Domain Index (ADI) ranges from 0 (not in AD) to 1 (in AD)

The ADI is calculated based on other indices, **each one taking into account a particular issue** of the applicability domain (AD)

3.2 Applicability Domain: Measured Applicability Domain Scores

☆☆☆

- ✓ Global AD Index
AD index = 0.965
Explanation: The predicted compound is into the Applicability Domain of the model.
- ✓ Similar molecules with known experimental value
Similarity index = 0.932
Explanation: Strongly similar compounds with known experimental value in the training set have been ..
- ✓ Accuracy of prediction for similar molecules
Accuracy index = 1
Explanation: Accuracy of prediction for similar molecules found in the training set is good..
- ✓ Concordance for similar molecules
Concordance index = 1
Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
- ✓ Atom Centered Fragments similarity check
ACF index = 1
Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

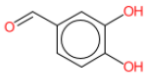
Symbols explanation:

- ✓ The feature has a good assessment, model is reliable regarding this aspect.
- ⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.
- ✗ The feature has a bad assessment, model is not reliable regarding this aspect.

VEGA Mutagenicity (Ames test) model (KNN-Read-Across) 1.0.1 page 14

1. Prediction Summary

Prediction for compound Molecule 0 -



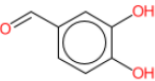
Prediction: ● Reliability: ☆☆☆

Prediction is NON-Mutagenic, the result appears reliable. Anyhow, you should check it through the evaluation of the information given in the following sections.

Compound: Molecule 0
Compound SMILES: O=Cc1ccc(O)c(O)c1
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
Molecules used for prediction: 4
Reliability: The predicted compound is into the Applicability Domain of the model
Remarks:
none

Number of considered similar molecules and number and type of indexes are **model dependent**

VEGA: example of the automated AD/reliability evaluation



Prediction: ● Reliability: ★ ★ ★

Prediction is NON-Mutagenic, the result appears reliable. Anyhow, you should check it through the evaluation of the information given in the following sections.

Compound: Molecule 0
Compound SMILES: O=Cc1ccc(O)c(O)c1
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
Molecules used for prediction: 4
Reliability: The predicted compound is into the Applicability Domain of the model
Remarks: none

- ✔ **Global AD Index**
AD index = 0.965
Explanation: The predicted compound is into the Applicability Domain of the model.
- ✔ **Similar molecules with known experimental value**
Similarity index = 0.932
Explanation: Strongly similar compounds with known experimental value in the training set have been ..
- ✔ **Accuracy of prediction for similar molecules**
Accuracy index = 1
Explanation: Accuracy of prediction for similar molecules found in the training set is good..
- ✔ **Concordance for similar molecules**
Concordance index = 1
Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
- ✔ **Atom Centered Fragments similarity check**
ACF index = 1
Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

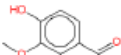
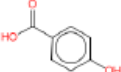
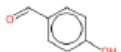
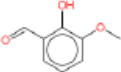
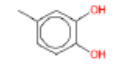
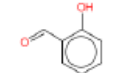
All the SCs are TN

All EXPs=prediction

3.1 Applicability Domain:

Similar Compounds, with Predicted and Experimental Values



	<p>Compound #1 CAS: 121-33-5 Dataset id:873 (Training Set) SMILES: <chem>O=Cc1ccc(O)c(OC)c1</chem> Similarity: 0.938 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic</p>
	<p>Compound #2 CAS: 99-96-7 Dataset id:5596 (Training Set) SMILES: <chem>O=C(O)c1ccc(O)cc1</chem> Similarity: 0.938 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic</p>
	<p>Compound #3 CAS: 123-08-0 Dataset id:932 (Training Set) SMILES: <chem>O=Cc1ccc(O)cc1</chem> Similarity: 0.927 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic</p>
	<p>Compound #4 CAS: 148-53-8 Dataset id:1548 (Training Set) SMILES: <chem>O=Cc1ccc(OC)c1(O)</chem> Similarity: 0.926 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic</p>
	<p>Compound #5 CAS: 452-86-8 Dataset id:3123 (Training Set) SMILES: <chem>Oc1ccc(cc1(O))C=O</chem> Similarity: 0.909 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic</p>
	<p>Compound #6 CAS: 90-02-8 Dataset id:5134 (Training Set) SMILES: <chem>O=Cc1cccc(O)c1</chem> Similarity: 0.909 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic</p>

3.2 Applicability Domain: Measured Applicability Domain Scores



Global AD Index
AD index = 0
Explanation: The predicted compound is outside the Applicability Domain of the model.



Similar molecules with known experimental value
Similarity index = 0.861
Explanation: Strongly similar compounds with known experimental value in the training set have been ..



Accuracy of prediction for similar molecules
Accuracy index = 1
Explanation: Accuracy of prediction for similar molecules found in the training set is good..



Concordance for similar molecules
Concordance index = 0
Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..

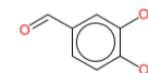






Atom Centered Fragments similarity check
ACF index = 0.85
Explanation: some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)..

1. Prediction Summary



Prediction for compound Molecule 0 -



Prediction:  Reliability:   

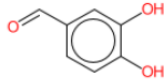
Prediction is Mutagenic, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:

- similar molecules found in the training set have experimental values that disagree with the predicted value
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)

The following alerts have been found: SA11 Simple aldehyde

Compound: Molecule 0
Compound SMILES: O=Cc1ccc(O)c(O)c1
Experimental value: -
Predicted Mutagen activity: Mutagenic
Structural Alerts: SA11 Simple aldehyde
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

VEGA: example of the automated AD/reliability evaluation



Prediction: ● Reliability: ★ ★ ★

Prediction is Mutagenic, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:

- similar molecules found in the training set have experimental values that disagree with the predicted value
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)

The following alerts have been found: SA11 Simple aldehyde

Compound: Molecule 0
Compound SMILES: O=Cc1ccc(O)c(O)c1
Experimental value: -
Predicted Mutagen activity: Mutagenic
Structural Alerts: SA11 Simple aldehyde
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks: none

3.2 Applicability Domain: Measured Applicability Domain Scores







✘	Global AD Index AD index = 0 Explanation: The predicted compound is outside the Applicability Domain of the model.
✔	Similar molecules with known experimental value Similarity index = 0.861 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
✔	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✘	Concordance for similar molecules Concordance index = 0 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
⚠	Atom Centered Fragments similarity check ACF index = 0.85 Explanation: some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)..

All the SCs are TN

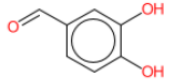
All EXPs≠prediction



VEGAMutagenicity (Ames test) model (ISS) 1.0.3page 6

3.1 Applicability Domain: Similar Compounds, with Predicted and Experimental Values

	Compound #1 CAS: 120-80-9 Dataset id:817 (Training Set) SMILES: <chem>Oc1ccccc1(O)</chem> Similarity: 0.866 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic
	Compound #2 CAS: 65-85-0 Dataset id:798 (Training Set) SMILES: <chem>O=C(O)c1ccccc1</chem> Similarity: 0.856 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic
	Compound #3 CAS: 123-31-9 Dataset id:673 (Training Set) SMILES: <chem>Oc1ccc(O)cc1</chem> Similarity: 0.852 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic
	Compound #4 CAS: 118-92-3 Dataset id:96 (Training Set) SMILES: <chem>O=C(O)c1ccccc1(N)</chem> Similarity: 0.849 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic
	Compound #5 CAS: 108-46-3 Dataset id:298 (Training Set) SMILES: <chem>Oc1ccccc1(O)c1</chem> Similarity: 0.849 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic
	Compound #6 CAS: 331-39-5 Dataset id:813 (Training Set) SMILES: <chem>O=C(O)C=Cc1ccc(O)c(O)c1</chem> Similarity: 0.835 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic

VEGA: example of the automated AD/reliability evaluation



Prediction:  Reliability: 


Prediction is Mutagenic, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:






- similar molecules found in the training set have experimental values that disagree with the predicted value
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)

The following alerts have been found: SA11 Simple aldehyde

Compound: Molecule 0
Compound SMILES: O=Cc1ccc(O)c(O)c1
Experimental value: -
Predicted Mutagen activity: Mutagenic
Structural Alerts: SA11 Simple aldehyde
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none


3.2 Applicability Domain: Measured Applicability Domain Scores



	<p>Global AD Index AD index = 0 Explanation: The predicted compound is outside the Applicability Domain of the model.</p>
	<p>Similar molecules with known experimental value Similarity index = 0.861 Explanation: Strongly similar compounds with known experimental value in the training set have been ..</p>
	<p>Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..</p>
	<p>Concordance for similar molecules Concordance index = 0 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..</p>
	<p>Atom Centered Fragments similarity check ACF index = 0.85 Explanation: some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)..</p>

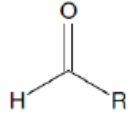
VEGA Mutagenicity (Ames test) model (ISS) 1.0.3 page 8

4.1 Reasoning: Relevant Chemical Fragments and Moieties



(Molecule 0) Reasoning on fragments/structural alerts :

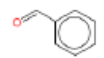


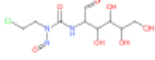

Fragment found: SA11 Simple aldehyde



R= aliphatic or aromatic carbon
 α, β unsaturated aldehydes are excluded

Aliphatic and aromatic aldehydes. The alfa,beta-unsaturated aldehydes are excluded

Following, the most similar compounds from the model's dataset having the same fragment.

	<p>CAS: 100-52-7 Dataset id:661 (Training Set) SMILES: <chem>O=Cc1ccccc1</chem> Similarity: 0.815</p> <p>Experimental value : NON-Mutagenic  Predicted value : Mutagenic </p> <p>Alerts (found also in the target): SA11 Simple aldehyde</p>
	<p>CAS: 54749-90-5 Dataset id:826 (Training Set) SMILES: <chem>O=NN(C(=O)NC(C=O)C(O)C(O)C(O)CO)CCCI</chem> Similarity: 0.48</p> <p>Experimental value : Mutagenic Predicted value : Mutagenic</p> <p>Alerts (found also in the target): SA11 Simple aldehyde Alerts (not found also in the target): SA8 Aliphatic halogens; SA21 Alkyl and aryl N-nitroso groups</p>
	<p>CAS: 107-20-0 Dataset id:820 (Training Set) SMILES: <chem>O=CCCl</chem> Similarity: 0.469</p> <p>Experimental value : Mutagenic Predicted value : Mutagenic</p> <p>Alerts (found also in the target): SA11 Simple aldehyde Alerts (not found also in the target): SA8 Aliphatic halogens</p>

• Lower similarity
• Other SAs not present in the target

01 Running **VEGA** models & **ToxRead**
module and results analysis

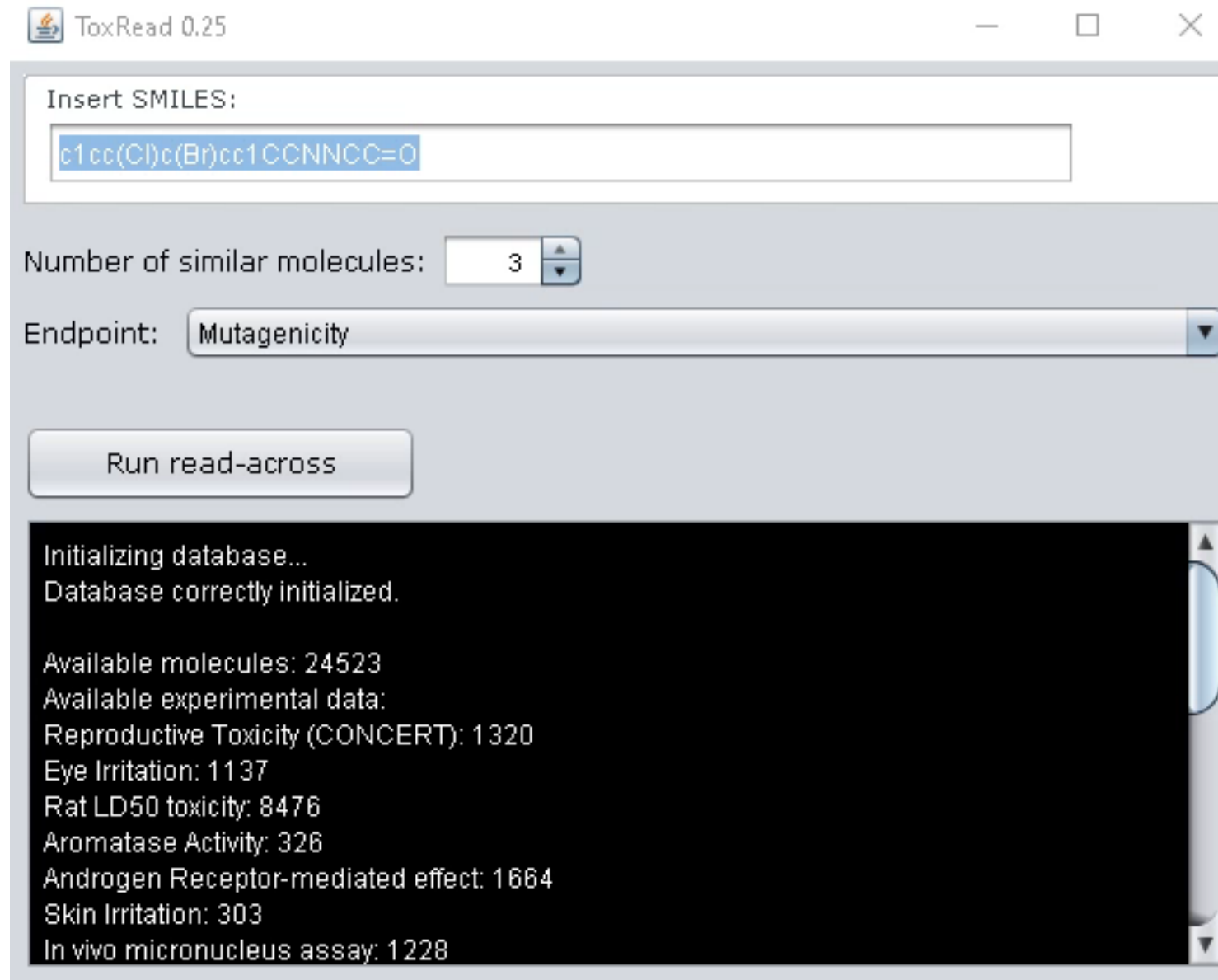
02 Using Danish (Q)SAR Database
and results analysis

03 Running OCHEM model &
ToxAlerts and results analysis

04 Using AMBIT database and results
analysis

TABLE OF CONTENTS

TOXREAD: running read-across



Insert SMILES:

```
c1cc(Cl)c(Br)cc1CCNNCC=O
```

Number of similar molecules: 3

Endpoint: Mutagenicity

Run read-across

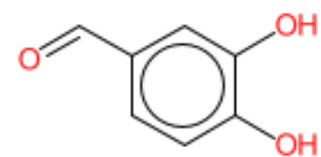
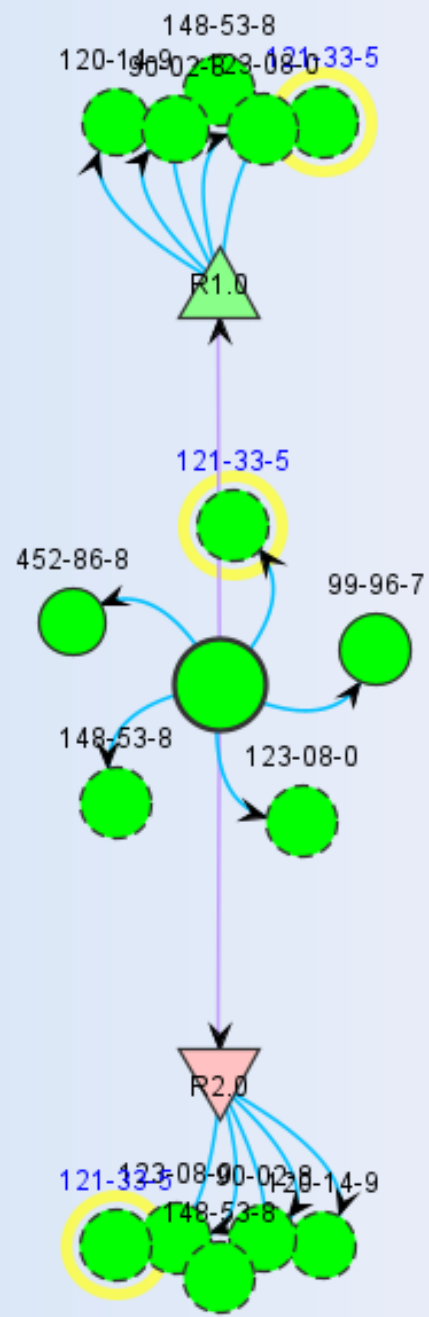
Initializing database...
Database correctly initialized.

Available molecules: 24523
Available experimental data:
Reproductive Toxicity (CONCERT): 1320
Eye Irritation: 1137
Rat LD50 toxicity: 8476
Aromatase Activity: 326
Androgen Receptor-mediated effect: 1664
Skin Irritation: 303
In vivo micronucleus assay: 1228

Dataset = 6060 substances
and their public data







4 different rulesets:

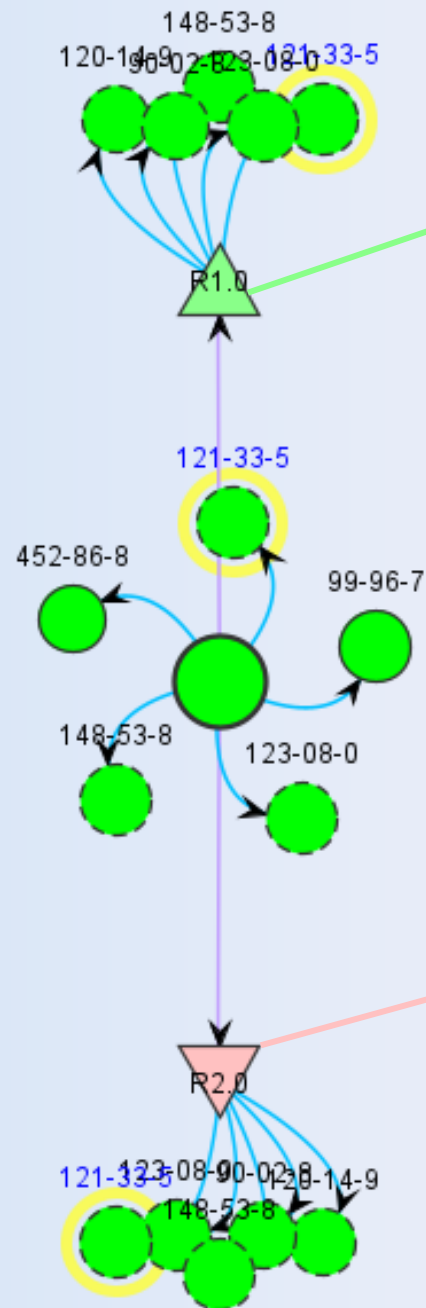
- ISS (46, 100% **P**)
- SARpy (205, 55% **P**)
- CSR4 (238, 41% **P**)
- IRFMN (282, 57% **P**)



Molecular Weight: 138.14
LogP (experimental): 1.09

Label	Rule Set
R2.0	Aliphatic and aromatic aldehydes. The alfa,beta-unsaturat...
R1.0	IRFMN alert n. 106 for NON-Mutagenicity, defined by the S...

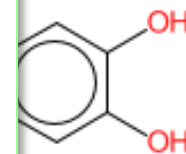
-  target compound
-  intensity of red related to the value of mutagenicity
-  intensity of green related to the value of non-mutagenicity
-  non-toxicity or toxicity rule, size related to the Fisher Test P-Value of the rule
-  circle size related to the similarity with the target compound
-  repeated compound



Details of rule MNM106

Details of rule MNM106

Name: MNM106
 Description: IRFMN alert n. 106 for NON-Mutagenicity
 Experimental accuracy: 0.71
 Fisher test p-value: 0.00319

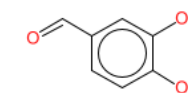
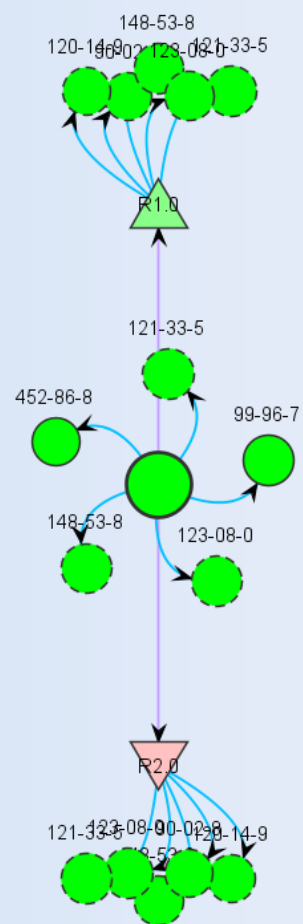


Details of rule SA11 Simple aldehyde

R= aliphatic or aromatic carbon
 α, β unsaturated aldehydes are excluded

Details of rule SA11 Simple aldehyde

Name: SA11 Simple aldehyde
 Description: Aliphatic and aromatic aldehydes. The alfa,beta-unsaturated aldehydes are excluded.
 Experimental accuracy: 0.43
 Experimental accuracy (only on biocide compounds): 1
 Fisher test p-value: 0.02957



Target molecule prediction

Overall assessment: **NON-MUTAGENIC**

Read-Across assessment: Non-Mutagenic
Read-Across Mutagenic score = 0
Read-Across Non-Mutagenic score = 1

QSAR consensus assessment: NON-Mutagenic (Consensus score: 0.675)
Predicted Consensus Mutagen activity = NON-Mutagenic
Mutagenic Score = 0.05
Non-Mutagenic Score = 0.675

Model Caesar assessment = NON-Mutagenic (GOOD reliability)
Model ISS assessment = Mutagenic (LOW reliability)
Model SarPy assessment = Possible NON-Mutagenic (GOOD reliability)
Model KNN assessment = NON-Mutagenic (GOOD reliability)

a-unsaturated aldehydes are excluded

defined by the SMARTS: [C;D2;H1](=O)a1aaaaa1



01 Running VEGA models & ToxRead module and results analysis

02 Using **Danish (Q)SAR Database** and results analysis

03 Running OCHEM model & ToxAlerts and results analysis

04 Using AMBIT database and results analysis

TABLE OF CONTENTS

Danish (Q)SAR Database: gathering (Q)SAR results

Danish (Q)SAR Database

Home Clear Information Contact QSAR2023

New search

Searches

Results

Substances

1. Input by structure

ID

Structure and name

PhysChem

Environment

ADME

Human health

AND
Intersect results

OR
Unite results

NOT
Complement results

MAX/MIN
More combinations

2. Import

3. Paste SMILES

4. Confirm SMILES

Single structure SMILES list

Substructure Similarity

Name search

O=Cc1ccc(O)c(O)c1

OK Cancel

5. Search the database for the target molecule

Substructure search

Exact match search

Cancel

Danish (Q)SAR Database: gathering (Q)SAR results

Danish (Q)SAR Database

4. Summary of the search performed

Home Clear Information Contact QSAR2023

New search

ID
Structure and name

PhysChem
Environment
ADME
Human health

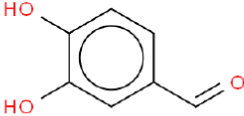
AND Intersect results
OR Unite results
NOT Complement results
MAX/MIN More combinations

Searches 1. Exact match: Results 1

Danish (Q)SAR Database, <http://qsar.food.dtu.dk> Date: 24-04-2023

(Q)SAR predicted profile

Structure (as used for QSAR prediction):



SMILES (used for QSAR prediction): Oc1ccc(O)cc1C=O

ID			
Registry Number	139-85-6	PubChem CID	
REACH EC Number (pre-registration, by 2013)	205-377-7	REACH EC Number (registration, 2019 or 2022)	
REACH registration (2022)		REACH registration cumulated minimum annual tonnage (2022)	
EU CLP Harmonized Classification*	DK-EPA / DTU QSAR-based CLP Advisory Classification	Acute Tox. 4; Skin Irrit. 2	
EU Biocide active substances		EU Pesticide active substances	
EU EFSA Botanical substances	Yes	US TSCA (Oct. 2021)	
Tox21 (2019)		ToxCast (Oct. 2021)	
Molecular Formula	C7 H6 O3	Molecular weight (g/mole)	138.12
Chemical Name	3,4-dihydroxybenzaldehyde		

(Annex VI to CLP up to and including the 9th ATP, and including Nordic Council of Minister SPIN list for group entries)

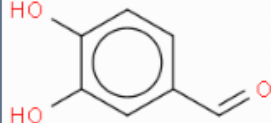

Melting point, Boiling point and Vapour pressure

1

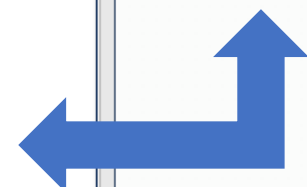
Substances Exact match: Page 1

Previous Next 1

Structures 1-1 of 1

Structure	Id	Similarity	+
	139-XX-X		

5. For each identified molecule, the (Q)SAR report can be downloaded in .rtf format



Danish (Q)SAR Database: results for in vitro gene mutation in bacteria

In vitro Genotoxicity - Bacterial Reverse Mutation Test (Ames test)

	Exp	Battery	CASE Ultra	Leadscope	SciQSAR
Ames test in <i>S. typhimurium</i> (<i>in vitro</i>)		NEG_IN	NEG_IN	NEG_IN	NEG_OUT
*Direct Acting Mutagens (without S9)	N/A	INC_OUT	NEG_OUT	INC_OUT	INC_OUT
*Base-Pair Ames Mutagens	N/A	NEG_OUT	NEG_OUT	NEG_IN	INC_OUT
*Frameshift Ames Mutagens	N/A	INC_OUT	POS_OUT	POS_IN	NEG_IN
*Potent Ames Mutagens, Reversions \geq 10 Times Controls	N/A	POS_IN	POS_IN	POS_IN	POS_IN

DTU-developed models

* The four models (Direct Acting mutagens (without S9), Base-Pair Ames Mutagens, Frameshift Ames Mutagens, Potent Ames Mutagens) should not be used to determine if substances are Ames mutagens, but can be used for indication of mechanism or potency for cases where the main Ames model (Ames test in *S. typhimurium* (*in vitro*)) is POS_IN.

The target molecule was evaluated as **compliant with AD** except for SciQSAR model.
The other four models should not be considered.

Within LIFE CONCERT REACH, results from the four VEGA models and the Consensus model have been integrated

VEGA	Mut. / Non-mut. scores	Used models
Mutagenicity consensus	0.05 / 0.68	4

Mutagenicity (Ames) consensus model version 1.0.2 contained in VEGA version 1.1.4 with calculation core version 1.2.4

Prediction: POS = Mutagenic, NEG = Non-mutagenic.

VEGA	CAESAR	SarPy	KNN
ISS	NEG_Good	POSS.NEG_Good	NEG_Good

Four individual models in mutagenicity consensus model version 1.0.2 contained in VEGA version 1.1.4 with calculation core version 1.2.4

Structural alerts identified by two endpoint-specific profilers present in the OECD QSAR Toolbox

DNA alerts for AMES by OASIS, alerts in:

- parent only No alert found

In vitro mutagenicity (Ames test) alerts by ISS, alerts in:

- parent only Simple aldehyde

OECD QSAR Toolbox v.4.2 profilers

Profiler predictions are supporting information to be used together with the relevant QSAR predictions

Danish (Q)SAR Database

Home Clear Information Contact QSAR2023

Danish (Q)

New search

ID

Structure and name

PhysChem

Environment

ADME

Human health

AND

Intersect results

OR

Unite results

NOT

Complement results

MAX/MIN

More combinations

Searches



1. Exact match:

Results

1

Substances

Exact match: Page 1

Edit

Single structure

SMILES list

Substructure

Similarity

Name search!



C

N

O

F

P

S

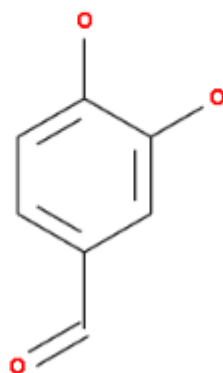
H

..

/

//

✎



The full database will be ordered
by similarity to the query chemical.

Display:

All structures

User-defined number of closest analogs:

10

Similarity

Cancel

139-XX-X



Danish (Q)SAR Database: identification of similar molecules

Danish (Q)SAR Database

Home Clear Information Contact QSAR2023

Danish (Q)SAR Models

New search

ID

Structure and name

PhysChem

Environment

ADME

Human health

AND
Intersect results

OR
Unite results

NOT
Complement results

MAX/MIN
More combinations

Select predictions

CAR inhibition at max. 50 uM (in vitro)

Developmental Toxicity

Teratogenic Potential in Humans

Genotoxicity

DNA Reactivity

Ashby Structural Alerts for DNA Reactivity

Profilers

DNA binding by OASIS, OECD QSAR Toolbox v.4.2

DNA binding by OECD, OECD QSAR Toolbox v.4.2

Ames test

Bacterial Reverse Mutation Test (Ames test in *S. typhimurium* (in vitro))

Direct Acting Ames Mutagens (without S9) - ONLY use for Ames POS

Base-Pair Ames Mutagens - ONLY use for Ames POS_IN

Frameshift Ames Mutagens - ONLY use for Ames POS_IN

Battery

CASE Ultra

Leadscope

SciQSAR

Experimental

Add

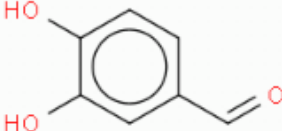
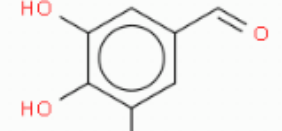
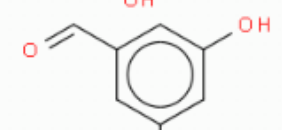
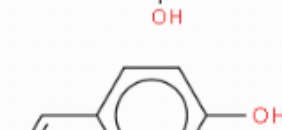
Cancel

Substances

Similarity, 10 closest analogs: Page 1

Previous Next 1

Structures 1-10 of 10

Structure	Id	Similarity	Bacterial Reve... Experimental	+
	139-XX-X	1.0		
	13677-XX-X	0.975		
	26153-XX-X	0.95		
	92203-XX-X	0.925		

By clicking on + it is possible to add info about exp/pred

Danish (Q)SAR Database: identification of similar molecules

Danish (Q)SAR Database

Home Clear Information Contact QSAR2023

Danish (Q)SAR Models

New search

ID
Structure and name

PhysChem
Environment
ADME
Human health

AND
Intersect results
OR
Unite results
NOT
Complement results
MAX/MIN
More combinations

Searches

1. Exact match: 1
2. Similarity, 10 closest analogs: 10

Results

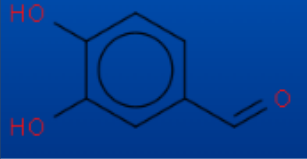
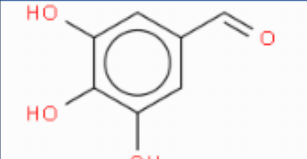
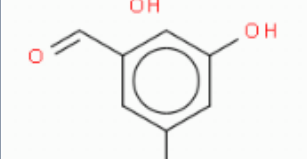
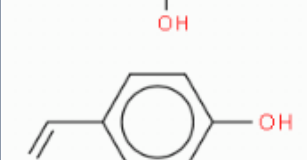
1
10

Substances

Similarity, 10 closest analogs: Page 1

Previous Next 1

Structures 1-10 of 10

Structure	Id	Similarity	Bacterial Reve... Experimental	Bacterial Reve... Battery	+
	139-XX-X ↓	1.0		NEG_IN	
	13677-XX-X ↓	0.975		NEG_IN	
	26153-XX-X ↓	0.95		NEG_IN	
	92203-XX-X ↓	0.925		NEG_IN	

The generated overview **can be used** to prepare a statement **for supporting the reliability** of the (Q)SAR prediction for the target molecule.

Danish (Q)SAR Database: identification of similar molecules

- **Stepwise approach**
- Danish (Q)SAR Database **can be searched** for molecules, based on available **experimental data, (Q)SAR predictions, structural alerts**, etc., for the endpoint of interest
- For each query, a **list of molecules is retrieved**
- The lists **can be merged**, using logical operators such as **AND** or **OR**

Danish (Q)SAR Database: identification of similar molecules

Home Clear Information Contact QSAR2023

New search

ID

Structure and name

PhysChem

Environment

ADME

Human health

AND
Intersect results

OR
Unite results

NOT
Complement results

MAX/MIN
More combinations

Searches



Results

Searches	Results
1. Exact match:	1
2. POS Bacterial Reverse Mutation Test (Ames test ...)	2361
3. NEG Bacterial Reverse Mutation Test (Ames test ...)	3971
4. NO alert in P: DNA alerts for AMES by OASIS, OECD QSAR Too...	594272
5. POS alert in P: In vitro mutagenicity (Ames test) alerts by...	185786
6. 1. OR 2. OR 3.	6333
7. 4. AND 5. AND 6.	1518

- 1: The target molecule (3,4-Dihydroxybenzaldehyde)- SMILES
- 2: All molecules from the database, **experimentally positive** for Ames test
- 3: All molecules from the database, **experimentally negative** for Ames test
- 4 to 5: All molecules with no alerts for DNA alerts for AMES by OASIS profiler and at least one alert for In vitro mutagenicity (Ames test) alerts by ISS.
- 6: Target + experimentally positive + experimentally negative (1, 2 and 3 combined with OR)
- 7: Subset of 6, including molecules with info about alerts for the 2 relevant profilers (6, 4 and 5 combined with AND)

The query from 2 to 5 is performed in the "Human Health" section

Danish (Q)SAR Database: identification of similar molecules

New search

Searches < > 1. Exact match: > Results 1

Substances Exact match: Page 1
Previous Next 1
Structures 1-1 of 1

Structure Id Similarity

Oc1ccc(O)cc1C=O 139-XX-X

Bacterial Reverse Mutation Test (Ames test in *S. typhimurium* (in vitro))

Search Model documentation

Bacterial Reverse Mutation Test (Ames test in *S. typhimurium* (in vitro))

Select predictions or experimental results:

- Battery (combines all three systems)
- CASE Ultra
- Leadscope
- SciQSAR
- Experimental (from training set)**

and search for structures experimentally tested:

Positive

Negative

Cancel

OR
Unite results

NOT
Complement results

MAX/MIN
More combinations

DNA Reactivity

Ames test

Bacterial Reverse Mutation Test (Ames test in *S. typhimurium* (in vitro))

Direct Acting Ames Mutagens (without S9) - ONLY use for Ames POS_IN

Base-Pair Ames Mutagens - ONLY use for Ames POS_IN

Frameshift Ames Mutagens - ONLY use for Ames POS_IN

Potent Ames Mutagens (Reversions ≥ 10 Times Controls) - ONLY use for Ames POS_IN

Profilers

Danish (Q)SAR Database: identification of similar molecules

New search

Searches



Results

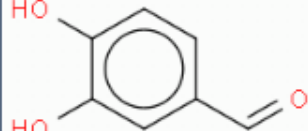
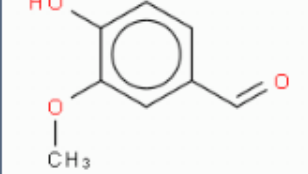
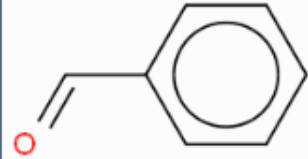
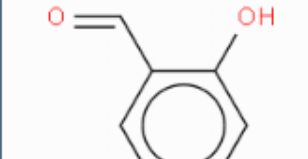
Search	Results
1. Exact match:	1
2. POS Bacterial Reverse Mutation Test (Ames test ...	2361
3. NEG Bacterial Reverse Mutation Test (Ames test ...	3971
4. NO alert in P: DNA alerts for AMES by OASIS, OECD QSAR Too...	594272
5. POS alert in P: In vitro mutagenicity (Ames test) alerts by...	185786
6. 1. OR 2. OR 3.	6333
7. 4. AND 5. AND 6.	1518

Substances

4. AND 5. AND 6.: Page 1

Previous Next 1 2 3 152

Structures 1-10 of 1518

Structure	Id	Similarity	Bacterial Reve... Experimental
	XXX-XX-X ↓	1.0	
	XXX-XX-X ↓	0.886	NEG
	XXX-XX-X ↓	0.846	NEG
	XX-XX-X ↓	0.829	NEG

ID

Structure and name

PhysChem

Environment

ADME

Human health

AND
Intersect results

OR
Unite results

NOT
Complement results

MAX/MIN
More combinations

Danish (Q)SAR Database: running predictions

Danish (Q)SAR Models

powered by Leadscope Predictive Data Miner



[Home](#) [New query](#) [Quick start guide](#) [Model documentation](#) [Contact](#)

Danish (Q)SAR Database

Molecule Id (optional):

Select models

Select all

Environmental

- Acute aquatic toxicity
 - Fish (Fathead minnow 96h LC50 (mg/L))
 - Daphnid (Daphnia magna 48h EC50 (mg/L))
 - Green algae (Pseudokirchneriella s. 72h EC50 (mg/L))
- Biodegradation
 - Not ready biodegradability (POS=Not Ready)

ADME

Endocrine/molecular

Endocrine/molecular 2

Genotoxicity/cancer

Other endpoints

Input structure

C
N
O
F
P
S
H
.
/
//
///
/

Predict

Select models and input a chemical structure.

(in vitro) NEW			(2*5-fold cross-validation)
Pregnane X Receptor (PXR) Activation (Human in vitro)	2176	Leadscope	Sens=89.1, Spec=98.6, BA=93.9 (2*5-fold cross-validation)
Pregnane X Receptor (PXR) Activation (Rat in vitro)	2330	Leadscope	Sens=86.5, Spec=97.4, BA=92.0 (2*5-fold cross-validation)
CYP3A4 Induction (Human in vitro)	2271	Leadscope	Sens=86.7, Spec=98.2, BA=92.5 (2*5-fold cross-validation)
Constitutive Androstane Receptor (CAR) activation at max. 20 μ M	924	Leadscope	Sens=72.2, Spec=93.5, BA=82.8 (2*5-fold cross-validation)
Constitutive Androstane Receptor (CAR) activation at max. 50 μ M	1903	Leadscope	Sens=78.4, Spec=91.4, BA=84.9 (2*5-fold cross-validation)
Constitutive Androstane Receptor (CAR) inhibition at max. 20 μ M	1408	Leadscope	Sens=58.4, Spec=97.1, BA=77.8 (2*5-fold cross-validation)
Constitutive Androstane Receptor (CAR) inhibition at max. 50 μ M	1870	Leadscope	Sens=72.4, Spec=91.6, BA=82.0 (2*5-fold cross-validation)
Bacterial reverse mutation test (Ames test in <i>S. typhimurium</i> in vitro)	4102	Leadscope	Sens=84.3, Spec=85.7, Conc=84.9
Chromosome aberrations in CHL cells (<i>in vitro</i>)	688	Leadscope	Sens=74.6, Spec=75.2, Conc=74.9
Mutations in thymidine kinase locus in mouse lymphoma cells (<i>in vitro</i>)	555	Leadscope	Sens=85.1, Spec=83.8, Conc=84.4
Mutations in HGPRT locus in CHO cells (<i>in vitro</i>)	239	Leadscope	Sens=81.7, Spec=78.4, Conc=80.5
Unscheduled DNA synthesis (UDS) in rat hepatocytes (<i>in vitro</i>)	415	Leadscope	Sens=74.1, Spec=70.1, Conc=72.4
Syrian hamster embryo (SHE) cell transformation (<i>in vitro</i>)	363	Leadscope	Sens=71.6, Spec=76.5, Conc=74.5
Sex-linked recessive lethal (SLRL) test in <i>Drosophila m.</i> (<i>in vivo</i>)	367	Leadscope	Sens=79.1, Spec=80.3, Conc=79.6

4. Defining the algorithm

4.1 Type of model

A categorical (Q)SAR model based on structural features and numeric molecular descriptors.

4.2 Explicit algorithm

This is a categorical (Q)SAR model made by use of partial logistic regression (PLR). The model is a composite model consisting of 2 submodels using all the negatives (1803 chemicals) in each of these and different sub-sets of the positives (see 4.5). The specific implementation is proprietary within the Leadscope software.

Danish (Q)SAR Database: running QSAR model

Danish (Q)SAR Models

powered by Leadscope Predictive Data Miner

Ministry of Health and Food

Home New query Quick start guide Model documentation Contact

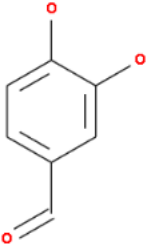
Molecule Id (optional):

1 model selected

1. Select the model from the list of endpoints

- Ames test
- Bacterial reverse mutation test (Ames) in *S. typhimurium* (in vivo)
- Other in vitro endpoints
 - Chromosome Aberrations in Chinese Hamster Lung Cells
 - Mutations in Thymidine Kinase Locus in Mouse Lymphoma Cells
 - Mutations in HGPRT Locus in Chinese Hamster Ovary Cells
 - Unscheduled DNA Synthesis in Rat Hepatocytes
 - Syrian Hamster Embryo Cell Transformation
- In vivo endpoints
 - Sex-Linked Recessive Lethal Test in *Drosophila* m.
 - Micronucleus Test in Mouse Erythrocytes
 - Dominant Lethal Mutations in Rodents
 - Sister Chromatid Exchange in Mouse Bone Marrow Cells
 - Comet Assay in Mouse
- Carcinogenicity
 - Liver specific cancer (rat/mouse in vivo)

2. Draw the molecule/paste the SMILES



3. Click predict

Select models and input a chemical structure.

Home New query Quick start guide Model documentation Contact

Molecule Id (optional):

QSAR Results

Model	Experimental	Probability	Prediction	Report
Bacterial reverse mutation test (Ames) in <i>S. typhimurium</i> (in vivo)		0.05995	NEG_IN	

Save table

Prediction generation completed.
Use to request a detailed report.

OK

The target chemical is predicted as Positive or Negative with respect to the Ames endpoint based on the **probability** associated with the **prediction** generated by a partial logistic regression (PLR) model. In this case, the model returns a probability of **0.05995**, leading to a negative prediction.

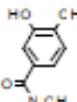
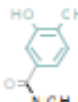
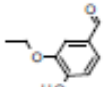
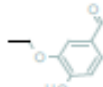
Danish (Q)SAR Database: running QSAR model

Model Features:
Danish_QSAR_DB_Bacterial_Reverse_Mutation_Test_Ames_mutagenicity_(S.typhimurium)_QSARmodels.food.dtu.dk v1


Model Features: DKDB_Ames_Mutagenicity_Multiple_Model-1 v1

Model Features: DKDB_Ames_Mutagenicity_Multiple_Model-2 v1


Analog Structures from Model Training Set

Structure	Similarity	Experimental Value - Danish_QSAR_DB_Bacterial_Reverse_Mutation_Test_Ames_mutagenicity_(S.typhimurium)_QSARmodels.food.dtu.dk	Danish_QSAR_DB_Bacterial_Reverse_Mutation_Test_Ames_mutagenicity_(S.typhimurium)_QSARmodels.food.dtu.dk - Highlights	Danish_QSAR_DB_Bacterial_Reverse_Mutation_Test_Ames_mutagenicity_(S.typhimurium)_QSARmodels.food.dtu.dk - Call	Danish_QSAR_DB_Bacterial_Reverse_Mutation_Test_Ames_mutagenicity_(S.typhimurium)_QSARmodels.food.dtu.dk - Prob.
 DKDB_2018_2561	0.7	Positive	 DKDB_2018_2561	Negative	0.287
	0.62	Negative		Negative	0.0271

Contact

Probability	Prediction	Report
0.05995	NEG_IN	

Save table

Prediction generation completed.
Use  to request a detailed report.

OK

01 Running VEGA models & ToxRead module and results analysis

02 Using Danish (Q)SAR Database and results analysis

03 Running **OCHEM** model & ToxAlerts and results analysis

04 Using AMBIT database and results analysis

TABLE OF CONTENTS

Checking the availability of the tools for in vitro gene mutation in bacteria in the CONCERT REACH gateway

HOME PROJECT RESULTS RESOURCES NEWS CONTACT

GATEWAY USER GUIDE

GATEWAY

8.4.1. In vitro gene mutation study in bacteria

ALL VEGA AND ToxRead DANISH QSAR DATABASE AMBIT **OCHEM**

End Point	Model	Type	Dataset size	Training set size	Test set size	Cross-validation procedure	Platform	Remarks
Ames test (OECD 471)	ASNN	Classification		4361	2181		OCHEM	

OCHEM
1 statistical model & ToxAlert match

Online chemical database
with modeling environment

Revision 2023-05-22 20:07:18 by 10.0.2.2 checked in on null. Built from null on null

Welcome, Dear Mrs.Raitano! [My account](#) [Logout](#)

Home Database Models

Model applicer X Model profile X Apply a model X

1. Search for the model by property

Models applicer browser

The complete list of models at OCHEM available for you is displayed below. If you are new here, you can also switch to a simplified [OCHEM predictor](#)

Model name or model ID: and property name: **Ames** Models visibility: Public and private Order by: creation time CONCERT members' models [refresh list](#)

1 - 1 of 1

	Ames levenberg published by midnighter	apply the model	predicts AMES using Ames challenge (training) (4361) validated by Ames challenge (test) (2181)	ASNN	2013-10-25
--	-------------------------------------------	---------------------------------	------------------------------------------------------------------------------------------------------	------	------------

1 - 1 of 1

[Next>>](#)

OCHEM: running model & ToxAlerts

Home Database Models A+ a- Privacy statement

Model applicator X Model profile X Apply a model X

Model profile

Statistical parameters, tables, charts - all the information related to the model.

2. Stats and info about AD

Overview Applicability domain

Model name: Ames levenberg, published in Applicability domain for <i>in silico</i> models to achieve accuracy of experimental measurements
Public ID is 1

Predicted property: **AMES** modeled in CLASS
Training method: ASNN

Data Set	#	Accuracy	Balanced Accuracy	MCC	AUC
Training set: Ames challenge (training)	4359 records	77.7% ± 0.6	77.5% ± 0.6	0.55 ± 0.01	0.852 ± 0.006
Test set: Ames challenge (test) [x]	2181 records	79.6% ± 0.8	79.5% ± 0.9	0.59 ± 0.02	0.86 ± 0.008

Show ROC curves

Real/Predicted →	inactive	active	Hit rate
inactive	1512	504	0.75
active	467	1876	0.801
Precision	0.764	0.788	

Training (Original)

Real/Predicted →	inactive	active	Hit rate
inactive	789	220	0.78
active	225	947	0.81
Precision	0.78	0.81	

Test (Original)

Number of compounds ignored because of errors in original model = 2

[Exclude duplicated records] Use optimal threshold for balanced accuracy

[Export this model](#) [View configuration XML](#) [Export configuration XML](#) [MMP-based analysis \(experimental\)](#)

Calculated in 2614 seconds
Size: 2554 Kb

APPLY THE MODEL TO NEW COMPOUNDS

Model applicator X Model profile X Apply a model X

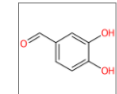
Model Applier

Provide the compound(s) to predict

Please provide compounds for which you want to predict the target property
Several options are available:

Upload compounds from a file
SDF, MOL2, SMILES or an Excel sheet Nessun file selezionato

Draw Molecule
click on depiction to the right to draw



[molecule profile]

Name/CASRN/SMILES:
e.g., "CC=CC" or "Aspirine" [load structure](#)

Choose a previously prepared set: [...]


Select molecules by a tag: [...]

Additional options

Disable prediction cache

3. Paste the SMILES


OCHEM: running model & ToxAlerts

 **Online chemical database**
with modeling environment

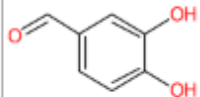
Home ▾ Database ▾ Models ▾

Model applier X Model profile X **Apply a model X** Prediction neighbors X Model's article X Export predictions X

OCHEM predictor - results ⓘ
Here you can browse the predictions for your compounds and export them in a variety of formats

 [Export results in a file \(Excel, CSV or SDF\)](#)

Sorting ▾ Ascending
1 - 1 of 1



molecule profile

AMES (Ames levenberg) = inactive (77.0% accuracy) **CACHED**

1 - 1 of 1

4. Online result

5. Export result: excel, CSV or SDF

SMILES	CASRN	EXTERNAIN	NAME	ARTICLEID	PUBMEDID	PAGE	TABLE	ERROR	AMES {predicted by ochem.eu/model/1}	Numeric prediction for AMES {predicted by ochem.eu/model/1}	PROB-STD{AMES by Ames levenberg}	ASNN-STDEV{AMES by Ames levenberg}	CLASS-LAG{AMES by Ames levenberg}	ASNN-CORREL{AMES by Ames levenberg}	Estimated accuracy{AMES predicted by Ames levenberg}	APPLICABILITY_DOMAIN	
OC1=CC=C(C=O)C=C1O				-	-	-	-		inactive	0.2437	0.18	0.09	0.49	0.26	0.77	TRUE	-

OCHEM: running model & ToxAlerts

Online chemical database
with modeling environment

Home Database Models

Welcome to ToxAlerts!

Structural alerts (also known as "toxicophores") are molecular patterns known to be associated with particular type of toxicity. The studies performed last decade has shown that structural alerts is an efficient technique to detect potentially toxic chemicals. Screening chemical compounds against known structural alerts can be a good practice to complement the QSAR models and to help interpreting their predictions.

ToxAlerts is a platform for screening chemical compounds against structural alerts. The platform allows to search structural alerts, introduce your own alerts and screen chemical libraries for alert-hitting compounds.

[View available alerts](#) [Upload new alerts](#) [Screen your molecules](#)

In case of any questions, ideas, or problems with the software, feel free do [drop us a message](#). We highly appreciate any feedback from you!

117 ToxAlerts for Genotoxic carcinogenicity, mutagenicity

Online chemical database
with modeling environment

Home Database Models

Welcome, Dear Mrs.Raitano [My account](#) [Logout](#)

v.4.2.476

ToxAlerts: Structural alerts browser

Here you can browser structural alerts for various toxicological endpoints

[Upload new alerts](#) [Screen compounds](#)

1 - 100 of 117 100 items on page 1 of 2 >>

	Aromatic nitro (general) Ar = any aromatic/heteroaromatic ring SMARTS: [a!r0][N(=O)=O] Endpoint: Genotoxic carcinogenicity, mutagenicity Kazius, J Derivation and validation of toxicophores for mutagenicity p... J. Med. Chem. 2005; 48 (1) 312-20 AlertID: 79521	15:52, 20 Feb 12 midnighter
	Aromatic amine (general) Ar = any aromatic/heteroaromatic ring SMARTS: [a!r0][N]([H]) Endpoint: Genotoxic carcinogenicity, mutagenicity Kazius, J Derivation and validation of toxicophores for mutagenicity p... J. Med. Chem. 2005; 48 (1) 312-20 AlertID: 79522	15:52, 20 Feb 12 midnighter
	Three-membered heterocycles (general) SMARTS: [N3]([H])([O])([S]) Endpoint: Genotoxic carcinogenicity, mutagenicity Kazius, J Derivation and validation of toxicophores for mutagenicity p... J. Med. Chem. 2005; 48 (1) 312-20 AlertID: 79523	15:52, 20 Feb 12 midnighter

ToxAlerts: Screening results
The compounds that matched any alerts grouped by endpoints, publications and by alerts themselves

ENDPOINTS

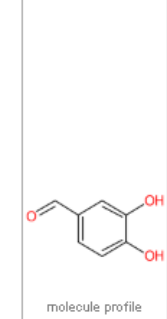
<input type="radio"/> Skin sensitization	1 compounds
<input type="radio"/> Genotoxic carcinogenicity, mutagenicity	1 compounds
<input type="radio"/> Reactive, unstable, toxic	1 compounds
<input type="radio"/> Acute Aquatic Toxicity	1 compounds
<input type="radio"/> Potential electrophilic agents	1 compounds
<input type="radio"/> Idiosyncratic toxicity (RM formation)	1 compounds
<input type="radio"/> Custom filters	1 compounds
<input type="radio"/> Extended Functional Groups (EFG)	1 compounds
<input type="radio"/> Promiscuity	1 compounds
<input type="radio"/> PAINS compounds	1 compounds
<input type="radio"/> Biodegradable compounds	1 compounds
<input type="radio"/> Nonbiodegradable compounds	1 compounds
<input type="radio"/> Non-genotoxic carcinogenicity	1 compounds
<input type="radio"/> UNIFAC	1 compounds
<input type="radio"/> Glaxo Wellcome	1 compounds
<input type="radio"/> Dundee	1 compounds
<input type="radio"/> LINT	1 compounds
<input type="radio"/> Inpharmatica	1 compounds
<input type="radio"/> ML5MR	1 compounds
<input type="radio"/> SureChEMBL	1 compounds
<input type="radio"/> SMARTCyp	1 compounds

PUBLICATIONS

<input type="radio"/> 1994 Barratt	1 compounds
<input type="radio"/> 2004 Gerner	1 compounds
<input type="radio"/> 1994 Payne	1 compounds
<input type="radio"/> 2008 Benigni	1 compounds
<input type="radio"/> 2008 Enoch	1 compounds
<input type="radio"/> 2011 ChemDiv	1 compounds
<input type="radio"/> 1990 Hermens	1 compounds
<input type="radio"/> 1992 Verhaar,H.J.M.	1 compounds
<input type="radio"/> 2011 Enoch	1 compounds
<input type="radio"/> 2005 Kalgutkar	1 compounds
<input type="radio"/> 2011 Life_Chemicals	1 compounds
<input type="radio"/> 2011 Enamine	1 compounds
<input type="radio"/> 2011 "Ontario"_filters	1 compounds
<input type="radio"/> 2011 Maybridge	1 compounds
<input type="radio"/> 2012 Tetko, I.V.	1 compounds
<input type="radio"/> 2005 CheckMol	1 compounds
<input type="radio"/> 2015 Salmina, ES	1 compounds
<input type="radio"/> 2006 Pearce, BC	1 compounds
<input type="radio"/> 2010 Baell, JB	1 compounds
<input type="radio"/> 2003 Environment	1 compounds
<input type="radio"/> 2013 Benigni, R	1 compounds
<input type="radio"/> 1991 Hansen	1 compounds
<input type="radio"/> 2003 Wittig	1 compounds
<input type="radio"/> 1999 Hann	1 compounds
<input type="radio"/> 2008 Brenk	1 compounds
<input type="radio"/> 2019 Blake JF	1 compounds
<input type="radio"/> 2017 Inpharmatica	1 compounds

View records for the filtered compounds Tag the 1 filtered molecules Export the screening results
1 - 1 of 1

Catechols and O-alkyl precursors (for Skin sensitization in 3894 Barratt)
Aldehydes (for Skin sensitization in 3894 Barratt)
Aldehydes (for Skin sensitization in 3904 Gerner)
Catechols, resorcinols, hydroquinones (for Skin sensitization in 3904 Gerner)
Aldehydes and precursors (for Skin sensitization in 3894 Payne)
Catechols (for Skin sensitization in 3894 Payne)
Di- or polyhydroxy aromatic compounds and their precursors (for Skin sensitization in 3894 Payne)
1,2-Dihydroxy aromatic compounds (for Skin sensitization in 3894 Payne)
Simple aldehyde (for Genotoxic carcinogenicity, mutagenicity in 3908 Benigni)
Aromatic aldehydes (for Skin sensitization in 3908 Enoch)
Ortho-disubstituted benzenes (for Skin sensitization in 3908 Enoch)
Aldehydes (for Reactive, unstable, toxic in 3911 ChemDiv)
Ortho-substituted phenols, primary and secondary amines (for Skin sensitization in 3908 Enoch)
Aldehydes (for Acute Aquatic Toxicity in 3890 Hermens)
Aldehydes (for Acute Aquatic Toxicity in 3892 Verhaar,H.J.M.)
Hydroquinones (for Potential electrophilic agents in 3911 Enoch)
Mono-aldehydes (for Potential electrophilic agents in 3911 Enoch)
Ortho- or parahydroquinones (for Idiosyncratic toxicity (RM formation) in 3905 Kalgutkar)
Aldehydes (for Reactive, unstable, toxic in 3911 Life_Chemicals)
Hydroquinone (for Reactive, unstable, toxic in 3911 Life_Chemicals)
Aldehydes (for Reactive, unstable, toxic in 3911 Enamine)
C, N, O, P and S atoms in unusual valence states (for Reactive, unstable, toxic in 3911 Enamine)
Aldehydes (for Reactive, unstable, toxic in 3911 "Ontario"_filters)
Catechols (for Reactive, unstable, toxic in 3911 "Ontario"_filters)
Aldehydes (including aminoformyl moieties) (for Reactive, unstable, toxic in 3911 Maybridge)



Simple anilines and phenols (for Reactive, unstable, toxic in 3911 Maybridge)
Atoms supported in ALOGPS program (for Custom filters in 3912 Tetko, I.V.)
Organic chemistry atoms (for Custom filters in 3912 Tetko, I.V.)
Carbonyl compounds: aldehydes or ketones (for Extended Functional Groups (EFG) in 3905 CheckMol)
Aldehydes (for Extended Functional Groups (EFG) in 3905 CheckMol)
Hydroxy compounds: alcohols or phenols (for Extended Functional Groups (EFG) in 3905 CheckMol)
Phenols (for Extended Functional Groups (EFG) in 3905 CheckMol)
Diphenols (for Extended Functional Groups (EFG) in 3915 Salmina, ES)
Aromatic compounds (for Extended Functional Groups (EFG) in 3905 CheckMol)
Arenes (for Extended Functional Groups (EFG) in 3915 Salmina, ES)
Nonmetals (for Extended Functional Groups (EFG) in 3915 Salmina, ES)
Chalcogens (oxygen group) (for Extended Functional Groups (EFG) in 3915 Salmina, ES)
Tetragens (carbon group) (for Extended Functional Groups (EFG) in 3915 Salmina, ES)
aldehyde (for Promiscuity in 3906 Pearce, BC)
Catechol_A (for PAINS compounds in 3910 Baell, JB)
1,2 - Diphenols (for Extended Functional Groups (EFG) in 3905 CheckMol)
Alcohols (for Biodegradable compounds in 3903 Environment)
Aldehyde (for Biodegradable compounds in 3903 Environment)
MoreThanTwoHydroxyOnAromaticRing (for Nonbiodegradable compounds in 3903 Environment)
Molecules with at least one carbon (for Custom filters in 3912 Tetko, I.V.)
Simple aldehyde (for Non-genotoxic carcinogenicity in 3913 Benigni, R)
3 - Aromatic carbon (for UNIFAC in 3891 Hansen)
8 - Aromatic carbon-alcohol (for UNIFAC in 3891 Hansen)
10 - Aldehyde (for UNIFAC in 3891 Hansen)
10 - ACH (3 - Aromatic carbon) (for UNIFAC in 3903 Wittig)
11 - AC (3 - Aromatic carbon) (for UNIFAC in 3903 Wittig)
18 - ACOH (8 - Aromatic carbon-alcohol) (for UNIFAC in 3903 Wittig)
21 - CHO (10 - Aldehyde) (for UNIFAC in 3903 Wittig)
A33 - phenol (for Glaxo Wellcome in 3899 Hann)

Paste the SMILES

Model applicer X Model profile X Apply a model X

Model Applicer

Provide the compound(s) to predict
Please provide compounds for which you want to predict the target property
Several options are available:

Upload compounds from a file
SDF, MOL, SMILES or an Excel sheet

Draw Molecule
click on depiction to the right to draw

Name/CASRN/SMILES:
e.g.: "CC=CC" or "Aspirin"

Choose a previously prepared set []

Select molecules by a tag []

Additional options


Disable prediction cache

[Next>>]

OCHEM: running model & ToxAlerts

Home Database Models A+ a- Privacy statement

OCHEM home page X Edit molecule X **Alert details X**

 **ToxAlerts: Structural alerts browser**
Here you can browser structural alerts for various toxicological endpoints

FILTERS

Article:

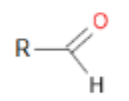
Endpoint / Filter type:

Name / Alert ID:

Show only approved alerts

[Upload new alerts](#) [Screen compounds](#) ✓ ✗

1 - 1 of 1



Simple aldehyde

R = aliphatic or aromatic carbon; α,β -unsaturated aldehydes are excluded

SMARTS: [CX3]([H])C(=O)X1[#1,#6&!\$(CX3)=CX3]

Endpoint: Genotoxic carcinogenicity, mutagenicity

[Benigni, R](#)
[Structure alerts for carcinogenicity, and the Salmonella ass...](#)
Mutat. Res. 2008; 659 (3) 248-61

Alert ID: 79368

15:52, 20 Feb 12 / 16:55, 6 Dec 12
midnighter / SALMINA1987

1 - 1 of 1

01

Running VEGA models & ToxRead module and results analysis

02

Using Danish (Q)SAR Database and results analysis

03

Running OCHEM model & ToxAlerts and results analysis

04

Using **AMBIT** database and results analysis

TABLE OF CONTENTS

Checking the availability of the tools for in vitro gene mutation in bacteria in the CONCERT REACH gateway

HOME PROJECT RESULTS RESOURCES NEWS CONTACT

GATEWAY USER GUIDE

GATEWAY

8.4.1. In vitro gene mutation study in bacteria

ALL VEGA AND ToxRead DANISH QSAR DATABASE AMBIT **OCHEM**

End Point	Model	Type	Dataset size	Training set size	Test set size	Cross-validation procedure	Platform	Remarks
Ames test (OECD 471)	ASNN	Classification		4361	2181		OCHEM	 

OCHEM
1 statistical model & ToxAlert match

ALL VEGA AND ToxRead DANISH QSAR DATABASE **AMBIT** OCHEM

End Point	Model	Type	Dataset size	Training set size	Test set size	Cross-validation procedure	Platform	Remarks
TOX 7.6.1. Genetic toxicity in vitro	Dataset		50366				AMBIT	

AMBIT
1 dataset

Using **AMBIT** database and results analysis

LRI AMBIT2 Read Across tool - new version!

Chemical substance database with read across workflow : IUCLID6 support; featuring OpenFoodToxData and VEGA integration

Simple search

Enter chemical name, identifiers, SMILES, InChI

Advanced: [Structure search](#) | [Search substances by identifiers](#) | [Search substances by endpoint data](#) | [Free text search](#)

Legal notice:

The LRI AMBIT - IUCLID tool is loaded with non-confidential REACH data supplied by ECHA.

The legal notice from the ECHA dissemination website <http://echa.europa.eu/web/guest/legal-notice#registration> applies to the AMBIT users.

In addition, [Cefic](#) disclaims any liability of whatsoever nature either direct or indirect regarding the use of the AMBIT-IUCLID tool or information / data contained in it.

[IdeaConsult](#) is a contractor of Cefic developing and hosting the AMBIT-IUCLID tool. Some data used may have been provided by Cefic. IdeaConsult has acted solely on the instructions of Cefic. IdeaConsult disclaims any liability of whatsoever nature, direct or indirect, regarding the use of any information/data by the AMBIT-IUCLID tool. IdeaConsult shall not have any liability of whatsoever nature for any use or misuse of the AMBIT-IUCLID tool.

Using **AMBIT** database and results analysis

New assessment ▶ Use of empty template

Existing assessments ▶ Use of Create new read across or category formation workflow

LRI AMBIT REACH - new version!

Chemical substance database with read across workflow : IUCLID6 support; featuring OpenFoodToxData and VEGA integration

Simple search

Enter chemical name, identifiers, SMILES, InChI

Advanced: [Structure search](#) | [Search substances by identifiers](#) | [Search substances by endpoint data](#) | [Free text search](#)

Legal notice:

The LRI AMBIT - IUCLID tool is loaded with non-confidential REACH data supplied by ECHA.

The legal notice from the ECHA dissemination website <http://echa.europa.eu/web/guest/legal-notice#registration> applies to the AMBIT users.

In addition, [Cefic](#) disclaims any liability of whatsoever nature either direct or indirect regarding the use of the AMBIT-IUCLID tool or information / data contained in it.

[IdeaConsult](#) is a contractor of Cefic developing and hosting the AMBIT-IUCLID tool. Some data used may have been provided by Cefic. IdeaConsult has acted solely on the instructions of Cefic. IdeaConsult disclaims any liability of whatsoever nature, direct or indirect, regarding the use of any information/data by the AMBIT-IUCLID tool. IdeaConsult shall not have any liability of whatsoever nature for any use or misuse of the AMBIT-IUCLID tool.

Using **AMBIT** database and results analysis

Home > All assessments >

1. Assessment identifier


Assessment identifier

Collect structures


Endpoint data used

Assessment details

Report

Assessment title* ⓘ:	3,4-Dihydroxybenzaldehyde
Owner* ⓘ:	Nelly
Purpose* ⓘ:	QSAR2023 course
Version ⓘ:	?.?
Version start date ⓘ:	
Version last modified on ⓘ:	
Status ⓘ:	
Assessment code* ⓘ:	read across1
Assessment Doclink(s)* ⓘ:	local 
Assessment ID ⓘ:	
Users with write access ⓘ:	<input type="text"/> <input type="button" value="Save"/>
Users with read access ⓘ:	<input type="text"/> <input type="button" value="Save"/>

Five steps assessment

category formation. [REACH guidance](#) 

The assessment ⓘ workflow is organized in five main tabs:

1. Assessment identifier ⓘ
2. Collect structures ⓘ
3. Endpoint data used ⓘ
4. Assessment details ⓘ
5. Report ⓘ

Start

Automatic generation of the ID assessment

Using **AMBIT** database and results analysis

[Home](#) > [All assessments](#) >

2. Collect structures, for the **target** as well as for the **analogues** and **category members**

Assessment identifier

Collect structures

Endpoint data used

Assessment details

Report

Collect structures

List collected

Search for the target: CAS N, SMILES, drawing etc..

Search

Exact structure

Similarity

Substructure

URL

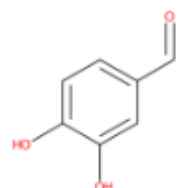

target

Enable fragment search

O=Cc1ccc(O)c(O)c1

showing from 1 to 1 in pages of 20 entries ◀ Previous Next ▶

Filter...

Diagram	CasRN	EC number	Names	Rationale
 	139-85-5	205-377-7	3,4-Dihydroxybenzaldehyde 3,4-dihydroxybenzaldehyde	Reason for selection_

Using **AMBIT** database and results analysis

2. Collect structures, for the target as well as for the analogues and category members

Assessment identifier Collect structures Endpoint data used Assessment details Report


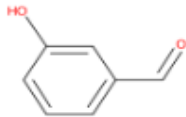



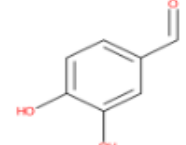
Collect structures List collected

Search

Exact structure Similarity Substructure URL

Only hits with substance data 0.8 O=Cc1ccc(O)c(O)c1

Showing from 1 to 3 in pages of 20 entries Previous Next

	Diagram	CasRN	EC number	Names	Similarity	Rationale
		100-83-4	202-892-9	3-hydroxybenzaldehyde Benzaldehyde, 3-hydroxy-	0.91	
		123-08-0	204-599-1	4-Hydroxybenzaldehyde 4-hydroxybenzaldehyde	0.85	
		139-85-5	205-377-7	3,4-Dihydroxybenzaldehyde 3,4-dihydroxybenzaldehyde	1	

Using **AMBIT** database and results analysis

Home > All assessments >

3. Endpoint data used

Assessment identifier

Collect structures

Endpoint data used

Assessment details

Report

Search substance(s)

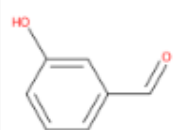

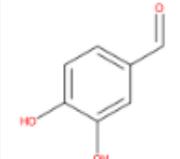
Selection of endpoints

Expand all

Collapse all

Showing from 1 to 3 in pages of 20 entries < Previous Next >

Filter...

	Diagram	CasRN	EC number	Rational	
- 1 -		100-83-4	202-892-9	hydroxybenzaldehyde Benzaldehyde, 3-hydroxy-	
- 2 -		123-08-0	204-599-1	4-Hydroxybenzaldehyde 4-hydroxybenzaldehyde	S
- 3 -		139-85-5	205-377-7	3,4-Dihydroxybenzaldehyde 3,4-dihydroxybenzaldehyde	T

B. Select the endpoint of interest

A. Expand all the info about the chemicals and then select

Using **AMBIT** database and results analysis



LIFE17 GIE/IT/000461



Search ▾ Assessments ▾ Import ▾ Enhanced functions ▾ Admin ▾ Help ▾

[giuseppa.raitano] Log out

Home > All assessments >

3. Endpoint data used

Assessment identifier

Collect structures

Endpoint data used

Assessment details

Report

Search substance(s)

Selection of endpoints

Show all endpoints

Filter...

▶ P-Chem	select all unselect all [91169]
▶ Env Fate	select all unselect all [55690]
▶ Eco Tox	select all unselect all [115597]
▼ Tox	select all unselect all [228703]

Showing 18 endpoint(s) (1 to 18)

<input type="checkbox"/>	7.2.1. Acute toxicity - oral	[26076] ⓘ
<input type="checkbox"/>	7.2.2. Acute toxicity - inhalation	[9213] ⓘ
<input type="checkbox"/>	7.2.3. Acute toxicity - dermal	[11852] ⓘ
<input type="checkbox"/>	7.3.1. Skin irritation / Corrosion	[21694] ⓘ
<input type="checkbox"/>	7.3.2. Eye irritation	[18430] ⓘ
<input type="checkbox"/>	7.4.1. Skin sensitisation	[16404] ⓘ
<input type="checkbox"/>	7.5.1. Repeated dose toxicity - oral	[20212] ⓘ
<input type="checkbox"/>	7.5.2. Repeated dose toxicity - inhalation	[6657] ⓘ
<input type="checkbox"/>	7.5.3. Repeated dose toxicity - dermal	[2139] ⓘ
<input checked="" type="checkbox"/>	7.6.1. Genetic toxicity in vitro	[50366] ⓘ
<input type="checkbox"/>	7.6.2. Genetic toxicity in vivo	[10232] ⓘ
<input type="checkbox"/>	7.7. Carcinogenicity	[6288] ⓘ

Using **AMBIT** database and results analysis

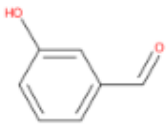
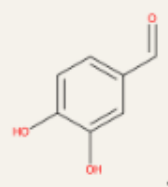
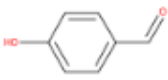
4. Assessment details

Assessment identifier Collect structures Endpoint data used **Assessment details** Report

Initial matrix Working matrix Final matrix

Identifiers TOX

Showing from 1 to 3 in pages of 20 entries < Previous Next >

Substance Name	ISUUID	Data source	Tag	Diagram	Constituents	7.6.1
3-hydroxybenzaldehyde	ECHA-8a...	-	S		Benzaldehyde hydroxy-	Structural Alert for <i>S. typhimurium</i> mutagenicity = YES Potential <i>S. typhimurium</i> TA100 mutagen based on QSAR = NO Mutagenicity (Ames test) model (CAESAR) - prediction = NON-Mutagenic No alerts for <i>S. typhimurium</i> mutagenicity = NO Unlikely to be a <i>S. typhimurium</i> TA100 mutagen based on QSAR = NO
3,4-DIHYDROXYBENZALDEHYDE	FOOD-64...	OpenFoodToxData	T		3,4-Dihydroxy	Structural Alert for <i>S. typhimurium</i> mutagenicity = YES Potential <i>S. typhimurium</i> TA100 mutagen based on QSAR = NO Mutagenicity (Ames test) model (CAESAR) - prediction = NON-Mutagenic No alerts for <i>S. typhimurium</i> mutagenicity = NO Unlikely to be a <i>S. typhimurium</i> TA100 mutagen based on QSAR = NO Negative (EFSA opinion) Negative (EFSA opinion) Negative (EFSA opinion) Negative (EFSA opinion) Negative (EFSA opinion)
4-HYDROXYBENZALDEHYDE	FOOD-a4...	OpenFoodToxData	S		4-Hydroxybe	Structural Alert for <i>S. typhimurium</i> mutagenicity = YES Potential <i>S. typhimurium</i> TA100 mutagen based on QSAR = NO Mutagenicity (Ames test) model (CAESAR) - prediction = NON-Mutagenic No alerts for <i>S. typhimurium</i> mutagenicity = NO

- 1) Initial matrix: all the available info is listed
- 2) Working matrix: it is possible to add or eliminate data
- 3) Final matrix

Using **AMBIT** database and results analysis



LIFE17 GIE/IT/000461



[Search](#) ▾ [Assessments](#) ▾ [Import](#) ▾ [Enhanced functions](#) ▾ [Admin](#) ▾ [Help](#) ▾

[\[giuseppa.raitano\]](#) [Log out](#)

[Home](#) > [All assessments](#) >

5. Report



[Search](#) ▾ [Assessments](#) ▾ [Import](#) ▾ [Enhanced functions](#) ▾ [Admin](#) ▾ [Help](#) ▾

[\[giuseppa.raitano\]](#) [Log out](#)

[Home](#) > [All assessments](#) > [This assessment](#) > [Report](#) >

[Create Word file](#)

Ambit Assessment Report

3,4-Dihydroxybenzaldehyde

Author: Nelly
Date: 03.06.2023
Assessment code: ecd6a503-83e3-40c8-a1f2-8bca2eee1036
Purpose: QSAR2023 course

1. Assessment Identifiers

Assessment title:	3,4-Dihydroxybenzaldehyde
Owner:	Nelly
Purpose:	QSAR2023 course
Version:	1
Status:	draft
Version start date:	03.06.2023
Version last modified on:	03.06.2023
Published:	draft
Assessment code:	read across1
Assessment Doc link:	local

2-day workshop

Monday 19/06, full day

Workshop presentations

Tuesday 20/06, morning

Training sessions



The screenshot shows the website for the ConcertReach project. At the top, there is a yellow header with the European Union flag and the text "LIFE17 GIE/IT/000461". Below this is the ConcertReach logo, which consists of a colorful circular pattern of dots and the text "CONCERTREACH CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH". To the right of the logo is a navigation menu with the following items: HOME, PROJECT, RESULTS, RESOURCES, NEWS, and CONTACT. The main content area has a dark grey background with a network diagram of white nodes and lines. A large black box on the left contains the text "Final Workshop" in white. Below this, there is a white box containing an illustration of two people (a man and a woman) and a computer monitor displaying code and data. To the right of the illustration, the text reads: "The EU LIFE CONCERT REACH project opens a web-based 'gateway' for the exploitation of (Q)SAR models in the (eco)toxicological evaluation of new compounds". Below this, it says "2 days Workshop Hybrid Event". Further down, it states "Organized by: Coordinating Beneficiary." and "It can be attended in person at the Mario Negri Institute, in Milan, Italy and virtually." At the bottom, it says "Admission is free of charge. Please make your registration in advance, since there is maximum number of participants. We will notify acceptance of the registered participants."

LIFE17 GIE/IT/000461

CONCERTREACH
CONCERTING EXPERIMENTAL DATA
AND IN SILICO MODELS FOR REACH

HOME PROJECT RESULTS RESOURCES NEWS CONTACT

Final Workshop



The EU LIFE CONCERT REACH project opens a web-based "gateway" for the exploitation of (Q)SAR models in the (eco)toxicological evaluation of new compounds

2 days Workshop
Hybrid Event

Organized by: Coordinating Beneficiary.

It can be attended in person at the Mario Negri Institute, in Milan, Italy and virtually.

Admission is free of charge. Please [make your registration](#) in advance, since there is maximum number of participants. We will notify acceptance of the registered participants.

Date
19 - 20 June 2023.

Venue
Istituto di Ricerche Farmacologiche Mario Negri
IRCCS.

<https://www.life-concertreach.eu/final-workshop-19-and-20-june-2023/>

THANKS

Does anyone have any questions?
<https://www.life-concertreach.eu/>



Acknowledgment:
My colleagues,
All the partners of the project,
Rodolfo Gonella Diaza from Knoell



ISTITUTO DI RICERCHE
FARMACOLOGICHE
MARIO NEGRI · IRCCS

giuseppa.raitano@marionegri.it

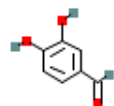
SEARCH FOR

3,4-Dihydroxybenzaldehyde



Treating this as a text search.

BEST MATCH



[3,4-Dihydroxybenzaldehyde](#); **Protocatechualdehyde**; **139-85-5**; **PROTOCATECHUIC ALDEHYDE**; **Rancinamycin IV**; **Benzaldehyde, 3,4-Dihydroxy-**; **4-Formyl-1,2-Dihydroxybenzene**; **3,4-Dihydroxybenzenecarbal**; ...

Compound CID: [8768](#)

MF: [C₇H₆O₃](#) **MW:** 138.12g/mol

IUPAC Name: [3,4-dihydroxybenzaldehyde](#)

Isomeric SMILES: [C1=CC\(=C\(C=C1C=O\)O\)O](#)

InChIKey: [IBGBGRVKPALMCQ-UHFFFAOYSA-N](#)

InChI: [InChI=1S/C7H6O3/c8-4-5-1-2-6\(9\)7\(10\)3-5/h1-4,9-10H](#)

Create Date: 2005-03-26

[Summary](#)

[Similar Structures Search](#)

[Related Records](#)

01

Welcome to the VEGA HUB

Offering a family of tools to evaluate chemical hazard: VEGA, ToxRead, SWAN, VERA, ToxWeight, ToxDelta, and JANUS.

VEGA is the QSAR software with tens of models for individual properties.



Do you need assistance for a property prediction?

CONTACT US

VEGA HUB ▾ QSAR ▾ Download ▾

Models and Tools

Assistance

VEGAHUB

Community News Contacts

Our philosophy

The In silico methods can be very useful, if correctly applied



We want to enhance

We want to enhance the capability of the in silico methods to assess the properties of chemical substances.

We sustain

We sustain the correct use of 4n silico methods

We provide

We provide computer tools to support the human experts

We support

We support the evaluation of chemicals safety

<https://www.vegahub.eu/download/>



VEGA QSAR

Try out VEGA QSAR
and its features



All the VEGA models are also available in a unique stand-alone application.

With the VEGA application you can easily execute all the models on your local machine without sending any information to our server. VEGA is the ideal application for batch processing large dataset. VEGA can be installed and used on any operative system supporting JAVA technology (for any doubt please visit [JAVA website](#)).

[Introduction](#)

[Screenshots](#)

[Interpretation](#)

[How to quote VEGA QSAR](#)

[VEGA QSAR for KNIME](#)

[Download](#)

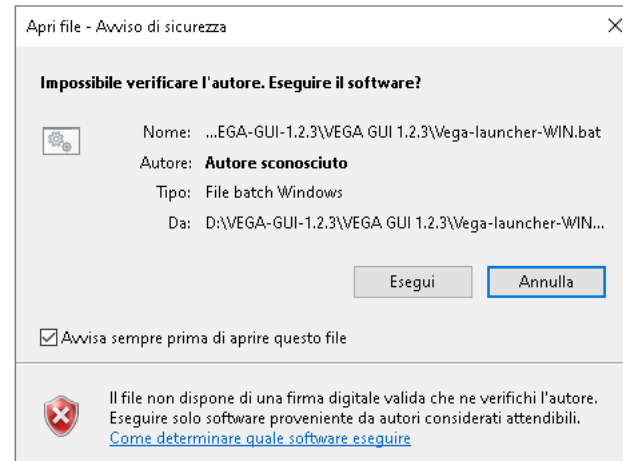
[Previous Versions](#)

Free download VEGA QSAR Application

Visit the link to download the application.

DOWNLOAD

Nome	Ultima modifica	Tipo	Dimensione
jdk-11.0.6	24/05/2023 09:18	Cartella di file	
README	24/05/2023 09:17	Documento di testo	2 KB
Vega-GUI-1.2.3	24/05/2023 09:18	Executable Jar File	254,025 KB
Vega-launcher-LINUX.sh	24/05/2023 09:18	File SH	1 KB
Vega-launcher-WIN	24/05/2023 09:18	File batch Windows	1 KB



VEGA in silico platform - version 1.2.3

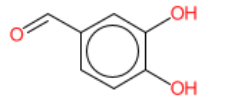
Insert chemicals

1. Add molecule(s) using SMILES notation

Insert SMILES:
O=Cc1ccc(O)c(O)c1 + Import File

ID	SMILES
Molecule 0	<chem>O=Cc1ccc(O)c(O)c1</chem>

2. Added molecules are listed and 2D structure can be visualized



Delete All Delete

VEGA in silico platform - version 1.2.3

Select models

Filter models: All available endpoints

Mutagenicity (Ames test)

- Select all models
- Mutagenicity (Ames test) model (CAESAR) - v. 2.1.14
- Mutagenicity (Ames test) model (ISS) - v. 1.0.3
- Mutagenicity (Ames test) model (SarPy-IRFMN) - v. 1.0.8
- Mutagenicity (Ames test) model (KNN-Read-Across) - v. 1.0.1
- Mutagenicity (Ames test) model for aromatic amines (CONCERT/IRFMN) - v. 1.0.0
- Mutagenicity (Ames test) CONSENSUS model - v. 1.0.4

Developmental toxicity

Select all models

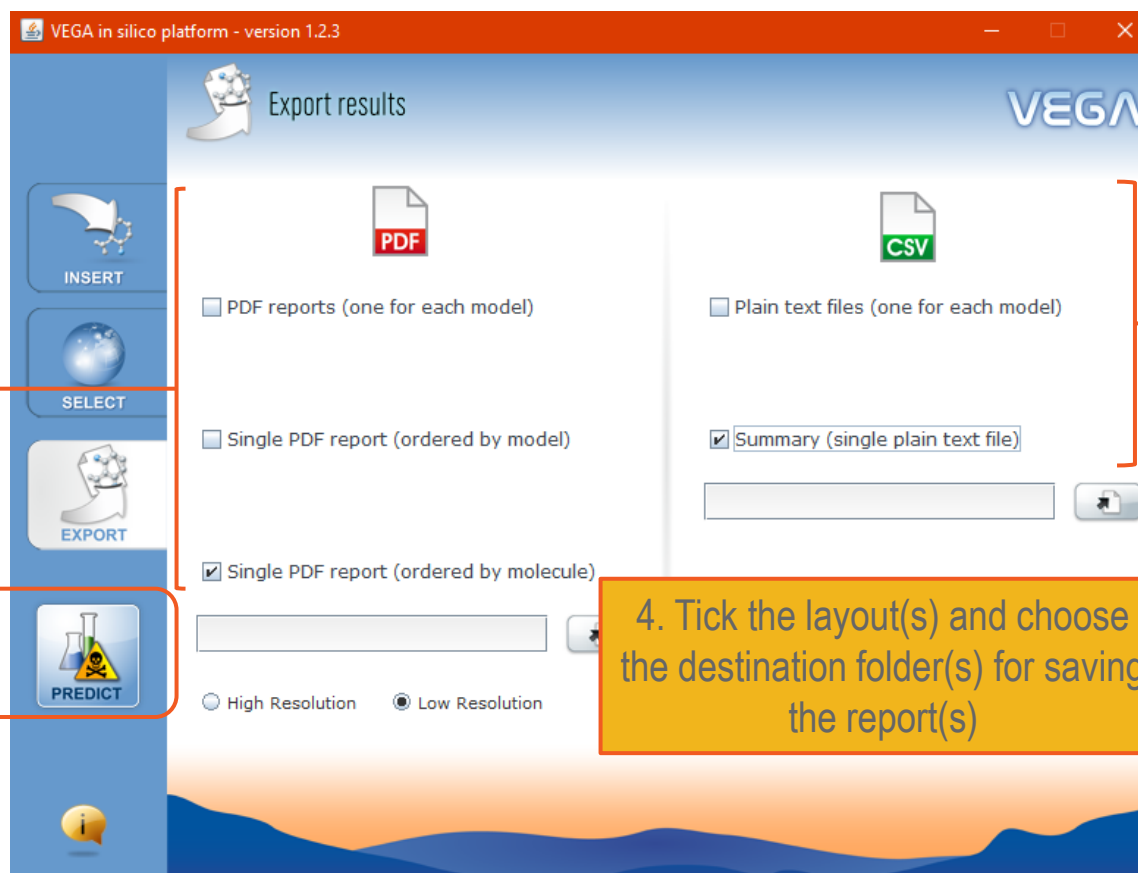
3. Select the model(s)

VEGA: running predictions

Full PDF reports:

- prediction(s) results
- applicability domain
- experimental data of the target (if any)
- most similar substances
- other supporting info (if any)

5. Click on «Predict»



Danish (Q)SAR Database: gathering (Q)SAR results

Danish (Q)SAR Database

Home Clear Information Contact QSAR2023

New search

Searches

Results

Substances

ID

Structure and name

1. Input by structure

PhysChem

Environment

ADME

Human health

AND
Intersect results

OR
Unite results

NOT
Complement results

MAX/MIN
More combinations

2. Import

3. Paste SMILES

4. Confirm SMILES

5. Search the database for the target molecule

Single structure SMILES list

Substructure Similarity

Name search

Substructure search

Exact match search

Cancel

OK Cancel

O=Cc1ccc(O)c(O)c1

Danish (Q)SAR Database: gathering (Q)SAR results

Danish (Q)SAR Database

Home Clear Information Contact QSAR2023

New search

ID
Structure and name

PhysChem
Environment
ADME
Human health

AND
Intersect results
OR
Unite results
NOT
Complement results
MAX/MIN
More combinations

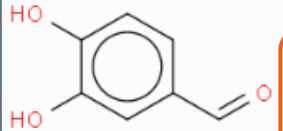

Searches 1. Results 1
Exact match: >

6. Summary of the search performed

Substances

Exact match: Page 1
Previous Next 1

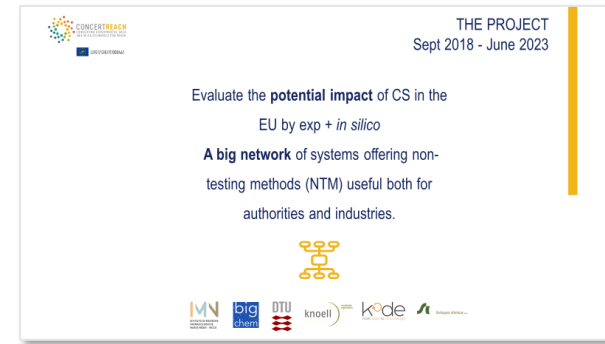
Structures 1-1 of 1

Structure	Id	Similarity	+
	139-XX-X 		

7. For each identified molecule, the (Q)SAR report can be downloaded in .rtf format

The GATEWAY

Why it will be useful both for regulators and the industry



Feedback from regulators & end-users (via workshops):

- Regulators recommend multiple systems.
 - Within the 4 platforms are available **several in silico models**
- **Difficulties**: access, quality/choice, use, interpretation, integration
 - Access. Which models? Commercial/public. **We cover both**
 - Quality. Any tool possible, providing doc. **We focus on 4 well-known systems**
 - Use. CONCERT **improves the use and explanation**
 - Integration. CONCERT **improves integration**

The GATEWAY

Why it will be useful both for regulators and the industry

- Website active for **5 years** after the end of the project
- Selection of the **endpoints for other regulations** then REACH
(cosmetics...)
- Continuous updating by adding **new tools and platforms**

in silico tool	Prediction	reliability
Mutagenicity (Ames test) CONSENSUS model-assessment	NON-Mutagenic (Consensus score: 0.675)	0.675
Mutagenicity (Ames test) model (CAESAR)	NON-Mutagenic (GOOD reliability)	0.96
Mutagenicity (Ames test) model (ISS)	Mutagenic (LOW reliability)	0
Mutagenicity (Ames test) model (SarPy-IRFMN)	Possible NON-Mutagenic (GOOD reliability)	0.96
Mutagenicity (Ames test) model (KNN-Read-Across)	NON-Mutagenic (GOOD reliability)	0.965
Toxread	NON-Mutagenic	
Battery	NEG_IN	
CASE Ultra	NEG_IN	
Leadscope	NEG_IN	aggiungere probability
SciQSAR	NEG_OUT	
OCHEM	INACTIVE	0.77
OCHEM ToxAlerts	Simple Aldehyde	
AMBIT	negative	