Istituto di Ricerche Farmacologiche Mario Negri IRFMN

Alessandra Roncaglioni

March 4th, 2024

AiChemist PI intro

The institute

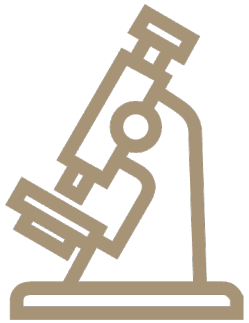
ISTITUTO DI RICERCHE FARMACOLOGICHE MARIO NEGRI · IRCCS

Private non-Profit biomedical research organization (est. 1963)

- Carrying out scientific research
- Training young people
- Disseminating the results of our research

3 locations (MI-BG)

600 people

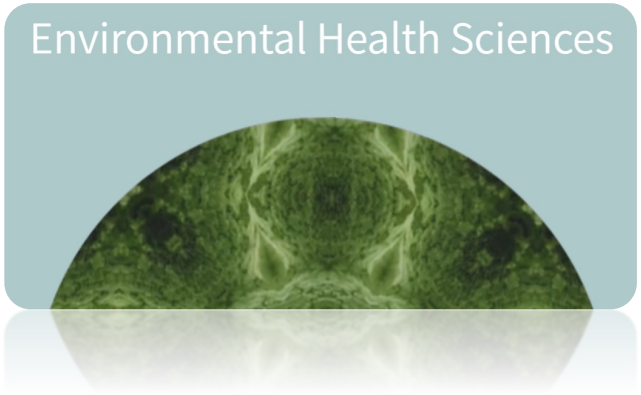


10 departments

Department of Environment Health Sciences
Lab. of Environmental Toxicology and Chemistry

ISTITUTO DI RICERCHE FARMACOLOGICHE MARIO NEGRI · IRCCS

Head of department:
Emilio Benfenati



4 labs + 1 unit

Laboratory of Chemistry and Environmental Toxicology

Head of laboratory
Alessandra Roncaglioni

About me

- MSc Environ. Sciences (UniMI Bicocca)
  - Intership@IRFMN
- PhD Open University
  - QSARs for endocrine disrupters
- Visiting scientist (Slovenia & France)
- Head Comp Tox unit @IRFMN
- Head of Lab @IRFMN

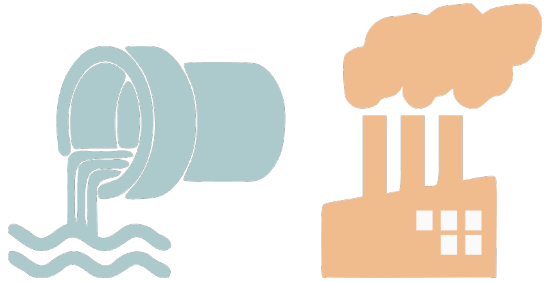
Laboratory of Chemistry and Environmental Toxicology
What we do

ISTITUTO DI RICERCHE FARMACOLOGICHE MARIO NEGRI · IRCCS

Identifying and preventing the negative effects of environmental factors and contaminants on health

COMPUTATIONAL TOXICOLOGY

RISK ASSESSMENT



ANALYTICAL MONITORING ACTIVITIES

Hazard assessment:
In silico modeling (QSAR, read-across) of (eco)tox data
(Eco)tox bioassays

Exposure assessment:
Environmental monitoring
Internal exposure estimation

Laboratory of Chemistry and Environmental Toxicology
What we do

ISTITUTO DI RICERCHE FARMACOLOGICHE MARIO NEGRI · IRCCS

RISK ASSESSMENT

Research activities in regulatory toxicology

NAMs — New approach methodologies: any non-animal technology, methodology, approach, or combination thereof that can be used to provide information on chemical hazard and risk assessment

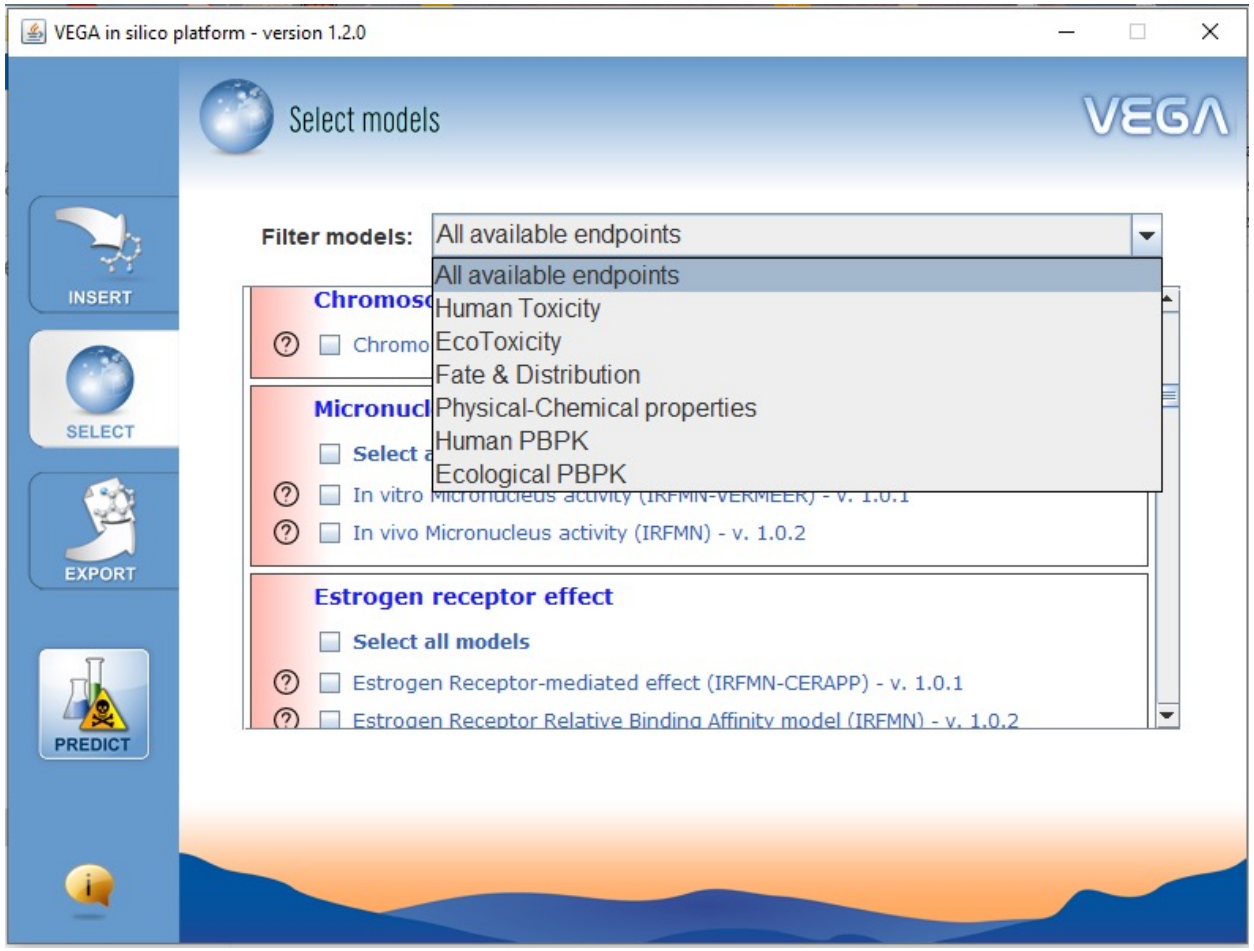
AOPs — Develop efficient and innovative safety testing strategies through a comprehensive mechanistic understanding of cause-consequence relationships of adverse chemical effects

COMPUTATIONAL TOXICOLOGY

A variety of tools providing computer simulations of complex biological phenomena:

- (Q)SAR
- Molecular modelling
- Bio/Chem Database
- Machine Learning
- Artificial Intelligence
- Read across
- ADME profiling

SW for QSAR



- Unified environment encoding tens of models for different effects: (eco)toxicity, environmental properties, phys-chem, etc….
  - Multiparametric applicability domain definition

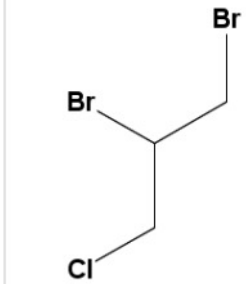
www.vegahub.eu

Applicability Domain Index (ADI)

ISTITUTO DI RICERCHE FARMACOLOGICHE MARIO NEGRI · IRCCS

1. Prediction Summary

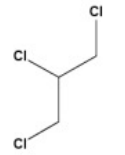
Prediction for compound 1 (Molecule 1)



Prediction: Reliability:

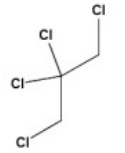
Model assessment: Prediction is logBCF = 1.65, the result appears reliable. Anyhow, you should check it through the evaluation of the information given in the following sections.

Compound: 1
Compound SMILES: C(C(CBr)Br)Cl
Experimental value: -
Prediction: 1.65 [log(L/kg)]
Prediction: 45 [L/kg]
Prediction of model 1 (HM): 1.75 [log(L/kg)]
Prediction of model 2 (GA): 1.61 [log(L/kg)]
Structural Alerts: -
Calculated LogP: 2.96 [log units]
Reliability: Compound is in model Applicability Domain
Remarks for the prediction:
none



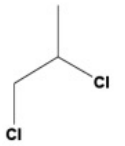
CAS: 96-18-4
Dataset id: 255 (training set)
SMILES: C(C(CCl)Cl)Cl
Similarity: 0.997

Experimental value: 0.96 [log(L/kg)]
Predicted value: 1.14 [log(L/kg)]



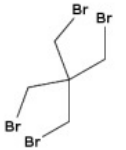
CAS: 13116-53-5
Dataset id: 256 (training set)
SMILES: C(C(CCl)(Cl)Cl)Cl
Similarity: 0.966

Experimental value: 1.57 [log(L/kg)]
Predicted value: 1.37 [log(L/kg)]



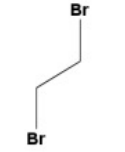
CAS: 78-87-5
Dataset id: 254 (training set)
SMILES: CC(CCl)Cl
Similarity: 0.963

Experimental value: 0.57 [log(L/kg)]
Predicted value: 0.9 [log(L/kg)]



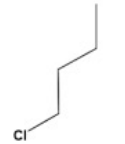
CAS: 3229-00-3
Dataset id: 253 (training set)
SMILES: C(C(CBr)(CBr)CBr)Br
Similarity: 0.944

Experimental value: 2.62 [log(L/kg)]
Predicted value: 2.88 [log(L/kg)]



CAS: 106-93-4
Dataset id: 249 (training set)
SMILES: C(CBr)Br
Similarity: 0.859

Experimental value: 0.35 [log(L/kg)]
Predicted value: 0.84 [log(L/kg)]



CAS: 109-69-3
Dataset id: 250 (training set)
SMILES: CCCCCl
Similarity: 0.858

Experimental value: 1.15 [log(L/kg)]
Predicted value: 1.05 [log(L/kg)]

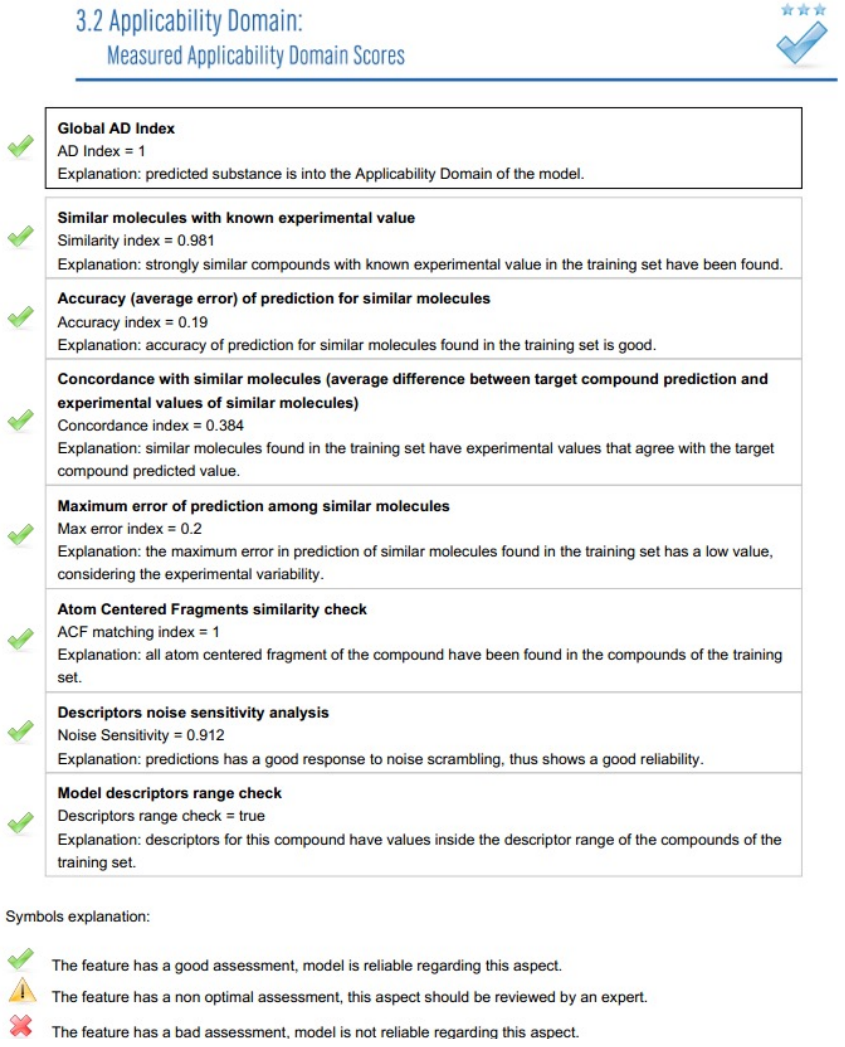
Applicability Domain Index (ADI)

ISTITUTO DI RICERCHE FARMACOLOGICHE MARIO NEGRI · IRCCS

- ✓ Visualization of similar substances
- ✓ Similarity index
- ✓ Descriptors range
- ✓ Atom centered-fragment

- ✓ Check of the descriptor sensitivity
- ✓ Uncertainty

- ✓ Fragments for outliers
- ✓ Prediction Accuracy
- ✓ Prediction Concordance



3.2 Applicability Domain:
Measured Applicability Domain Scores

Global AD Index
AD Index = 1
Explanation: predicted substance is into the Applicability Domain of the model.

Similar molecules with known experimental value
Similarity index = 0.981
Explanation: strongly similar compounds with known experimental value in the training set have been found.

Accuracy (average error) of prediction for similar molecules
Accuracy index = 0.19
Explanation: accuracy of prediction for similar molecules found in the training set is good.

Concordance with similar molecules (average difference between target compound prediction and experimental values of similar molecules)
Concordance index = 0.384
Explanation: similar molecules found in the training set have experimental values that agree with the target compound predicted value.

Maximum error of prediction among similar molecules
Max error index = 0.2
Explanation: the maximum error in prediction of similar molecules found in the training set has a low value, considering the experimental variability.

Atom Centered Fragments similarity check
ACF matching index = 1
Explanation: all atom centered fragment of the compound have been found in the compounds of the training set.

Descriptors noise sensitivity analysis
Noise Sensitivity = 0.912
Explanation: predictions has a good response to noise scrambling, thus shows a good reliability.

Model descriptors range check
Descriptors range check = true
Explanation: descriptors for this compound have values inside the descriptor range of 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.

ISTITUTO DI RICERCHE
FARMACOLOGICHE
MARIO NEGRI · IRCCS

Thank you for your attention!