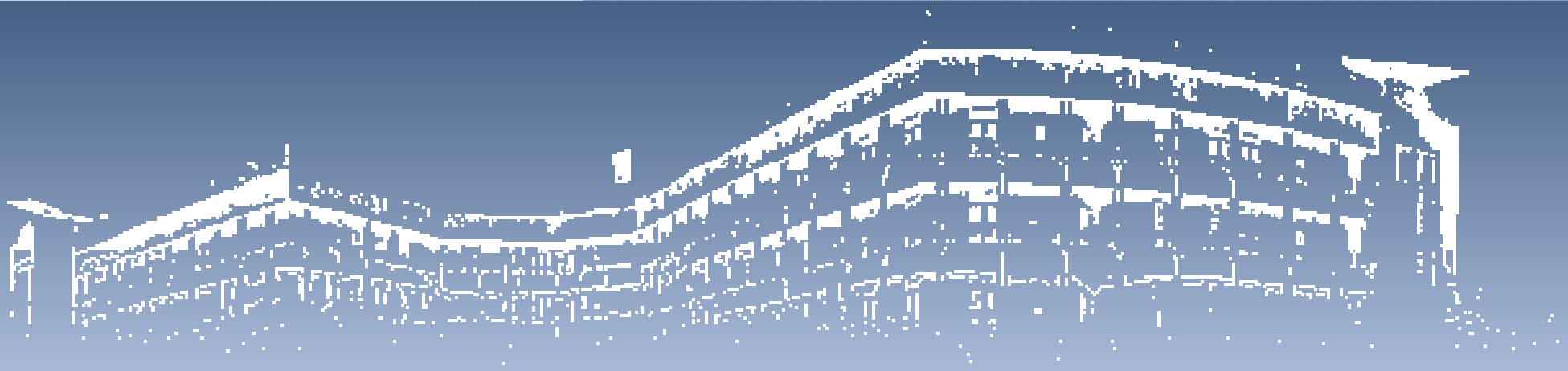
Emilio BENFENATI

Istituto di Ricerche Farmacologiche Mario Negri
Laboratory of Environmental Chemistry & Toxicology



Rethinking the equation REACH = Regulation

15 dicembre 2011

La nuova piattaforma VEGA per i metodi di non-testing

MARIO NEGRI ISTITUTO DI RICERCHE FARMACOLOGICHE

ECT Environment Chemistry Toxicology

ORCHESTRA

antares

antares

Alternative Non-Testing methods Assessed for REACH Substances

LIFE08
ENV/IT/000435

MARIO NEGRI
ISTITUTO DI RICERCHE FARMACOLOGICHE

ISTITUTO SUPERIORE DI SANITÀ

FEDERCHIMICA
CONFINDUSTRIA

POLITECNICO DI MILANO

KnowledgeMiner Software

www.antares-life.eu

ORCHESTRA

Organising dissemination on Results of projects on CHemicals Evaluation, Spreading Techniques for Risk Assessment

www.orchestra-qsar.eu

AND

According to *REACH regulation* (Annex XI) a (Q)SAR is VALID if:

- the model is recognized *scientifically valid*;
- the substance is included in the *applicability domain* of the model;
- results are *adequate* for *classification and labelling and for risk assessment*;
- *adequate documentation* of the methods provided.

VEGA QSAR Models

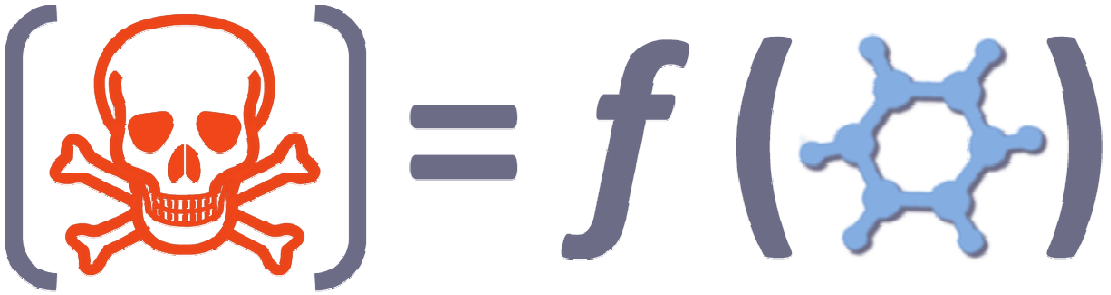
VEGA QSAR Models

OUR VISION

- To increase the documentation about the QSAR results
- To provide guidance on the reliable use of QSAR
- To make explicit when results are questionable
- To combine the experience of the models developers, of the model expert users, and of the stakeholders
- The establish a network about the improvement of the QSAR acceptance
- To incorporate stakeholder needs

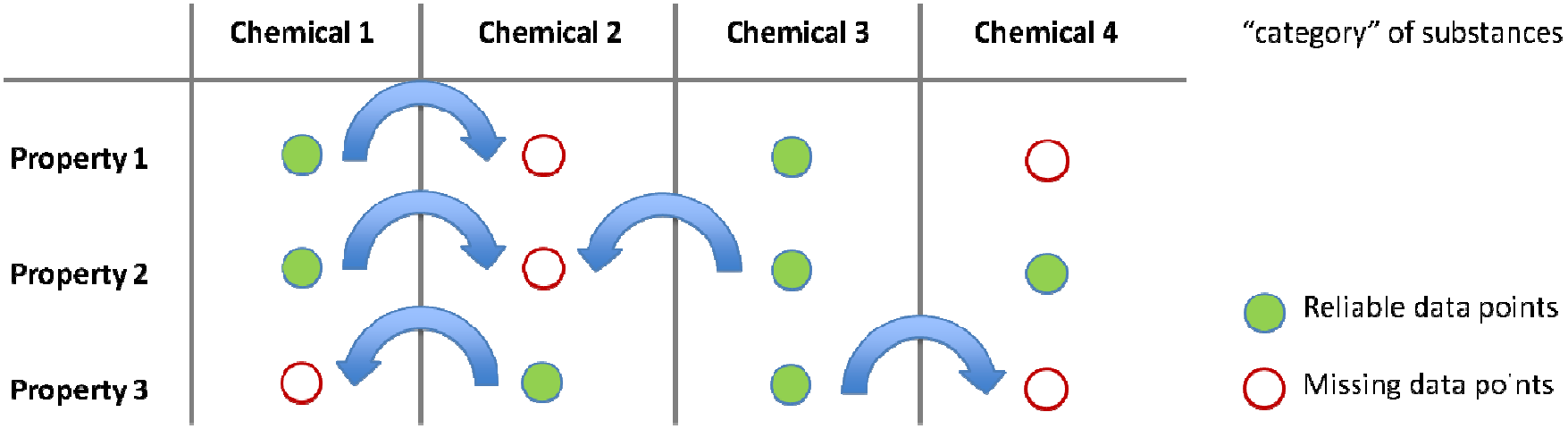
(Q)SAR

(QUANTITATIVE) STRUCTURE-ACTIVITY RELATIONSHIP



(Q)SAR SIMILAR TO READ-ACROSS

VEGA COMBINES THEM



VEGA STRATEGY

- VEGA combines QSAR and read across
- QSAR and read across are based on independent software
- VEGA automatically evaluates the prediction reliability
- Effort to make objective some evidences
- VEGA assists the human expert
- VEGA = collaboration between computer and expert
- The user should always use its/her experience
- Expert can override QSAR using read across

THE VEGA OUTPUT

- *Summary: value and reliability*
- *Uncertainty and possible uses*
- *The applicability domain: visualisation and score*
- *The documentation and reasoning: specific features and general parameters*
- *The references*

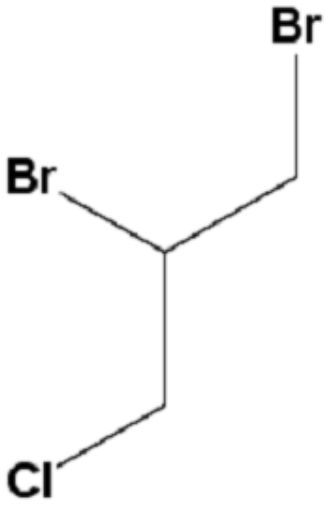



THE APPLICABILITY DOMAIN INDEX

The different checks done for the Applicability Domain

- **Visualisation of similar substances**
- **Similarity index** (chemical; sub-indices)
- **Chemiometric check** (descriptor space)
- **Atom centred-fragment** (chemical)
- **Check of the descriptor sensitivity** (algorithm)
- **Uncertainty** (algorithm)
- **Fragments for outliers** (output space)
- **Prediction Accuracy** (output space)
- **Prediction Concordance** (tox exploration)

THE APPLICABILITY DOMAIN INDEX



Compound: 138
Compound SMILES: C(C(CBr)Br)Cl
Prediction: 1.649 [log units]
Prediction: 45 [L/Kg]
Prediction from model 1 (HM): 1.754 [lo
Prediction from model 2 (GA): 1.614 [lo
Structural Alerts: -
Calculated LogP: 2.957 [log units]
Experimental value: -
Reliability: Compound could be out of n
Remarks for the prediction:

Global AD Index
AD Index = 0.7
Explanation: predicted substance could be out of the Applicability Domain of the model.

Similar molecules with known experimental value
Similarity index = 0.753
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 = 0.295
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 = 1.025
Explanation: similar molecules found in the training set have experimental values that completely disagree with the target compound predicted value.

Maximum error of prediction among similar molecules
Max error index = 0.33
Explanation: the maximum error in prediction of similar molecules found in the training set has a low value.

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.

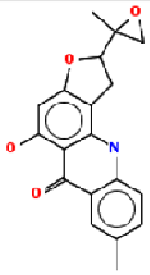
Descriptors noise sensitivity analysis
Noise Sensitivity = 0.922
Explanation: predictions has a good response to noise scrambling, thus shows a good reliability.

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.

VEGA

ADI : SIMILARITY SEARCH

Prediction for the compound no. 1: Cc1ccc2Nc3c4CC(Oc4cc(O)c3C(=O)c2c1)C1(C)CO1

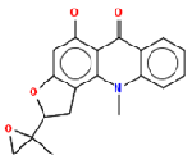


Activity: Mutagen

Remarks for the prediction:

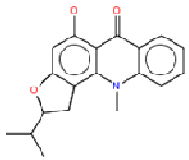
VISUALIZATION OF SIMILAR SUBSTANCES

The following chemicals similar to the query compound have been identified in the CAESAR database:



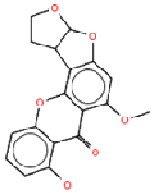
Dataset id: 10
SMILES: O=C5c1ccccc1N(c3c5(c(O)cc2OC(Cc23)C4(OC4)(C)))C
Similarity: 0.99

Experimental class: Mutagen
Predicted class: Mutagen



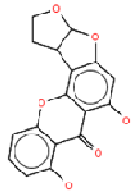
Dataset id: 772
SMILES: O=C2c1ccccc1N(c4c2c(O)cc3OC(C(=C)C)Cc34)C
Similarity: 0.922

Experimental class: Mutagen
Predicted class: Mutagen



Dataset id: 1963
SMILES: O=C1c5c(O)cccc5(Oc3c1c(OC)cc2OC4OCCC4(c23))
Similarity: 0.828

Experimental class: NON-Mutagen
Predicted class: NON-Mutagen



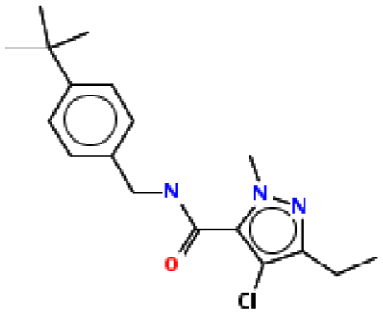
Dataset id: 3769
SMILES: O=C4c5c(O)cccc5(Oc2c4(c(O)cc1OC3OCCC3(c12)))
Similarity: 0.825

Experimental class: NON-Mutagen
Predicted class: NON-Mutagen

ADI : UNCERTAINTY

CAESAR QSAR model for Carcinogenicity - version 1.0

Prediction for the compound no. 1: CCC1=NN(C)C(C(=O)NCC2=CC=C(C=C2)C(C)(C)C)=C1Cl



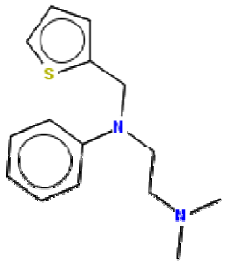
Carcinogenic: Non-Positive

Class indices:

Remarks for the prediction:

Borderline output, the predicted carcinogenicity probability values show a **VERY HIGH UNCERTAINTY IN THE PREDICTION**

The following chemicals similar to the query compound have been identified in the CAESAR database:



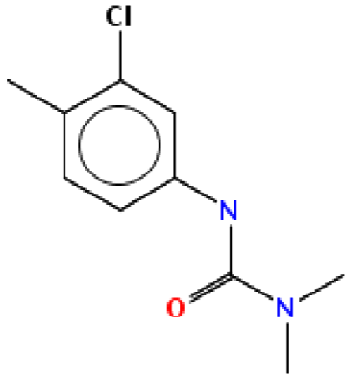
Dataset id: 433
SMILES: CN(CCN(c1ccccc1)Cc1cccs1)C
Similarity: 0.693

Experimental class: Non-Positive
Predicted class: Non-Positive

ADI : ACCURACY

PREDICTION FOR SIMILAR MOLECULES

Prediction for the compound no. 1: CN(C)C(=O)NC1=CC(Cl)=C(C)C=C1

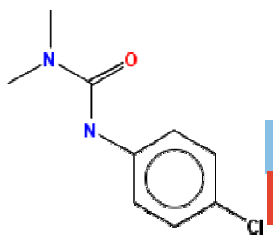


Carcinogenic: Non-Positive

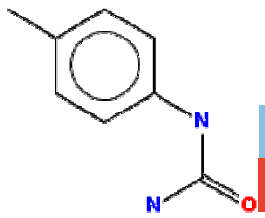
Class indices: Positive=0.079, Non-Positive=0.921

Remarks for the prediction:

The following chemicals similar to the query compound have been identified in the CAESAR database:



Dataset id: 162
SMILES: Clc1ccc(cc1)NC(=O)N(C)C
Similarity: 0.884



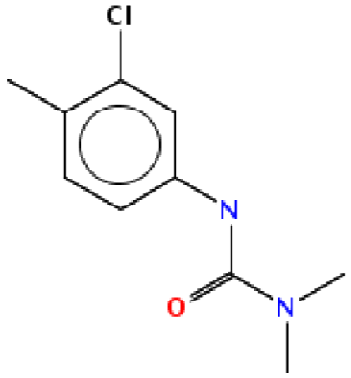
Dataset id: 754
SMILES: O=C(Nc1ccc(cc1)C)N
Similarity: 0.857

ADI : CONCORDANCE

WITH EXPERIMENTAL VALUE OF SIMILAR COMPOUNDS

CAESAR QSAR model for Carcinogenicity - version 1.0

Prediction for the compound no. 1: CN(C)C(=O)NC1=CC(Cl)=C(C)C=C1

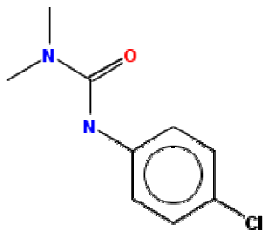


Carcinogenic: Non-Positive

Class indices: Positive=0.079, Non-Positive=0.921

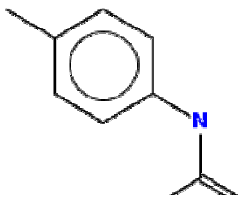
Remarks for the prediction:

The following chemicals similar to the query compound have been identified in the CAESAR database:



Dataset id: 162
SMILES: Clc1ccc(cc1)NC(=O)N(C)C
Similarity: 0.884

Experimental class: Positive
Predicted class: Non-Positive

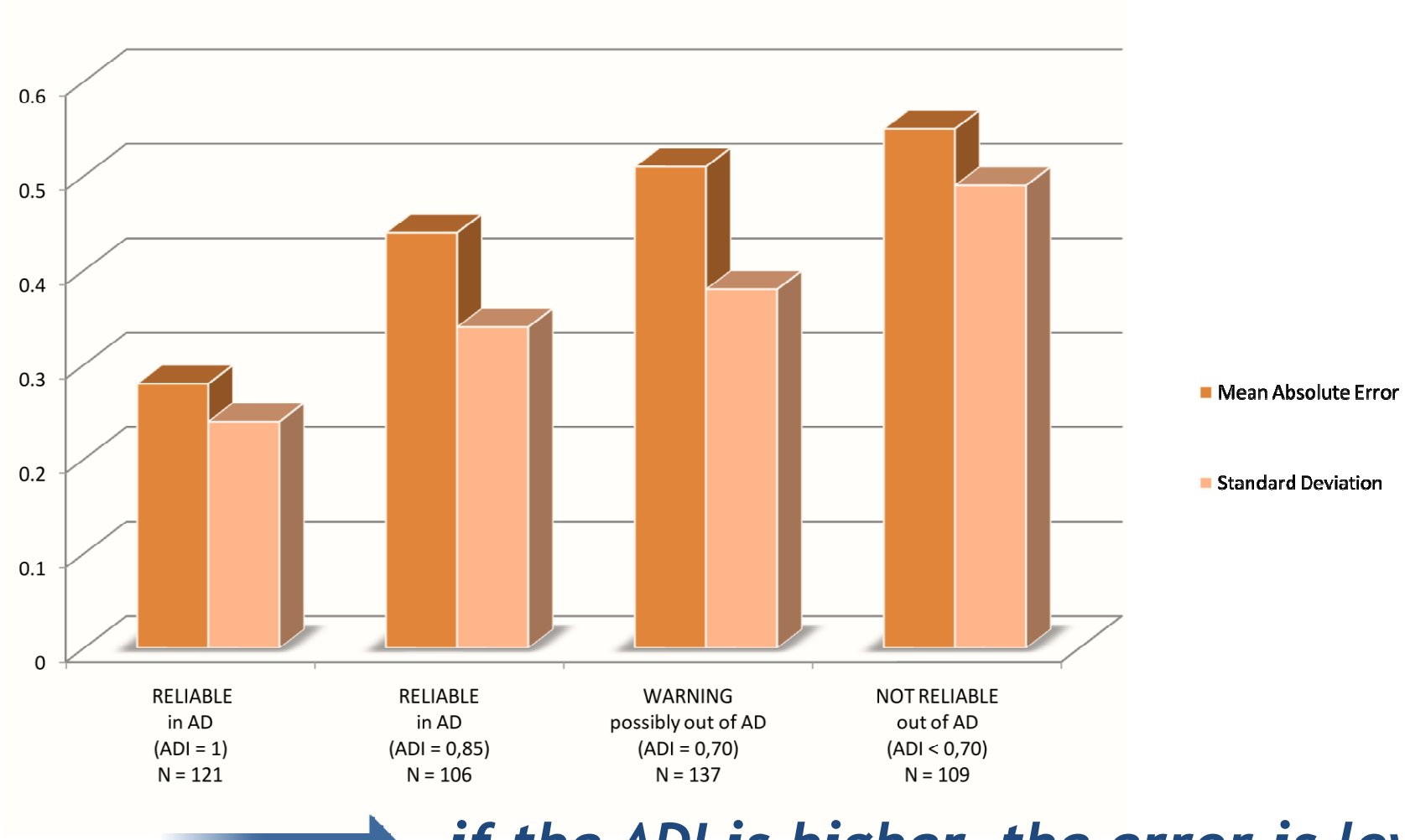


Dataset id: 754
SMILES: O=C(Nc1ccc(cc1)C)N
Similarity: 0.857

Experimental class: Non-Positive
Predicted class: Positive

VEGA

ADI : APPLIED TO BCF

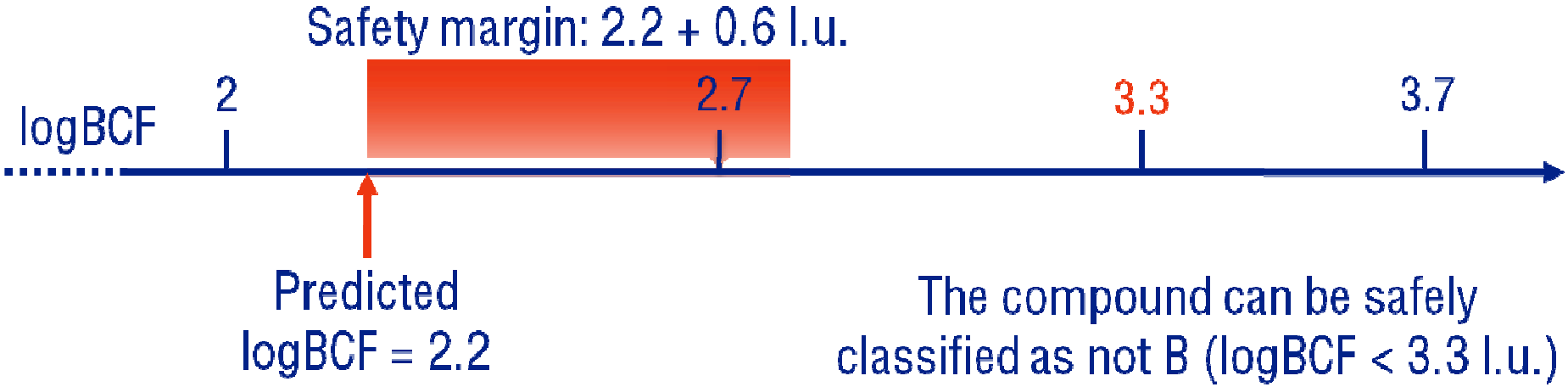


if the ADI is higher, the error is lower

ADI : ADEQUATE MODEL

Evaluation of the Bioaccumulation threshold (2000 = 3.3 l.u.)

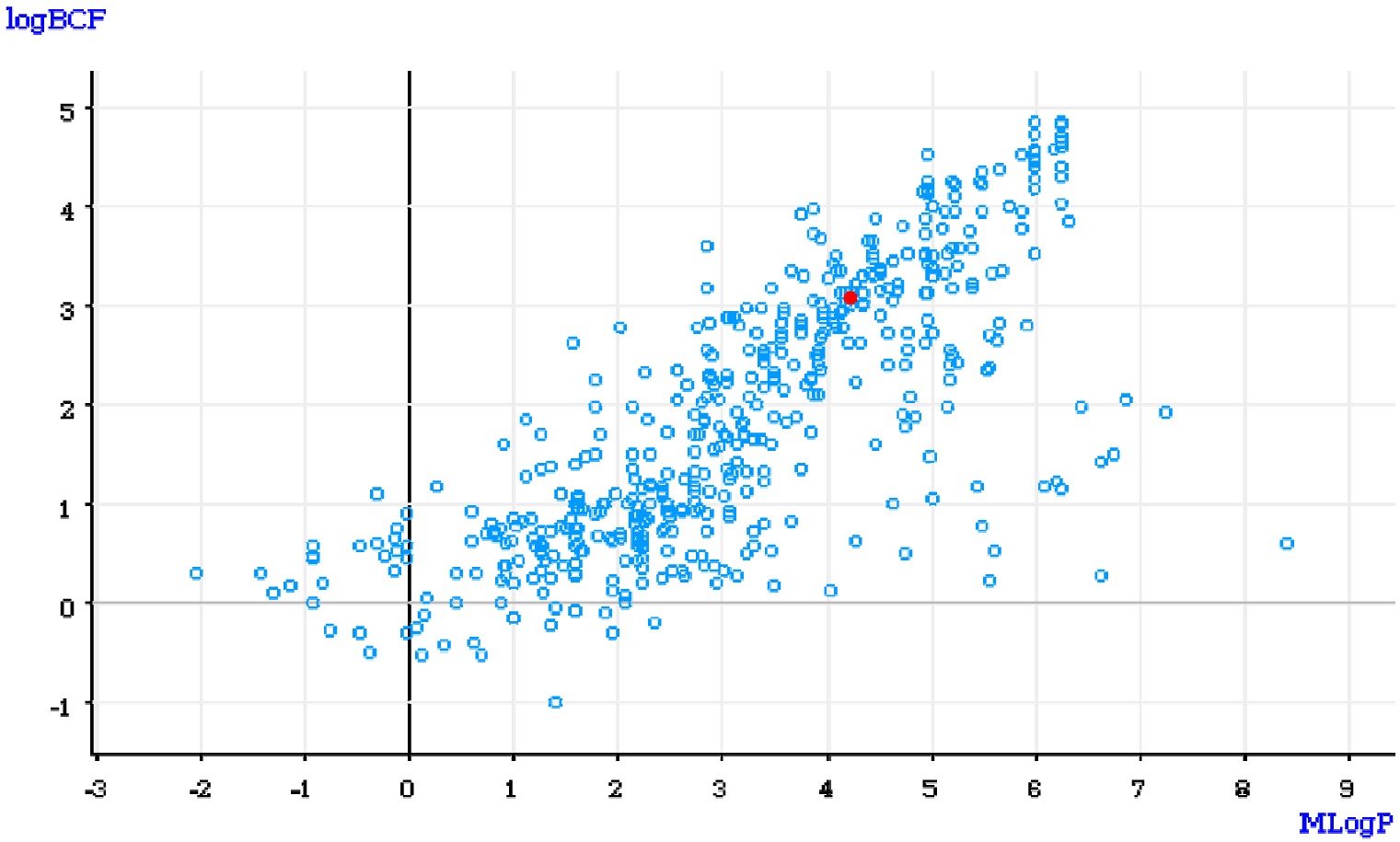
B threshold evaluation (3.3 l.u.)



No. Comp. = 492		Exp. logBCF	
		nB	B/vB
Pred. logBCF	nB	359	0
	B/vB	60	73

SUPPORTING DOCUMENTATION

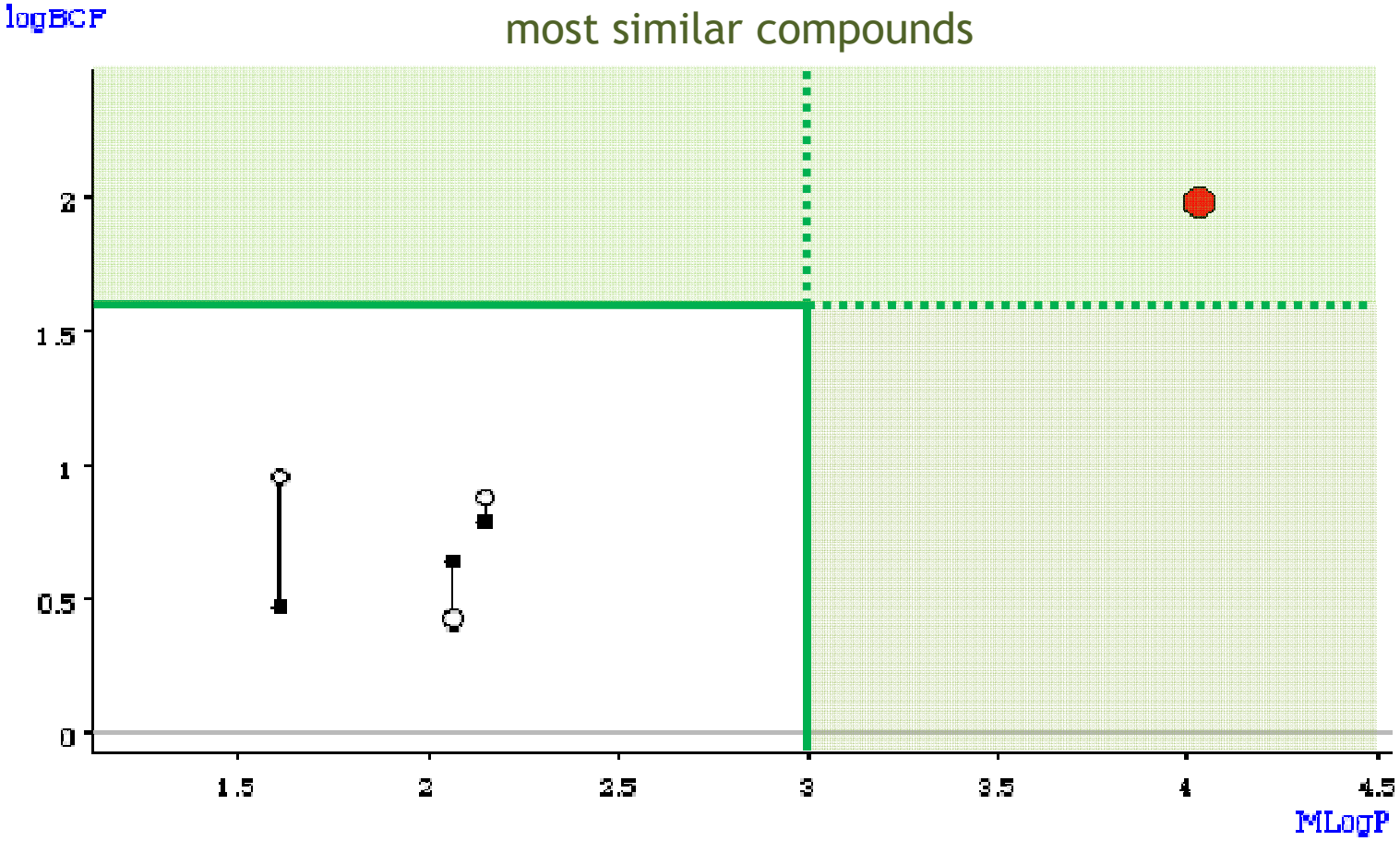
OVERVIEW OF THE DESCRIPTORS *logP and BCF*



SUPPORTING DOCUMENTATION

DETAIL ON SIMILAR COMPOUNDS

Difference between logP of the target and logP of the 3 most similar compounds → NO



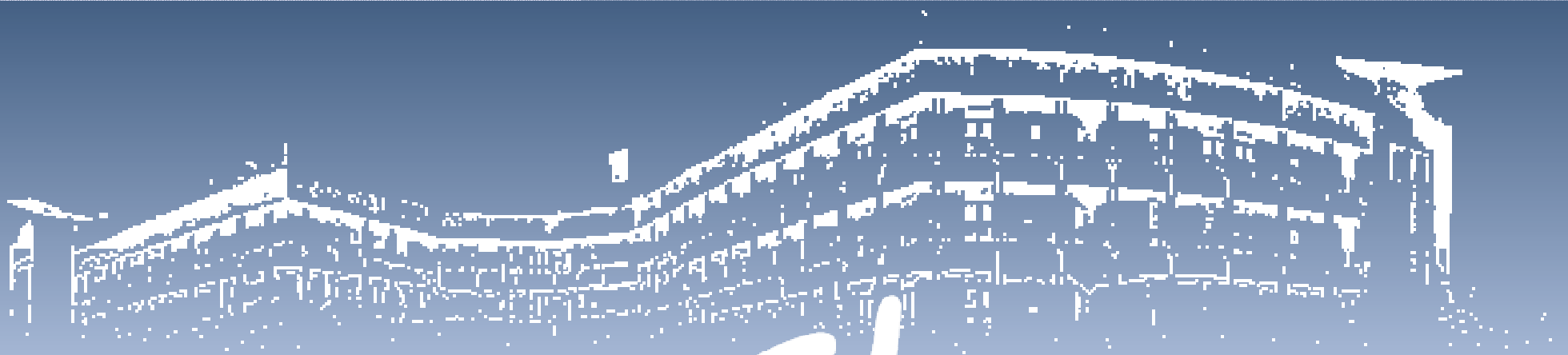
NO

↑

Difference between predicted logBCF and experimental logBCF of the 3 most similar compounds

Emilio BENFENATI

Istituto di Ricerche Farmacologiche Mario Negri
Laboratory of Environmental Chemistry & Toxicology



GRAZIE!

Emilio Benfenati

MARIO NEGRI
ISTITUTO DI RICERCHE FARMACOLOGICHE

ECT
Environment Chemistry Toxicology

ORCHESTRA

antares