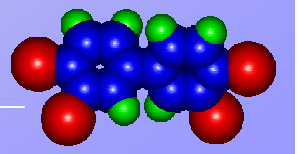
QSAR modelling of toxicity endpoints of emerging pollutants: Fragrances and Perfluorinated compounds

Barun Bhhatarai, Paola Gramatica, Mara Luini, Ester Papa

QSAR Research Unit in Environmental Chemistry and Ecotoxicology
DBSF -**University of Insubria, Varese – Italy**

e-mail: paola.gramatica@uninsubria.it
ester.papa@uninsubria.it
barun.bhhatarai@uninsubria.it
mara.luini@uninsubria.it

http://www.qsar.it



FP7- EU project CADASTER

CAse studies on the Development and Application of in-Silico Techniques for Environmental hazard and Risk assessment

Willie Peijnenburg, RIVM, The Netherlands	**Coordinator**
Mojca Durjava, Public Health Institute Maribor, Slovenia	**WP2 Leader** Database Experimental Parameters/(Q)SARs Chem/Biol.Endpoints Generation of new data
Paola Gramatica, University of Insubria (Varese), Italy	**WP3 Leader** Development and validation of QSARs
Thomas Öberg, University of Kalmar, Sweden	**WP4 Leader** Integration of QSARs with risk assessment
Igor Tetko, HMGU, Germany	**WP5 Leader** QSPR-THESAURUS: Web site and standalone tools
Andreas Woldegiorgis, IVL, Sweden	
Nina Jeliazkova, IDEA, Bulgaria	
Mike Comber, MCC. Belgium	
Mark Huijbregts, RUN, The Netherlands	

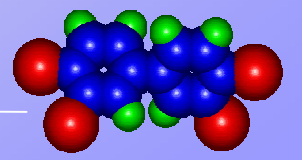
FP7- EU project CADASTER

4 classes of emerging pollutants studied:

Flame retardants, Fragrances, PFCs and (benzo) Triazoles (REACH)

WP3: QSAR model development and validation

- **DRAGON descriptors (from Hyperchem), selected by GA**
- **MLR models**
- **External Validation by *a priori* splitting of data (random and by SOM)**
- **Applicability Domain**



FRAGRANCES

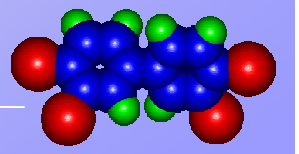
Mara Luini

Introduction

- Fragrances are used in a wide variety of consumer products such as creams, lotions, detergents, and various other **personal and household products**
- The low cost synthesis and increased resistance to light were the main reasons for their **extensive use**
- Human exposure to these agents is **widespread and often involuntary**
- Fragrances are believed to have possible **toxic effects** on humans
- Little is known about the environmental fate and toxicity

=> **their potential effects on humans and aquatic ecosystems are not yet clearly understood**

Need to use predictive QSAR approaches to fill this data gap and characterize the environmental and toxicological profiles of these compounds by minimizing animal tests



Health concerns

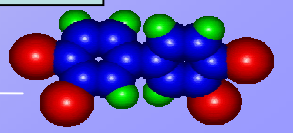
SKIN → fragrances have been recognized as skin allergens and irritants

RESPIRAT → fragrances can induce respiratory … as asthma, …blems, rhinitis…

NEUROLOG → …an impact …rvous system

SYSTEMIC → fragrances can enter the body through numerous routes and once inside the body can impact any organ or system

- Some fragrance materials have been found to accumulate in adipose tissue and are present in breast milk
- Others are suspected of being hormone disruptors



Environmental concerns

AIR

Fragrances are complex mixtures of volatile organic compounds (VOCs).

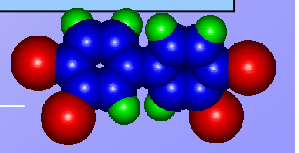
Once in the air they can break down and form new compounds.

WATER

A large portion of fragrances ends up in wastewater, but most wastewater treatment methods do not remove them so they end up in streams and rivers from sewage treatment plans.

WILDLIFE

Musk compounds tend to accumulate and break down slowly; they persist in the aquatic environment and accumulate in the fatty tissue of aquatic wildlife.



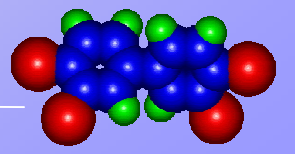
Data Sets

Toxicological properties

Dataset	N° of available exp-data (→ modelled)	Bibliography
Log1/LD50 Oral mouse	24 → 23	D.R.Bickers et al. 2002 D.Belsito et al. 2007 ChemIDPlus
LogEC50 NADH-Ossidase	20 → 18	D.E.Griffith et al. 2005
LogEC50 Δψm (effect on membrane potential)	20 → 15	D.E.Griffith et al. 2005

Inhibition of mithocondrial NADH Ossidase complex in rat cells liver

Inhibition of mithocondrial membrane potential in rat cells liver



Results: models for toxicological endpoints

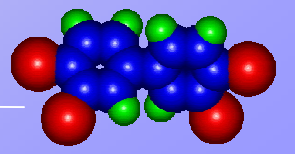
Toxicological Endpoints	Model	Train Obj.	Test Obj.	Variables	R^2	Q^2_{LOO}	Q^2_{BOOT}	Q^2_{EXT}	RMSE Train	RMSE Test	R^2-YScr
Log 1/LD50 Oral Mouse	Full Model	23	-	**nR=Cs H-047**	89	86.2	81	-	-	-	9.2
	Random	13	10		91.4	86.2	80.5	73.4	0.265	0.226	-
	SOM	15	8		88.8	83.4	72.2	90.2	0.256	0.186	-
LogEC50 ΔΨm	Full Model	15	-	ATS4v MATS2m	91.7	88.6	83.9	-	-	-	14.3
	Random	12	3		90.3	84.3	74.0	98.2	0.200	0.061	-
	SOM	13	2		90.9	85.8	78.8	97.3	0.193	0.070	-
LogEC50 NADH-Ox	Full Model	20	-	nC R5u+	85.8	82.4	76.6	-	-	-	10.5
	Random	16	4		84.4	79.7	73.3	89.1	0.321	0.320	-
	SOM	16	4		86.5	81.8	77.8	80.8	0.292	0.440	-

Focus on the following aspects of interest:

- **VALIDATION**
- **APPLICABILITY DOMAIN**

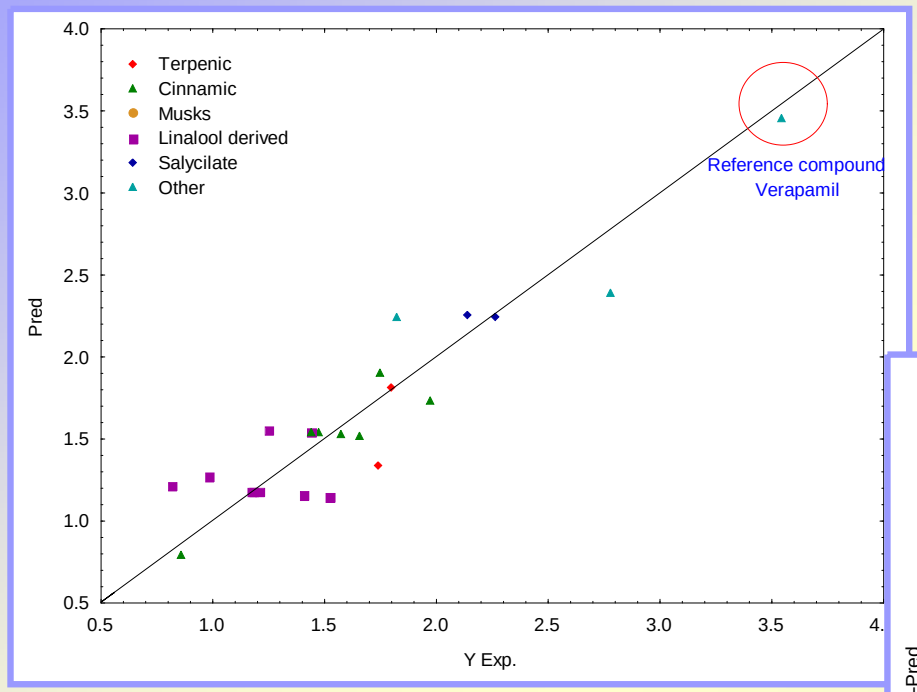
P. Gramatica, Principles of QSAR models validation: internal and external

***QSAR Comb.Sci.* 2007, 26(5), 694-701**



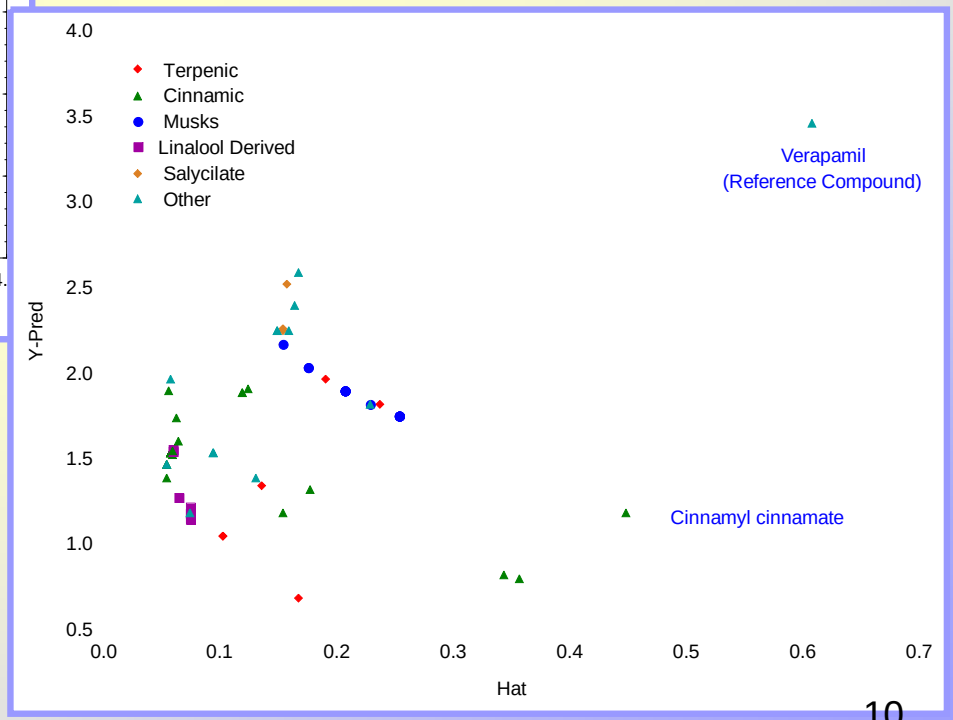
Model for Log1/LD50 oral mouse

Log1/LD50 Oral Mouse = 1.746 + 0.0705 H-047 – 0.4247 nR=Cs



Variables	R^2 %	Q^2 %	Q^2_{boot} %
nR=Cs H-047	89.0	86.2	81.0

Applicability Domain



nR=Cs : it is among the functional group counts. It corresponds to number of aliphatic secondary C (sp2).

H-047: it is among the atom-centred fragments. It corresponds to H attached to C1(sp3)/C0 (sp2), linked (1) or not (0) to heteroatoms.

Conclusions on Fragrances

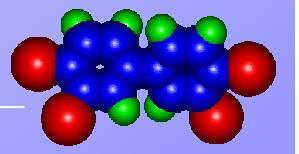
- **Limited availability of experimental data useful for QSAR (in particular SIDS endpoints for CADASTER project).**

•New QSAR and QSPR models have been developed for the prediction of 3 toxicological endpoints:

acute oral mouse toxicity, and 2 endpoints related to mitochondrial toxicity

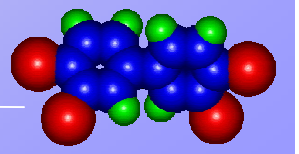
•Despite the limited amount of available data, all the models where carefully internally and externally validated.

At our knowledge, no other QSAR models are available in literature for these endpoints.



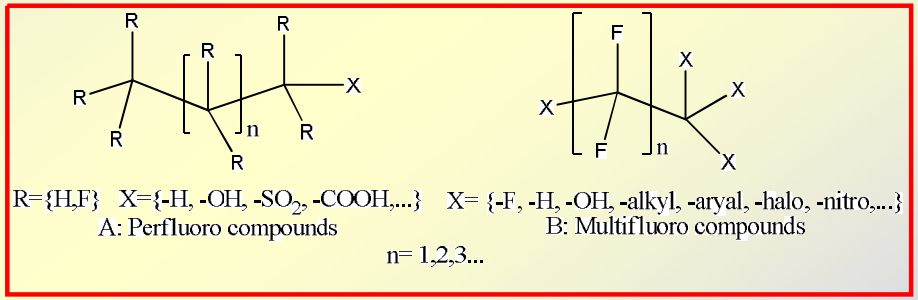
PERFLUORINATED COMPOUNDS

Barun Bhhatarai, PhD



Introduction

- Perfluorinated compounds (PFCs) are chemicals containing a long fluorinated carbon tail attached to different functional groups
- PFCs as perfluoro-octanesulfonate (PFOS), perfluoro-octanoate (PFOA) and perfluoro- octane sulfonylamide (PFOSA) are stable chemicals with a wide range of industrial and consumer applications [Inoue 2004]
- Degradable products of commercial PFCs are found in environment and biota and diPAPs (a group of PFCs used on food wrappers) was recently reported in human blood [*Renner 2009*]
- PFCs are considered emerging pollutants and are believed to have potential toxic effects in humans and wildlife
- PFCs along with Polyfluoro compounds are studied for LC_{50} inhalation toxicity of Mouse and Rat



R={H,F} X={-H, -OH, -SO_2, -COOH,...} X= {-F, -H, -OH, -alkyl, -aryal, -halo, -nitro,...}

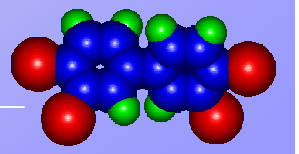
A: Perfluoro compounds B: Multifluoro compounds

n= 1,2,3...

Predictive QSAR approaches is used to fill the data gap and to predict toxicity of 250 PFCs on two different species viz. Mouse and Rat

Results: QSAR models for LC_{50} inhalation

	Splitting	Compounds	Variables selected	R^2 (%)	Q^2_{LOO}	Q^2_{BOOT}	Q^2 ext	R^2-Y_{Scrm}
Mouse Inhalation **56 compounds**	SOM 28.5%	Train: 40 Test: 16	**X3v; H-048; MLOGP; F01[C-C]**	***83.0***	***78.1***	***75.5***	***71.6***	***10.3***
	Random by Activity 20%	Train: 44 Test: 12		***77.1***	***71.7***	***69.9***	***85.1***	***9.0***
	Full model			**79.8**	**76.3**	**75.4**	**-**	***7.0***
Rat Inhalation **52 compounds**	SOM 18.87%	Train: 42 Test: 10	**Jhetv: PCR; ALOGP; B02[Cl-Cl]**	***79.4***	***73.9***	***71.9***	***72.5***	***9.6***
	Activity 20%	Train: 42 Test: 10		***79.8***	***74.7***	***73.4***	***70.3***	***10.6***
	Full model			**78.5**	**74.2**	**73.3**	**-**	***7.6***

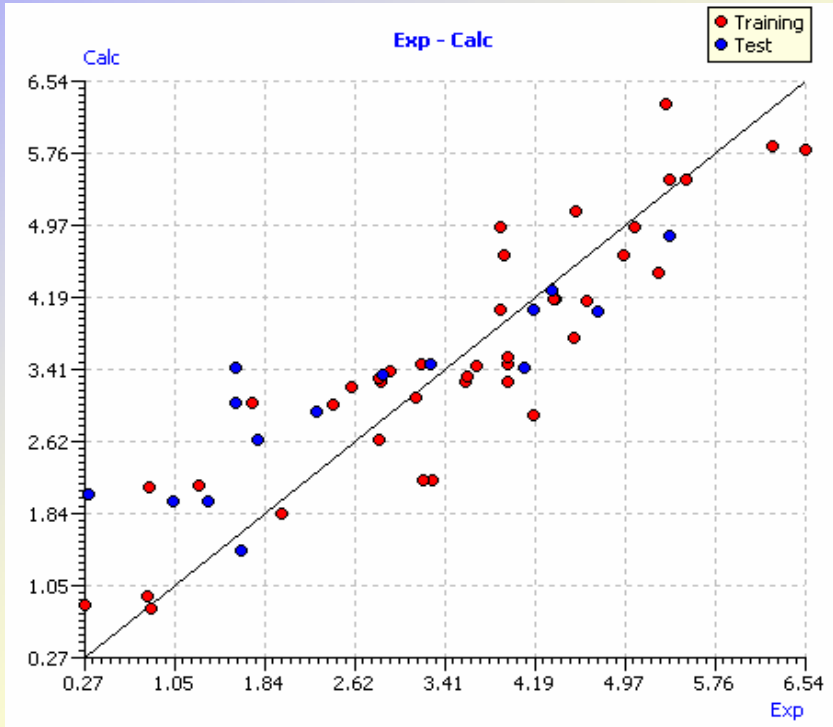


Regression plots for the models on datasets split by SOM

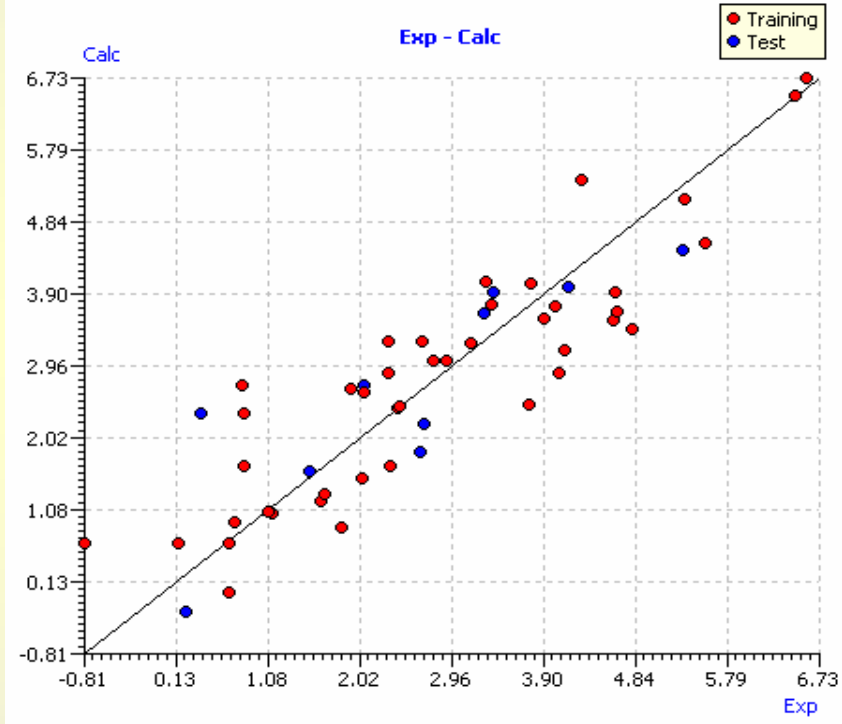
Mouse

$\log 1/LC_{50}$ = 4.21 – 1.27 (±0.31) MlogP + 1.43 (±0.46) X3v + 0.38 (±0.13) F01[C-C] – 1.14 (±0.37) H-048

n=56, s=0.72, r^2=79.83, F=50.5, Kx=42.34, Kxy=50.40



Mouse

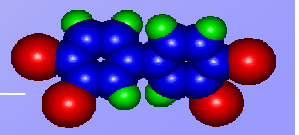


Rat

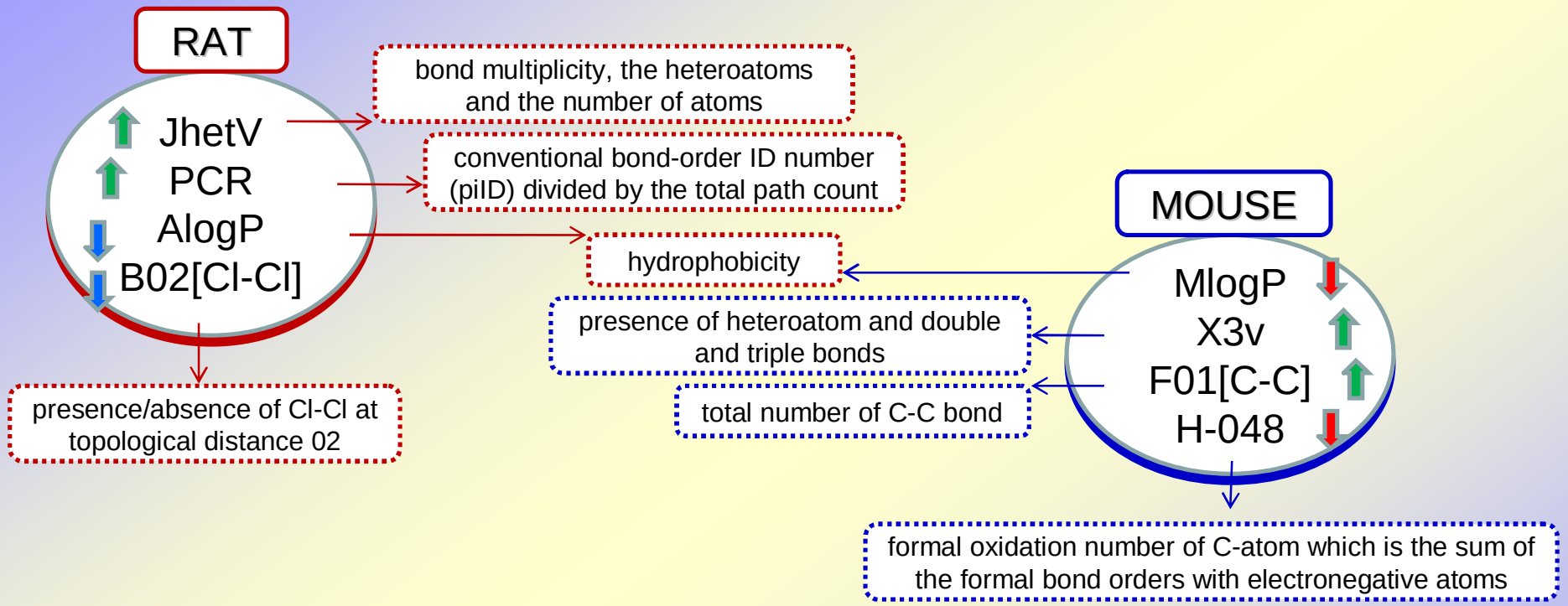
Rat

$\log 1/LC_{50}$ = – 11.14 + 2.09 (±0.43) Jhetv + 9.57 (±2.31) PCR – 0.66 (±0.26) AlogP – 1.58 (±0.80) B02[Cl-Cl]

n=52, s=0.82, r^2=78.53, F=42.98, Kx=25.36, Kxy=34.92



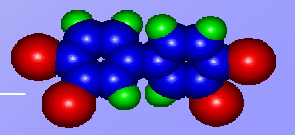
Descriptor analysis



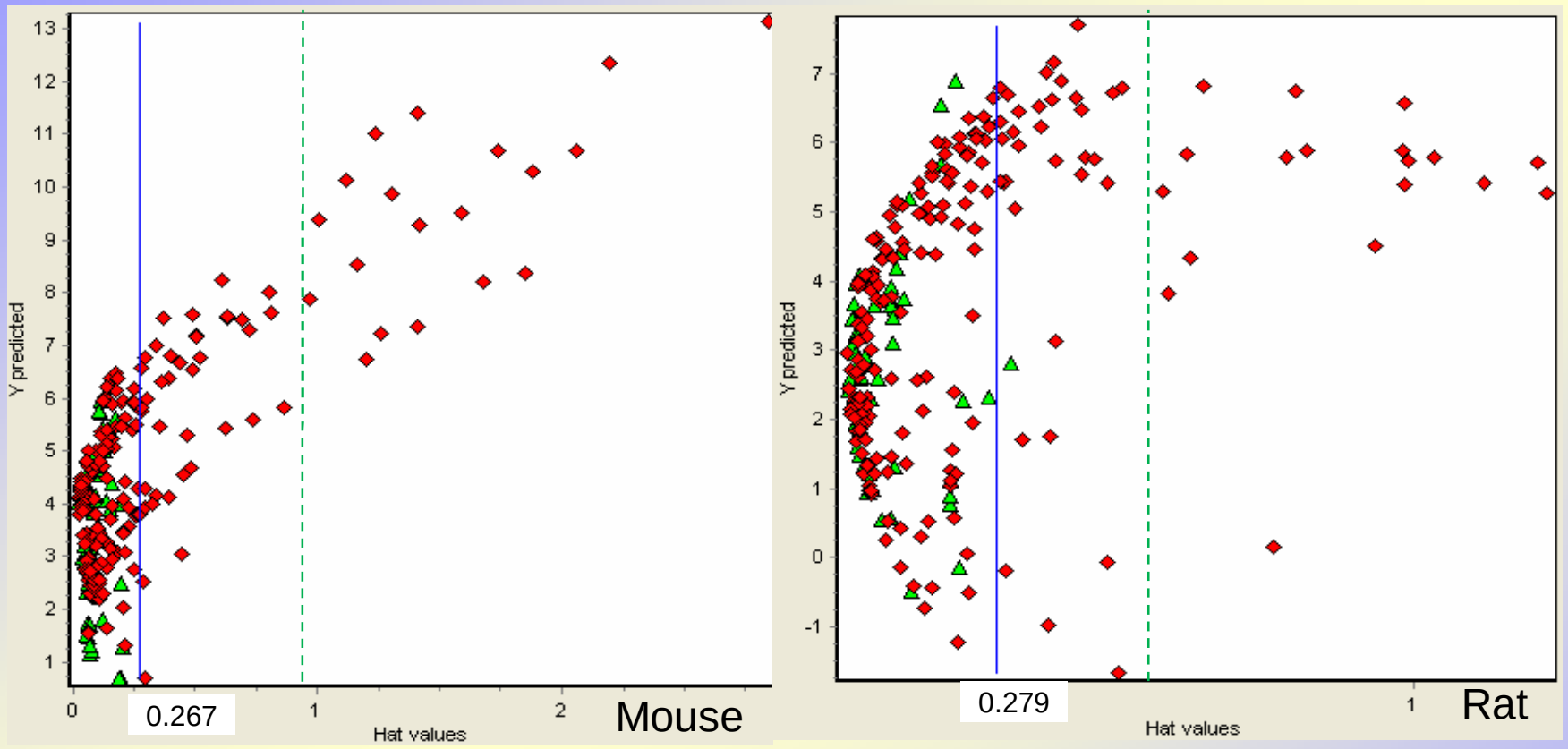
- **Common descriptor characterizing Hydrophobicity was negative for both species**
  MlogP vs AlogP = 0.847

- **JhetV and X3v have similar chemical meanings and are positive for both species**
  JhetV vs X3v r= 0.780

•B02[Cl-Cl] present for 5 of 52 compounds –
fitting (?) descriptor to include all Freons

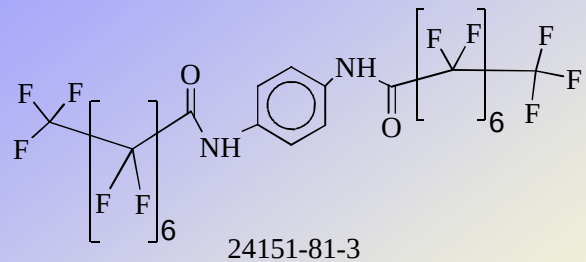


Applicability Domain (AD) study on 250 PFCs

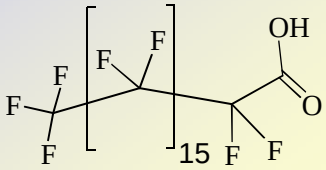


- **61 compounds are out of domain in Mouse model (75.6% coverage of PFCs) and 53 in Rat model (78.8% coverage).**
- **Arbitrary cutoff at 1.0 for Mouse and 0.5 for Rat (green lines): 11 common compounds.**

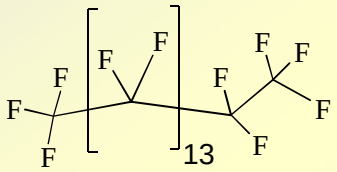
Focus on AD: Common Out-of-domain compounds



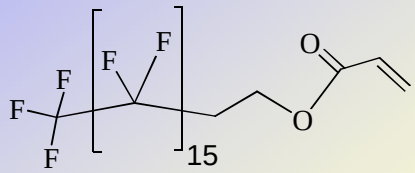
24151-81-3



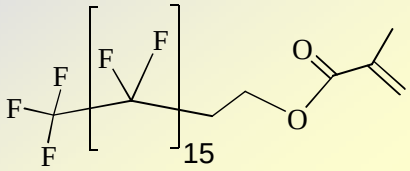
16517-11-6



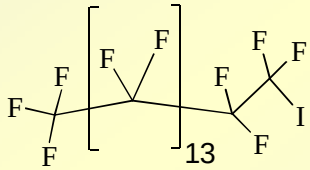
355-49-7



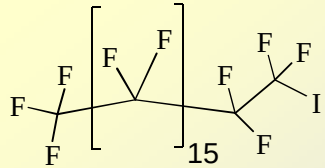
65150-93-8



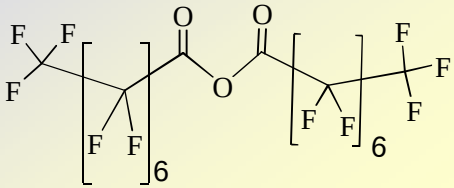
59778-97-1



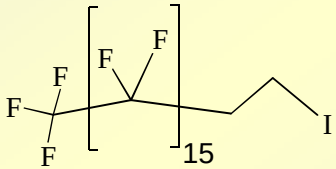
355-50-0



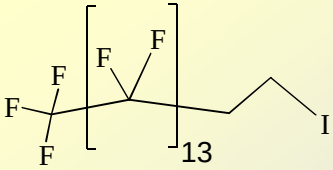
29809-35-6



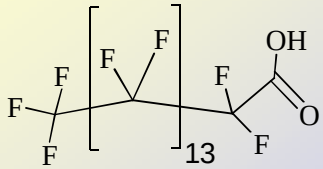
33496-48-9



65150-94-9

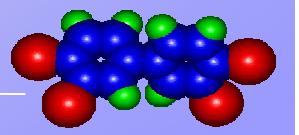


65510-55-6

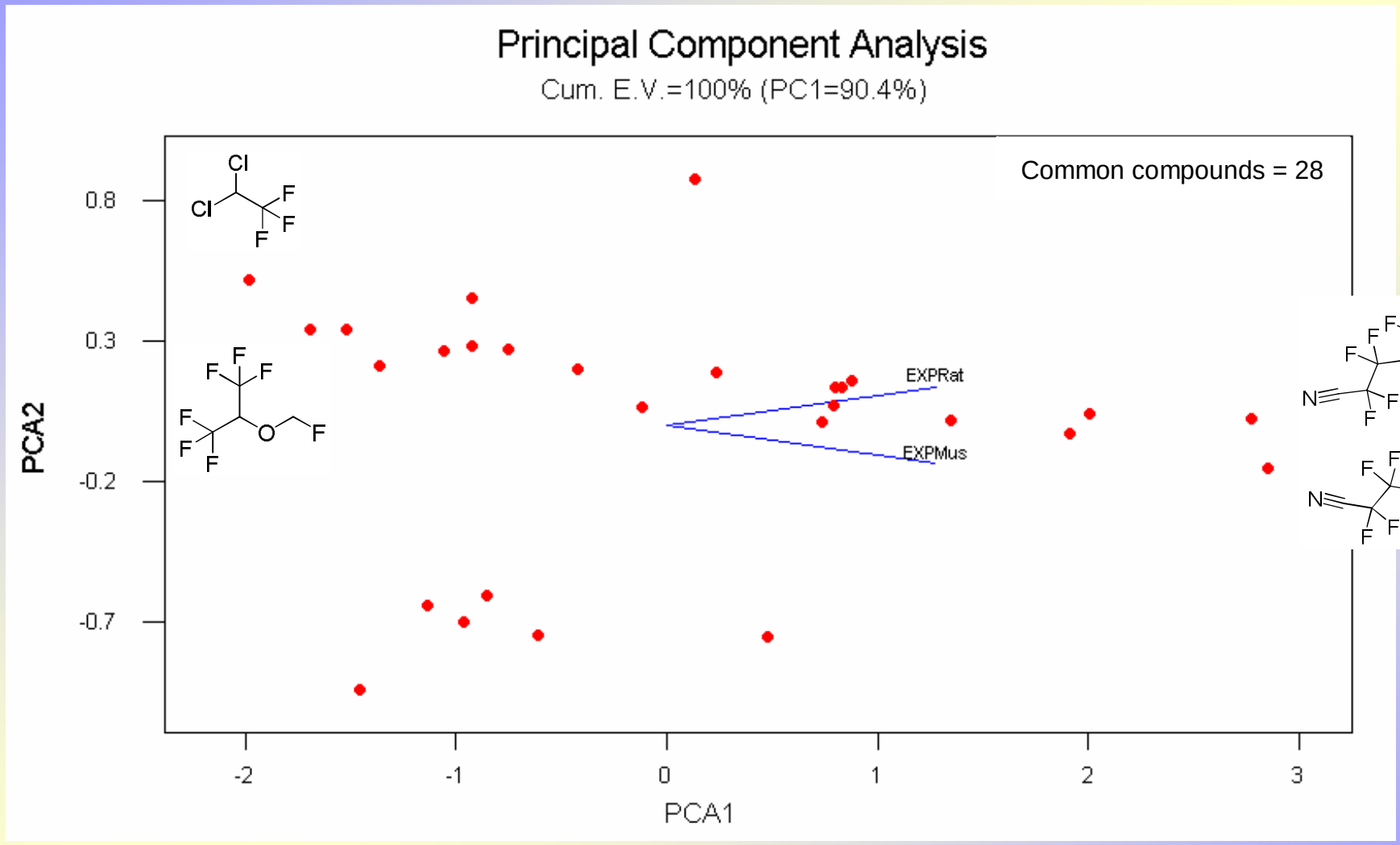


67905-19-5

- **Predicted compounds out of applicability domain of both Mouse and Rat model are long chain PFCs (>15-Carbon)**
- **They are probably extrapolated as the longest compounds in the training sets are with 7-Carbon**

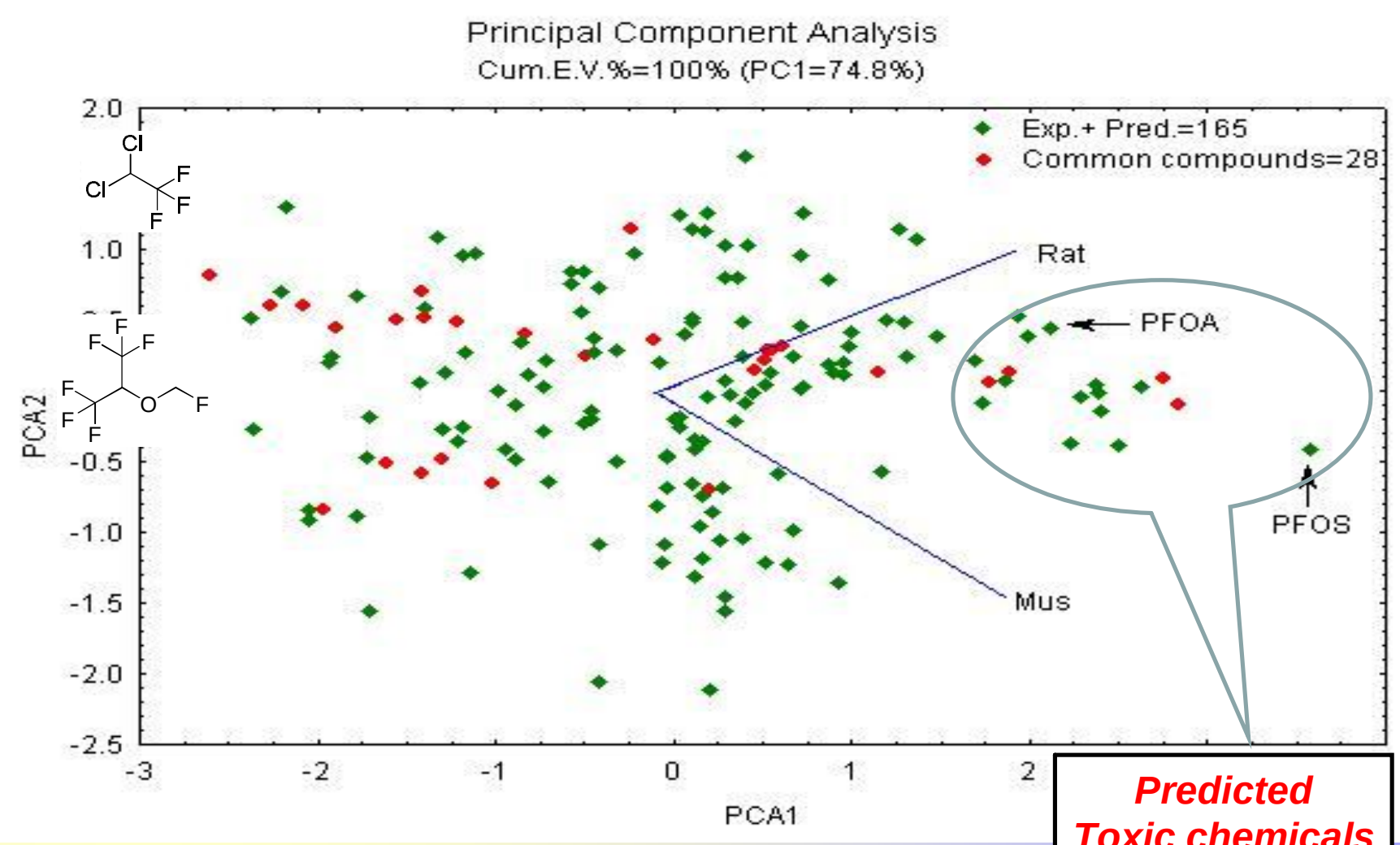


Toxicity Trend



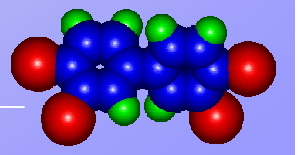
Increasing Toxicity

Toxicity Trend

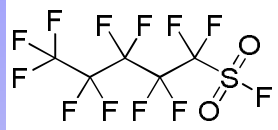


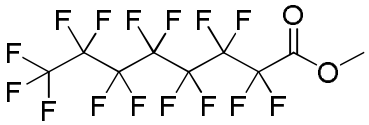
Predicted Toxic chemicals

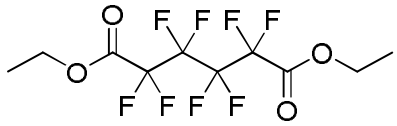
Increasing Toxicity

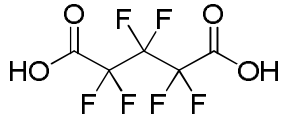


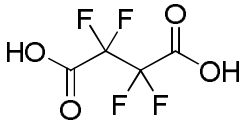
Toxic Chemicals Predicted: by PCA analysis

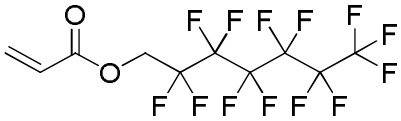

375-81-5

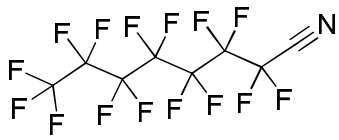

376-27-2

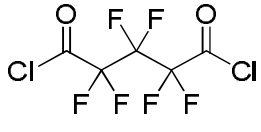

376-50-1

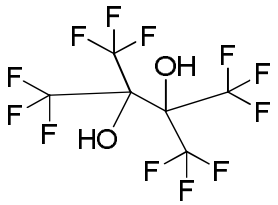

376-73-8

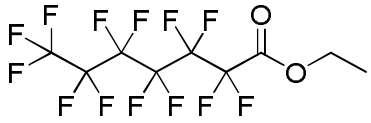

377-38-8

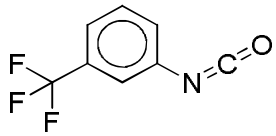

559-11-5

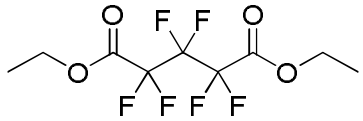

647-12-1

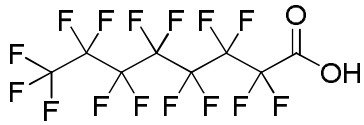

678-77-3

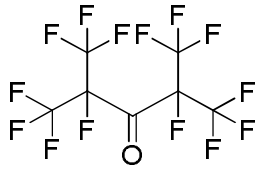

918-21-8

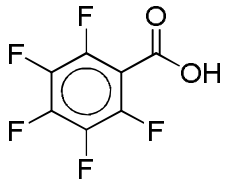

41430-70-0

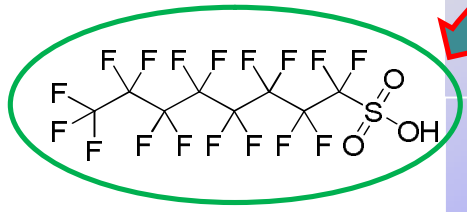

329-01-1


424-40-8


335-67-1


813-44-5

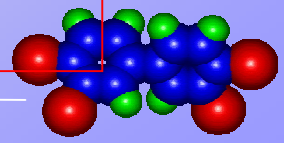

602-94-8


1763-23-1

PFOS – slightly out of AD but predicted Toxic

PFOA is under investigation as toxic

These chemicals have been suggested to the Partners for experimental tests



This study has been financially supported by the FP7th-EU Project CADASTER (n. 212668)

Thanks for your attention

http://www.qsar.it