Perfluoro Toxicity (LC_{50} inhalation) in Rattus and Mus species using QSAR

Barun Bhhatarai, Paola Gramatica

QSAR Research Unit in Environmental Chemistry and Ecotoxicology, DBSF, University of Insubria, Varese, Italy. e-mail: paola.gramatica@uninsubria.it

Introduction

- PFOS and PFOA are a class of perfluoro alkylated compounds which are categorized by US-EPA and EU-REACH as toxic chemicals.
- Other perfluoro alkylated compounds are widely distributed in environments, few are classified as 'emerging pollutants'
- Few experimental data on environmental and bio-toxicity are available
- Toxicity profiles are found different for types of animals and species used.
- This class of chemicals is studied in the European FP7th Project CADASTER.
- QSAR is applied to understand the inter-species toxicity of short and long chain perfluoro alkylated compounds by modeling inhalation (LC_{50}) data on *Rattus* (Rat) and *Mus* (Mouse).

Data Set

ChemIDplus Advance database: *Mus* 56 and *Rattus* 52 compounds used.

❖ Training and Prediction set were prepared by splitting using: Self Organizing Map and random selection through activity sampling [6]. 250 congeneric per- fluorinated chemicals with no toxicity data were predicted.

MOLECULAR DESCRIPTORS:

❖ 0D to 3D molecular descriptors including Quantum-chemical were calculated by DRAGON [3] XYZ coordinate and HYPERCHEM package [4] respectively.

QSAR Modeling and Validation

❖ GA based Variable Subset Selection method for descriptor selection [5]. *Ordinary Least Squares* regression (OLS) based Mutiple Linear Regression (MLR) models.

❖ Prediction set was selected based on a) the structure and b) the responseModels were developed by following OECD principles for QSAR validation [7]. Applicability Domain (AD for 250 PFCs) was verified by leverage approach.

❖ Internal (Q^2_{LOO}, Q^2_{BOOT}, R^2Y_{scram}) and external validation (Q^2ext-F1, Q^2ext F3, R^2ext) were checked.

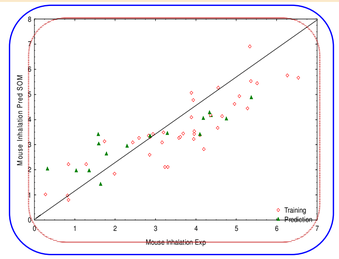
❖ Residuals and Williams plot to check the quality of best models and to find compounds which are out of AD

❖ Mechanistic explanation based on the selected descriptors.

Results and Discussions

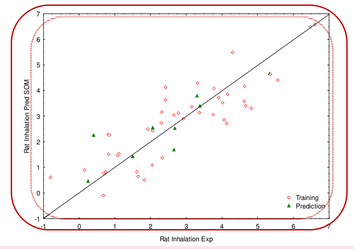
log $1/LC_{50}$ = –12.76 + 1.87 (±0.20) Jhetv + 11.43 (±1.27) PCR – 0.60 (±0.12) MlogP – 1.41 (±0.40) B02[Cl-Cl]
n=52, s=0.82, r^2=78.14, F=41.99, Kx=23.55, Kxy=30.86, R^2_{Yscram}=7.64

Response (Compounds)	Descriptors: Input	Descriptors: Selected	Splitting criteria	Compounds	R^2	Q^2_{LOO}	Q^2_{BOOT}	Q^2 ext F1	Q^2 ext F3	$RMSE_{Train}$	$RMSE_{EXT}$	R^2_{Yscram}
Mouse Inhalation LC_{50} (56)	537	X3v: H-048: MlogP: F01[C-C]	SOM 28.5%	Train: 40, Predict: 16	***82.99***	***78.09***	***75.46***	***71.62***	***64.46***	***0.61***	***0.69***	***10.32***
			Random by Activity 20%	Train: 44, Predict: 12	***77.07***	***71.73***	***69.89***	***85.11***	***85.11***	***0.70***	***0.78***	***8.99***
			Full model without splitting		***79.83***	***76.31***	***75.38***	-	-	***0.68***	***0.74 (cv)***	***7.05***

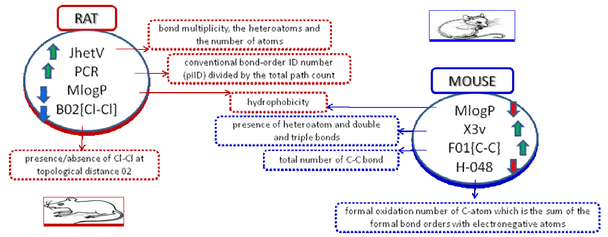


log $1/LC_{50}$ = 4.21 – 1.27 (±0.31) MlogP + 1.43 (±0.46) X3v + 0.38 (±0.13) F01[C-C] – 1.14 (±0.37) H-048
n=56, s=0.717, K_x=42.34, K_{xy}=50.40, $RMSE_{train}$=0.68, R^2(Yscram)=0.394,

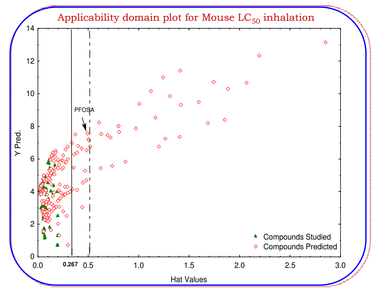
Response (Compounds)	Descriptors: Input	Descriptors: Selected	Splitting criteria	Compounds	R^2	Q^2_{LOO}	Q^2_{BOOT}	Q^2 ext F1	Q^2 ext F3	$RMSE_{Train}$	$RMSE_{EXT}$	R^2_{Yscram}
Rat Inhalation LC_{50} (52)	360	Jhetv: PCR: MlogP: B02[Cl-Cl]	SOM 18.87%	Train: 43 Test: 10	***78.36***	***72.99***	***71.95***	***75.47***	***79.67***	***0.80***	***0.89***	***8.75***
			Random by Activity 20%	Train: 43 Test: 10	***80.01***	***75.21***	***74.12***	***66.70***	***75.41***	***0.77***	***0.86***	***9.91***
			Full model without splitting		***78.12***	***73.83***	***73.26***	-	-	***0.79***	***0.86(cv)***	***7.64***



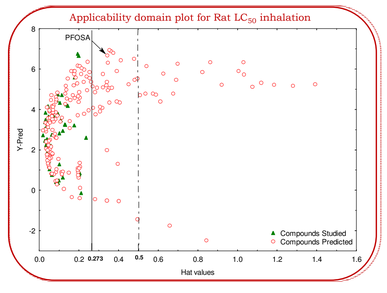
Descriptor Analysis, Mechanistic Interpretation and Applicability Domain



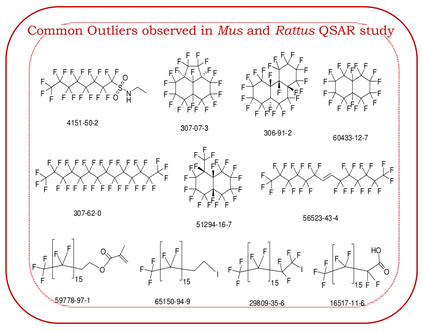
Applicability domain plot for Mouse LC_{50} inhalation



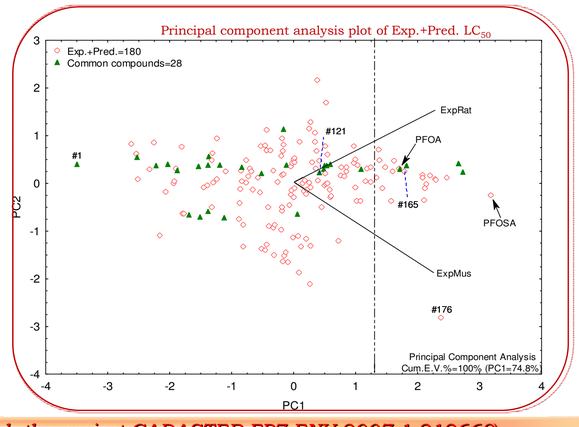
Applicability domain plot for Rat LC_{50} inhalation

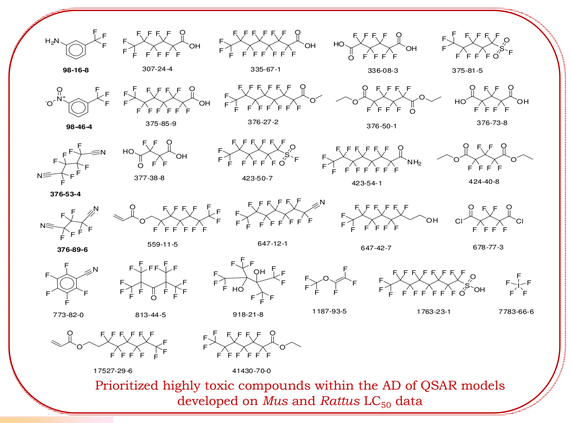


Common Outliers observed in *Mus* and *Rattus* QSAR study



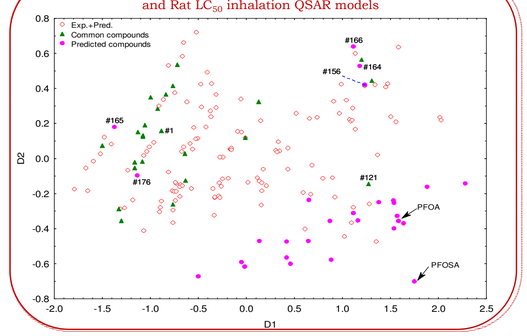
Principal component analysis plot of Exp.+Pred. LC_{50}





Prioritized highly toxic compounds within the AD of QSAR models developed on *Mus* and *Rattus* LC_{50} data

MDS plot on 7 molecular descriptors which appeared in Mouse and Rat LC_{50} inhalation QSAR models



References

1) Hekster F.M. et al. *Rev. Environ. Contam. Toxicol.* **2003**, 179, 99-121.
2) Renner R. *Environ. Sci. Technol.* **2003**, *37*, 201A. 3
3) Todeschini R., Consonni V. and Pavan M., 2001. *DRAGON – Software for the calculation of molecular descriptors*, rel. 1.12 for Windows. Free download available at http://www.disat.unimib/chm;
4) HYPERCHEM. Release 7.03 for Windows, 2002. Hypercube, Inc. Florida, USA.
5) Todeschini, R., 2001. *Moby Digs - Software for multilinear regression analysis and variable subset selection by Genetic Algorithm*, rel. 2.3 for Windows, Talete srl, Milan (Italy).
6) Gramatica, P. QSAR Comb. Sci. 2007, 26, 694-701.
7) OECD Guidelines, http://www.oecd.org/dataoecd/33/37/37849783.pdf.
8) Bhhatarai, B., Gramatica, P. Per- and Poly-fluoro Toxicity (LC_{50} inhalation) Study in Rat and Mouse using QSAR Modeling, *Chem. Res. Tox.* (accepted)

(Financial support by European Union through the project CADASTER FP7-ENV-2007-1-212668)