QSAR and QSPR models for emerging pollutants: WP3 activities within the FP7 European Project CADASTER

S. Kovarich[1], B. Bhhatarai[1], E. Papa[1], M. Rahmberg[2], S. Nilsson[2], T. Liu[3], T. Öberg[3], N. Jeliazkova[4], N. Kochev[5], O. Pukalov[5], W. Teetz[6], S. Brandmaier[6], I.V. Tetko[6], P. Gramatica[1]

[1] QSAR Research Unit in Environmental Chemistry and Ecotoxicology, DBSF, University of Insubria (**UI**), Varese, Italy. [2] IVL Swedish Environmental Research Institute Ltd.(**IVL**) [3] School of Pure and Applied Natural Sciences, Linnaeus University (**LNU**), Kalmar, Sweden. [4] Ideaconsult Ltd. (**IDEA**), Sofia, Bulgaria. [5] Department of Analytical and Computer Chemistry, University of Plovdiv, Plovdiv, Bulgaria. [6] Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum Muenchen - German Research Center for Environmental Health (**HMGU**), Neuherberg, Germany. E-mail: paola.gramatica@uninsubria.it

The EU-REACH regulation encourages the use of alternative *in vitro* and *in silico* methods in order to minimize animal testing, costs and time. Among these, quantitative structure-activity relationships (QSARs) represent a useful tool to predict unknown activities/properties for existing or even not yet synthesized chemicals. The development and validation of QSAR models for four classes of emerging pollutants (brominated flame retardants, fragrances, perfluorinated compounds and (benzo)triazoles) is the central topic of Work Package 3 (WP3) within the FP7 European project CADASTER (CAse studies on the Development and Application of in-Silico Techniques for Environmental hazard and Risk assessment). The final goal of the project is to exemplify the integration of information, models and strategies for carrying out hazard and risk assessments for large numbers of substances, organized in the four representative chemical classes.
The aim of this poster is to summarize the WP3 activities within CADASTER project and the QSAR/QSPR models developed so far for the four classes of compounds under investigation. This modeling activity involved different project partners in universities and research institutes across Europe (University of Insubria, Linnaeus University, IVL Swedish Environmental Research Institute, Ideaconsult Ltd. and Helmholtz Zentrum München), and was realized by different modeling approaches. For each class, *ad hoc* QSARs were developed for all the available experimental data (i.e. physico-chemical properties, environmental and mammalian toxicity) in order to characterize environmental behavior and activity profile of the chemicals. In agreement with the OECD principles for the validation of QSARs for regulatory purposes, all the proposed models were checked for their robustness, external predictivity and applicability domain. QSAR predictions, together with structural analysis (e.g. similarity analysis and multivariate ranking methods), were used for the identification of priority compounds (also among the ECHA pre-registration list) to optimize the experimental testing to be performed in WP2.

FINAL GOAL of CADASTER (1)

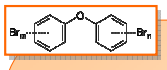
to exemplify the integration of information, models and strategies for carrying out hazard and risk assessments for four classes of emerging pollutants:

- **Brominated Flame Retardants**
- **Fragrances**
- **Perfluorinated Compounds**
- **Triazoles / benzotriazoles**

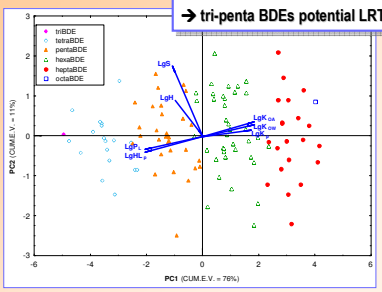
WP2 Collection of data and models
WP3 Development and validation of QSARs
WP4 Integration of QSARs within hazard and risk assessment
WP5 Development of website and stand-alone tools for dissemination and project results

- **D3.1** Chemical structures and molecular descriptor
- **D3.2** Evaluation of existing QSARs according to OECD principles
- **D3.3** Gap analysis
- **D3.4** Prioritization through similarity analysis and ranking methods
- **D3.5** New QSARs for relevant end-points
- **D3.6** Multi-model approaches
- **D3.7** External validation of models with new experimental data (WP2)

BROMINATED FLAME RETARDANTS (BFRs)



Characterization of the environmental behaviour
→ tri-penta BDEs potential LRT



Models for PHYS-CHEM properties & FATE

ENDPOINT (N_{obj})	SIDS	MODELS	Ref.
Henry (7)	✓	Partner: UI	
Melting Point (25)	✓		
Vapor Pressure (34)	✓	DESCRIPTORS:	*a-b*
Water Solubility (12)	✓	Dragon 5.5	
logKoa (30)			
logKow (20)	✓	METHOD:	
K and HL Photolysis (15)	✓	MLR-OLS	
VP (30)	✓	LNU - Dragon 1.11 - PLS	*c*

Comparison of MP, logKoa and logKow models with general models (e.g. EPI Suite, MLOGP)
→ Higher accuracy in prediction of local models especially for highly brominated PBDEs
→ Check of structural AD interpolation or extrapolation?

Models for TOXICITY

ENDPOINT (N_{obj})	SIDS	Partner	DESC.	METHOD	Ref	
AMES Test (4361)	✓	HMGU	E-state	ASNN	*d*	→ General model including 10 BFRs
ED activity (~20)		UI	Dragon 5.5	MLR/K-NN	*e-f*	→ Local model for BFRs

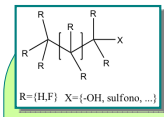
***Ad hoc* local models for BFRs**

Predictions and AD for 209 PBDEs + other BFRs

Experimental and Predicted data used for Prioritization (D 3.4)
→ 11 BFRs suggested for experimental testing

PUBLICATIONS
[a] Papa, E.; Kovarich, S.; Gramatica P., *QSAR Comb. Sci.* (2009) 28, 790-796.
[b] Papa, E.; Kovarich, S.; Gramatica P., *Mol. Info.* (2011) 30, 232-240.
[c] Oberg T., *Environ Sci & Pollut Res* (2002) 9, 405-411.
[d] Sushko, I. et al., *J. Chem. Inf. Model.* (2010) 50, 2094-2111.
[e] Papa, E.; Kovarich, S.; Gramatica, P., *Chem. Res. Toxicol.* (2010) 23, 946-954.
[f] Kovarich, S.; Papa, E.; Gramatica, P., *J Haz. Mat.* (2011) on-line

PERFLUORINATED COMPOUNDS (PFCs)



ENDPOINT (N_{obj})	SIDS	UI	LNU	IDEA	HMGU	Ref.
MP (94)	✓	Dragon desc.	Dragon desc.	Topol.+Fragments	E-State	*a*
BP (93)	✓	MLR-OLS	PLSR	MLR-OLS	ASNN	

CONSENSUS

Models for PHYS-CHEM properties

ENDPOINT	N_{obj}	SIDS	Partner	DESCRIPTORS	METHOD	Ref.	
VP	870+5	✓	LNU	Dragon 5.4	PLS	*b*	→ Extension to PFCs of an existing general QSPR model
VP	35	✓	UI	Dragon 5.5	MLR	*c*	→ LOCAL models developed *ad hoc* for PFCs
WS	20	✓					
CMC	10	✓					

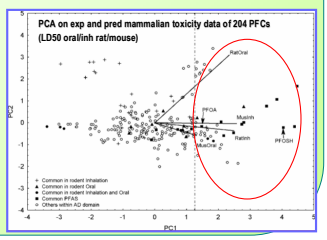
RESULTS reported in POSTER n° TU-347

Models for TOXICITY

ENDPOINT	N_{obj}	SIDS	Partner	DESC.	METHOD	Ref	
AMES Test	4361	✓	HMGU			*d*	→ General model including 25 PFCs
ED activity	24		UI	Dragon 5.5	K-NN (*ad hoc* models)	*e*	
					Application of existing classification models for ED		
Rodent acute tox	~55	✓	UI	Dragon 5.5	MLR-OLS	*f-g*	→ Local model for PFCs

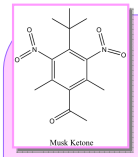
Predictions and AD for ~220 PFCs (ECHA list)

Prioritization based on toxicity profile (D 3.4)
→ 22 long chain PFCs were identified as the most toxic for all the modelled endpoints and suggested for experimental testing



PUBLICATIONS
[a] Bhhatarai, B. et al. (WP3 partners), *Molecular Informatics* (2011) 30, 189-204.
[b] Oberg,T.; Liu, T., Chemometr. Intell. Lab. Syst. (2011) in press.
[c] Bhhatarai, B.; Gramatica, P., *Environ. Sci. Technol.* (2010) Online first
[d] Sushko, I. et al., *J. Chem. Inf. Model.* (2010) 50, 2094-2111.
[e] Kovarich, S.; Gramatica, P; Papa, E., POSTER presented at SETAC-Europe 2010
[f] Bhhatarai, B.; Gramatica, P. , *Molecular Diversity* (2011) 15, 467-476.
[g] Bhhatarai, B.; Gramatica P., *Chem. Res. Toxicol.* (2010) 23, 528-539.

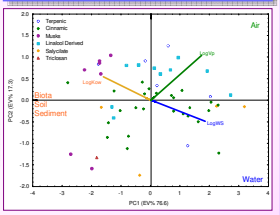
FRAGRANCES



Models for PHYS-CHEM properties & FATE

ENDPOINT (N_{obj})	SIDS	MODELS	Ref.
Vapor Pressure (37)	✓	Partner: UI	
Water Solubility (37)	✓	DESCRIPTORS: Dragon 5.5	*a*
logKow (52)	✓	METHOD: MLR-OLS	

Characterization of the environmental partitioning profile



✓ Better results than EPI Suite and check of structural AD

Models for TOXICITY

ENDPOINT (N_{obj})	SIDS	PARTNER	DESC	METHOD	Ref.	
AMES test (4361 (65 fr.)	✓	HMGU	E-state	ASNN	*b*	General models including Fragrances
T. Pyriformis (644 (36fr.)			E-state	ASNN	*c*	
Rodent acute toxicity (23)	✓	UI	Dragon 5.5	MLR	*d*	
Cytotoxicity (~20)						

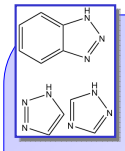
→ Local models developed *ad hoc* for Fragrances

Lack of experimental data for endpoints of interest for CADASTER, to be modelled by QSAR.

Prioritization based on toxicity profile (exp + pred data) and structural representativity (MDS, factorial design) (D 3.4)

PUBLICATIONS
[a] Papa, E.; Luini, M.; Gramatica, POSTER presented at SETAC-Europe 2009.
[b] Sushko, I. et al., *J. Chem. Inf. Model.* (2010) 50, 2094-2111.
[c] Tetko, I.V. Et al., J. Chem. Inf. Model. (2008) 48, 1733-1746.
[d] Papa, E.; Luini, M.; Gramatica, P. *SAR QSAR Environ. Res.* (2009) 20, 767–779.

TRIAZOLES & BENZO-TRIAZOLES (B-TAZs)



Models for PHYS-CHEM properties

ENDPOINT (N_{obj})	SIDS	MODELS	Ref.
Melting Point (56)	✓	Partner: UI	Bhhatarai B., P. Gramatica, Water Res. (2011) 45, 1463-1471
Vapor Pressure (33)	✓	DESCRIPTORS: Dragon 5.5	
Water Solubility (49)	✓		
logKow (64)	✓	METHOD: MLR-OLS	

✓ Characterization of the environmental behaviour → identification of most soluble compounds
✓ Better results than EPI Suite
✓ Check of structural AD

RESULTS reported in POSTER n° TU-347

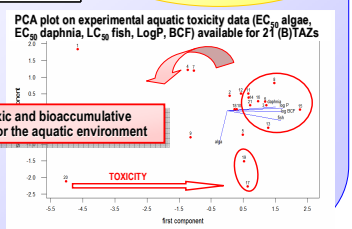
Models for TOXICITY

ENDPOINT (N_{obj})	SIDS	Partner	DESCRIPTORS	METHOD	Ref.	
AMES Test (4361)	✓	HMGU	E-state	ASNN	*Sushko et al.*	→ General model including 17 (B)TAZs

ENDPOINT	SIDS	UI	IVL	Ref.
EC_{50} Algae	✓	DESCRIPTORS:	DESCRIPTORS:	RESULTS reported in POSTER n° TU-348
EC_{50} Daphnia	✓	Dragon 5.5 / on-line	Dragon 6 / on-line	
LC_{50} Fish	✓	METHOD: MLR-OLS	METHOD: PLS	

Predictions and AD for > 350 (B)TAZs (ECHA list)

Prioritization based on eco-toxicity profile (D 3.4)



Most toxic and bioaccumulative B-TAZ for the aquatic environment

Identification of structurally similar (B)TAZs without experimental data — MDS +K-ANN

23 (B)TAZs suggested for experimental testing

Financial support, given by European Union through the project CADASTER (FP7-ENV-2007-1-212668), is gratefully acknowledged (www.cadaster.eu)

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