

QSAR Modeling of Cytotoxicity of Nanoparticles

Andrey A. Toropov¹, Alla P. Toropova¹, Emilio Benfenati¹,
Giuseppina Gini², Danuta Leszczynska³, Jerzy Leszczynski⁴



¹ Laboratory of Environmental Chemistry and Toxicology, Istituto di Ricerche Farmacologiche Mario Negri - Via Giuseppe La Masa 19 - 20156 - Milan - ITALY

² Department of Electronics and Information, Politecnico di Milano - Via Ponzio 34/5 - 20133 - Milan - ITALY

³ Interdisciplinary Center for Nanotoxicity, Department of Civil and Environmental Engineering, Jackson State University - 1325 J. R. Lynch Street - Jackson, MS - 39217-0510 - USA

⁴ Interdisciplinary Center for Nanotoxicity, Department of Chemistry and Biochemistry, Jackson State University - 1400 J. R. Lynch Street - Jackson, MS - 39217-0510, USA



Abstract

CORAL (<http://www.insilico.eu/CORAL>) has been used to build up QSAR models for prediction of cytotoxicity of metal oxide nanoparticles to bacteria *Escherichia coli* (pEC50). Six random splits into the training and test set were examined. It has shown that the CORAL is a satisfactory tool to build up a QSAR of the pEC50.

andrey.toropov@marionegri.it

Method

Data

The numerical data on cytotoxicity of metal oxide nanoparticles (n=17) to bacteria *Escherichia coli* (pEC50) have been taken in the literature [1]. Six random splits into the training and test sets were examined.

Optimal descriptors

The SMILES-based optimal descriptors [2-5] have been calculated as the following:

$$DCW(T) = \sum_{k=1}^E W(1S_k) + \sum_{k=1}^{E-1} W(2S_k) + \sum_{k=1}^{E-2} W(3S_k)$$

where $1S_k$, $2S_k$, and $3S_k$ are one-, two-, and three-elements SMILES attributes; E is the total number of SMILES elements for a given molecular structure; $W(1S_k)$, $W(2S_k)$, and $W(3S_k)$ are the correlation weights of the attributes. The SMILES element comprises one or two symbols which should be examined as one (e.g., 'Cl', 'Br', etc.).

The threshold is a value used to classify attributes as either rare or active. For instance, if the threshold is 5, then attributes found in four (or fewer) SMILES structures of the training set should be classified as rare. The correlation weights of rare attributes are blocked with their values fixed at zero. E is the number of $1S_k$. Better prediction for all random splits takes place if T=2.

TABLE 1

Oxide	SMILES	pEC50
ZnO	O=[Zn]	3.45
CuO	[Cu]=O	3.20
V ₂ O ₅	O=[V]O[V]=O	3.14
Y ₂ O ₃	O=[Y]O[Y]=O	2.87
Bi ₂ O ₃	O=[Bi]O[Bi]=O	2.82
In ₂ O ₃	O=[In]O[In]=O	2.81
Sb ₂ O ₃	O=[Sb]O[Sb]=O	2.64
Al ₂ O ₃	O=[Al]O[Al]=O	2.49
Fe ₂ O ₃	O=[Fe]O[Fe]=O	2.29
SiO ₂	O=[Si]=O	2.20
ZrO ₂	O=[Zr]=O	2.15
SnO ₂	O=[Sn]=O	2.01
TiO ₂	O=[Ti]=O	1.74
CoO	[Co]=O	3.51
NiO	[Ni]=O	3.45
Cr ₂ O ₃	O=[Cr]O[Cr]=O	2.51
La ₂ O ₃	O=[La]O[La]=O	2.87

TABLE 2

Split	Probe	n	r ²	RMSE	F	n	r ²	RMSE	F
Split1	1	11	0.7407	0.234	26	6	0.9397	0.205	62
	2	11	0.7407	0.234	26	6	0.9400	0.205	63
	3	11	0.7407	0.234	26	6	0.9409	0.204	64
	average		0.7407	0.234	26	6	0.9402	0.204	63
Split2	1	11	0.8217	0.232	41	6	0.9650	0.237	110
	2	11	0.8217	0.232	41	6	0.9647	0.236	109
	3	11	0.8217	0.232	41	6	0.9648	0.236	110
	average		0.8217	0.232	41	6	0.9648	0.236	110
Split3	1	11	0.8215	0.170	41	6	0.8325	0.336	20
	2	11	0.8213	0.170	41	6	0.8407	0.338	21
	3	11	0.8215	0.170	41	6	0.8357	0.337	20
	average		0.8214	0.170	41	6	0.8363	0.337	20
Split4	1	10	0.7779	0.261	28	7	0.9466	0.140	89
	2	10	0.7779	0.261	28	7	0.9469	0.139	89
	3	10	0.7779	0.261	28	7	0.9470	0.139	89
	average		0.7779	0.261	28	7	0.9468	0.139	89
Split5	1	11	0.8172	0.207	40	6	0.9268	0.270	51
	2	11	0.8172	0.207	40	6	0.9285	0.270	52
	3	11	0.8169	0.207	40	6	0.9227	0.271	48
	average		0.8171	0.207	40	6	0.9260	0.270	50
Split6	1	11	0.8377	0.190	46	6	0.8494	0.294	23
	2	11	0.8377	0.190	46	6	0.8463	0.296	22
	3	11	0.8377	0.190	46	6	0.8505	0.293	23
	average		0.8377	0.190	46	6	0.8487	0.294	22

Results

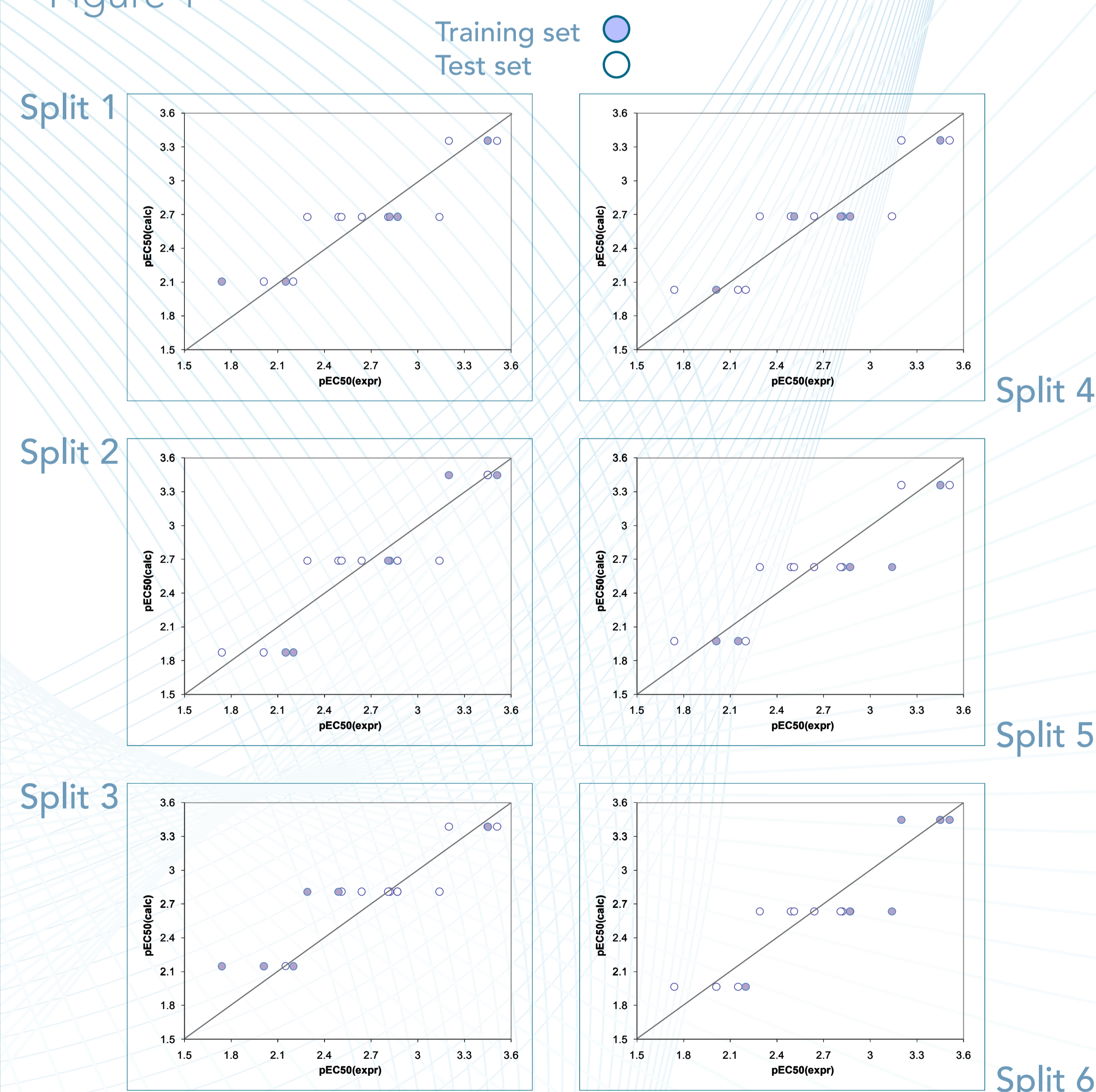
TABLE 1 shows pEC50 values and SMILES representing the molecular structure of the metal oxides.

TABLE 2 contains the statistical characteristics of the pEC50 models for six random splits.

Figure 1 shows best models for each split graphically.

There is significant difference of the predictability for these models (TABLE 1), but all models can be estimated as satisfactory.

Figure 1

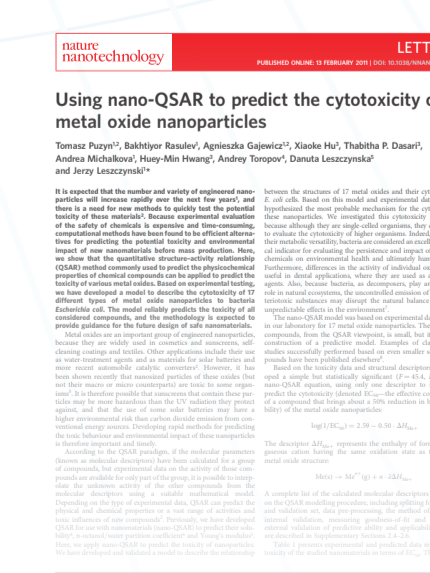


Conclusions

CORAL is a satisfactory tool for QSAR modelling of the pEC50 of the metal oxide nanoparticles cytotoxicity to bacteria *Escherichia coli*. It provides results similar to those we published on *Nature Nanotechnology* [1].

References

- [1] Puzyn, T., Rasulev, B., Gajewicz, A., Hu, X., Dasari, T.P., Michalkova, A., Hwang, H.-M., Toropov, A., Leszczynska, D., Leszczynski, J., Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles, *Nat. Nanotechnol.*, 6 (2011) : 175-178
- [2] Toropova, A.P., Toropov, A.A., Benfenati, E., Leszczynska, D., Leszczynski, J., QSAR modeling of measured binding affinity for fullerene-based HIV-1 PR inhibitors by CORAL, *J. Math. Chem* 47 (2010) : 959-987
- [3] Toropov, A.A., Toropova, A.P., Benfenati, E., Leszczynska, D., Leszczynski, J., SMILES-Based Optimal Descriptors: QSAR Analysis of Fullerene-Based HIV-1 PR Inhibitors by Means of Balance of Correlations, *J. Comput. Chem.* 31 (2010) : 381-392
- [4] Toropov, A.A., Toropova, A.P., Benfenati, E., Leszczynska, D., Leszczynski, J., InChI-based optimal descriptors: QSAR analysis of fullerene[C60]-based HIV-1 PR inhibitors by correlation balance, *Eur. J. Med. Chem.* 45 (2010) :1387-1394
- [5] Toropova, A.P., Toropov, A.A., Benfenati, E., Gini, G., Leszczynska, D., Leszczynski, J., CORAL: QSPR models for solubility of [C60] and [C70] fullerene derivatives, *Mol. Divers in press*



Financial support from OSIRIS (Contract no. GOCE-CT-2007-037017) and ORCHESTRA (Contract no. FP7-COOPERATION-226521) projects is gratefully acknowledged

