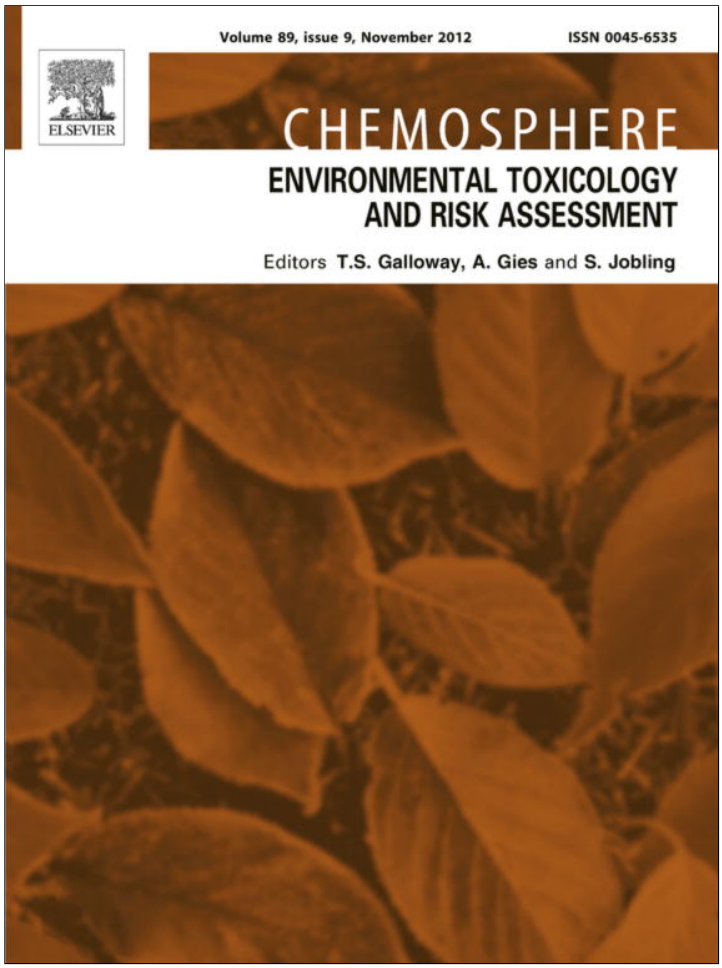

Chemosphere 89 (2012) 1098–1102

Contents lists available at SciVerse ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Novel application of the CORAL software to model cytotoxicity of metal oxide nanoparticles to bacteria *Escherichia coli*

Andrey A. Toropov [a,*], Alla P. Toropova [a], Emilio Benfenati [a], Giuseppina Gini [b], Tomasz Puzyn [c], Danuta Leszczynska [d], Jerzy Leszczynski [e]

[a] Istituto di Ricerche Farmacologiche Mario Negri, Via La Masa 19, 20156 Milano, Italy
[b] Department of Electronics and Information, Politecnico di Milano, piazza Leonardo da Vinci 32, 20133 Milano, Italy
[c] Faculty of Chemistry, Laboratory of Environmental Chemometrics, University of Gdansk, ul. Sobieskiego 18/19, Gdansk 80-952, Poland
[d] Interdisciplinary Nanotoxicity Center, Department of Civil and Environmental Engineering, Jackson State University, 1325 Lynch Street, Jackson, MS 39217-0510, USA
[e] Interdisciplinary Nanotoxicity Center, Department of Chemistry and Biochemistry, Jackson State University, 1400 JR Lynch Street, P.O. Box 17910, Jackson, MS 39217, USA

HIGHLIGHTS

- The CORAL software for the building up of QSPR/QSAR models is suggested.
- The CORAL model for cytotoxicity of metal oxide nanoparticles is demonstrated.
- The model is a mathematical function of the numbers of oxygen atoms and double bonds.

ARTICLE INFO

Article history:
Received 1 February 2012
Received in revised form 20 April 2012
Accepted 16 May 2012
Available online 15 June 2012

Keywords:
QSAR
CORAL software
Cytotoxicity to bacterium *Escherichia coli*
Metal oxide nanoparticle

ABSTRACT

Convenient to apply and available on the Internet software CORAL (http://www.insilico.eu/CORAL) has been used to build up quantitative structure–activity relationships (QSAR) for prediction of cytotoxicity of metal oxide nanoparticles to bacteria *Escherichia coli* (minus logarithm of concentration for 50% effect pEC50). In this study six random splits of the data into the training and test set were examined. It has been shown that the CORAL provides a reliable tool that could be used to build up a QSAR of the pEC50.

1. Introduction

Nanomaterials gradually become vital constituent of modern industry. They are visible on the shells of stores as components of the everyday necessities (medicine, cosmetics, etc.). These widespread applications of nanomaterials rise questions concerning their environmental and health effects. It was recently argued that only direct prediction of toxicity of nanomaterials provides a reliable tool for evaluation of their effects on environment and humans (Leszczynski, 2010). Such issues could be (at least in part) addressed using modern scientific techniques. Among them are methods of computational chemistry, and the most appropriate approach to study such phenomena the quantitative structure–property/activity relationships (QSPRs/QSARs) (Randic, 1991; Cosentino et al., 2000; Balaban et al., 2005; Ivanciuc et al., 2006; Enoch et al., 2008; Tetko et al., 2008; Bhhatarai et al., 2010; Das and Trinajstic, 2010; Leszczynski, 2010; Katritzky et al., 2010; Mitra et al., 2010; Afantitis et al., 2011; Basak, 2011; Duchowicz et al., 2011; Furtula and Gutman, 2011).

The QSAR studies of nanomaterials, such as the metal oxide nanoparticles (Puzyn et al., 2011) fullerene C_{60} (Toropova et al., 2010; Toropov et al., 2010a,b), and C_{70} (Toropova et al., 2011) become an integral part of the QSPR/QSAR theory and applications of various approaches. Among many possible tools for the QSPR/QSAR analysis of the nanomaterials is CORAL software (http://www.insilico.eu/CORAL). The aim of the present study is evaluation of the CORAL as a tool for the QSAR analysis of the toxicity of metal oxide nanoparticles.

2. Method

2.1. Data

The numerical data on cytotoxicity to bacteria *E. coli* (minus logarithm of concentration for 50% effect, $n = 17$, pEC50) have been

* Corresponding author. Tel.: +39 0239014805.
E-mail address: andrey.toropov@marionegri.it (A.A. Toropov).

0045-6535/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.chemosphere.2012.05.077

Table 1
Percentage of identity for kth and jth splits $P_{kj} = \frac{N_{train}+N_{test}}{17} \times 100\%$, where N_{train} is the number of oxides which are placed in training set for both kth and jth splits, N_{test} is the number of oxides which are placed in test set for both kth and jth splits.

	Split1	Split 2	Split 3	Split 4	Split 5	Split 6
Split 1	100	52.9	52.9	58.8	64.7	41.2
Split 2		100	41.2	47.1	41.2	64.7
Split 3			100	47.1	52.9	41.2
Split 4				100	70.6	47.1
Split 5					100	64.7
Split 6						100
Oxide						
ZnO	Test	Training	Test	Test	Test	Training
CuO	Training	Test	Training	Training	Training	Test
V_2O_3	Training	Training	Training	Training	Test	Test
Y_2O_3	Training	Training	Training	Test	Test	Training
Bi_2O_3	Test	Test	Training	Test	Training	Training
In_2O_3	Training	Test	Training	Test	Training	Training
Sb_2O_3	Training	Training	Training	Training	Training	Training
Al_2O_3	Training	Training	Test	Training	Training	Training
Fe_2O_3	Training	Training	Test	Training	Training	Test
SiO_2	Training	Test	Test	Training	Training	Test
ZrO_2	Test	Test	Training	Training	Test	Training
SnO_2	Training	Training	Test	Test	Test	Training
TiO_2	Test	Training	Test	Training	Training	Training
CoO	Training	Test	Training	Training	Training	Test
NiO	Test	Training	Training	Test	Test	Test
Cr_2O_3	Training	Training	Training	Test	Training	Training
La_2O_3	Test	Training	Training	Training	Training	Training

Table 2
Structures and cytotoxicity to bacteria *Escherichia coli* of oxides.

Oxide	SMILES	pEC50 (mol/L)
ZnO	O = [Zn]	3.45
CuO	[Cu] = O	3.20
V_2O_3	O = [V]O[V] = O	3.14
Y_2O_3	O = [Y]O[Y] = O	2.87
Bi_2O_3	O = [Bi]O[Bi] = O	2.82
In_2O_3	O = [In]O[In] = O	2.81
Sb_2O_3	O = [Sb]O[Sb] = O	2.64
Al_2O_3	O = [Al]O[Al] = O	2.49
Fe_2O_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_2O_3	O = [Cr]O[Cr] = O	2.51
La_2O_3	O = [La]O[La] = O	2.87

taken from the literature (Puzyn et al., 2011). Six random splits into the training and test sets were examined. Two principles were used in order to prepare these splits (i) ranges of pEC50 for training and test set are as equivalent as possible; and (ii) percentage of splits identity is as small as possible (Table 1). Table 2 contains the details of structures of the considered metal oxides.

Table 3
Statistical quality of QSAR models of the toxicity calculated with CORAL software. Best predictions have been obtained with the threshold = 2 (these are given in bold).

T	N_{act}	Probe	Training set				Test set			
			n	r^2	RMSE	F	n	r^2	RMSE	F
Split 1										
0	119	1	11	0.9999	0.004	153934	6	0.8641	0.567	25
0	119	2	11	0.9999	0.005	70019	6	0.7268	0.675	11
0	119	3	11	0.9999	0.005	63146	6	0.5409	0.575	5
0		Average		0.9999	0.005	95699		0.7106	0.606	14
1	14	1	11	0.9999	0.004	94246	6	0.3107	0.773	2
1	14	2	11	0.9999	0.005	91578	6	0.6290	0.756	7
1	14	3	11	0.9999	0.005	70216	6	0.5214	0.735	4
1		Average		0.9999	0.005	85347		0.4870	0.755	4
2	**3**	**1**	**11**	**0.7407**	**0.234**	**26**	**6**	**0.9397**	**0.205**	**62**
2	**3**	**2**	**11**	**0.7407**	**0.234**	**26**	**6**	**0.9400**	**0.205**	**63**
2	**3**	**3**	**11**	**0.7407**	**0.234**	**26**	**6**	**0.9409**	**0.204**	**64**
2		Average		**0.7407**	**0.234**	**26**		**0.9402**	**0.204**	**63**
Split 2										
0	119	1	11	0.9996	0.011	21798	6	0.3722	0.557	2
0	119	2	11	0.9999	0.006	84400	6	0.3626	0.467	2
0	119	3	11	0.9999	0.006	75565	6	0.6247	0.498	7
0		Average		0.9998	0.008	60588		0.4532	0.507	4
1	14	1	11	0.9998	0.007	57609	6	0.9657	0.872	113
1	14	2	11	0.9999	0.004	156263	6	0.6788	0.643	8
1	14	3	11	0.9999	0.005	104526	6	0.9435	0.737	67
1		Average		0.9999	0.005	106133		0.8627	0.751	63
2	**3**	**1**	**11**	**0.8217**	**0.232**	**41**	**6**	**0.9650**	**0.237**	**110**
2	**3**	**2**	**11**	**0.8217**	**0.232**	**41**	**6**	**0.9647**	**0.236**	**109**
2	**3**	**3**	**11**	**0.8217**	**0.232**	**41**	**6**	**0.9648**	**0.236**	**110**
2		Average		**0.8217**	**0.232**	**41**		**0.9648**	**0.236**	**110**
Split 3										
0	119	1	11	0.9999	0.004	84785	6	0.4563	0.619	3
0	119	2	11	0.9999	0.004	76448	6	0.3982	0.647	3
0	119	3	11	0.9998	0.006	45466	6	0.0010	0.732	0
0		Average		0.9999	0.005	68900		0.2852	0.666	2
1	14	1	11	0.9999	0.005	69455	6	0.4778	0.505	4
1	14	2	11	0.9999	0.005	71171	6	0.5347	0.521	5
1	14	3	11	0.9999	0.005	63169	6	0.6919	0.534	9
1		Average		0.9999	0.005	67932		0.5681	0.520	6
2	**3**	**1**	**11**	**0.8215**	**0.170**	**41**	**6**	**0.8325**	**0.336**	**20**
2	**3**	**2**	**11**	**0.8213**	**0.170**	**41**	**6**	**0.8407**	**0.338**	**21**

(*continued on next page*)

Table 3 (*continued*)

T	N_{act}	Probe	Training set				Test set			
			n	r^2	RMSE	*F*	*n*	r^2	RMSE	*F*
2	**3**	**3**	**11**	**0.8215**	**0.170**	**41**	**6**	**0.8357**	**0.337**	**20**
2		Average		**0.8214**	**0.170**	**41**		**0.8363**	**0.337**	**20**
Split 4										
0	119	1	10	0.9997	0.009	28593	7	0.0063	0.578	0
0	119	2	10	0.9998	0.007	52545	7	0.8827	0.503	38
0	119	3	10	0.9998	0.008	35489	7	0.8345	0.527	25
0		Average		0.9998	0.008	38876		0.5745	0.536	21
1	13	1	10	0.9999	0.005	87592	7	0.7126	0.758	12
1	13	2	10	1.0000	0.003	275957	7	0.0971	0.864	1
1	13	3	10	0.9997	0.010	26515	7	0.7932	0.923	19
1		Average		0.9999	0.006	130021		0.5343	0.848	11
2	**3**	**1**	**10**	**0.7779**	**0.261**	**28**	**7**	**0.9466**	**0.140**	**89**
2	**3**	**2**	**10**	**0.7779**	**0.261**	**28**	**7**	**0.9469**	**0.139**	**89**
2	**3**	**3**	**10**	**0.7779**	**0.261**	**28**	**7**	**0.9470**	**0.139**	**89**
2		Average		**0.7779**	**0.261**	**28**		**0.9468**	**0.139**	**89**
Split 5										
0	119	1	11	0.9997	0.008	34909	6	0.7779	0.568	14
0	119	2	11	0.9999	0.006	67200	6	0.8649	0.588	26
0	119	3	11	0.9998	0.006	57434	6	0.9750	0.719	156
0		Average		0.9998	0.006	53181		0.8726	0.625	65
1	16	1	11	0.9999	0.006	68393	6	0.9269	0.930	51
1	16	2	11	0.9999	0.006	62183	6	0.9271	0.825	51
1	16	3	11	0.9994	0.012	15885	6	0.6893	1.178	9
1		Average		0.9997	0.008	48820		0.8478	0.978	37
2	**3**	**1**	**11**	**0.8172**	**0.207**	**40**	**6**	**0.9268**	**0.270**	**51**
2	**3**	**2**	**11**	**0.8172**	**0.207**	**40**	**6**	**0.9285**	**0.270**	**52**
2	**3**	**3**	**11**	**0.8169**	**0.207**	**40**	**6**	**0.9227**	**0.271**	**48**
2		Average		**0.8171**	**0.207**	**40**		**0.9260**	**0.270**	**50**
Split 6										
0	119	1	11	0.9996	0.009	22722	6	0.4711	0.727	4
0	119	2	11	0.9998	0.006	59324	6	0.9072	0.599	39
0	119	3	11	0.9998	0.007	37320	6	0.8546	0.688	24
0		Average		0.9997	0.008	39789		0.7443	0.671	22
1	16	1	11	0.9997	0.009	26143	6	0.6517	1.194	7
1	16	2	11	0.9999	0.006	66438	6	0.9076	1.004	39
1	16	3	11	0.9996	0.009	22457	6	0.4862	1.056	4
1		Average		0.9997	0.008	38346		0.6818	1.085	17
2	**3**	**1**	**11**	**0.8377**	**0.190**	**46**	**6**	**0.8494**	**0.294**	**23**
2	**3**	**2**	**11**	**0.8377**	**0.190**	**46**	**6**	**0.8463**	**0.296**	**22**
2	**3**	**3**	**11**	**0.8377**	**0.190**	**46**	**6**	**0.8505**	**0.293**	**23**
2		Average		**0.8377**	**0.190**	**46**		**0.8487**	**0.294**	**22**

2.2. Optimal descriptors

The SMILES-based optimal descriptors (Toropov et al., 2010a,b) have been calculated as follows:

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

where ^1S_k, ^2S_k, and ^3S_k are one-, two-, and three-elements SMILES attributes; the E is the total number of SMILES elements for a given molecular structure; $W(^1S_k)$, $W(2S_k)$, and $W(3S_k)$ present 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 which are 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. The *E* is the number of ^1S_k. If a SMILES represents a sequence of element 'ABCDE', then the construction ^1S_k, ^2S_k, and ^3S_k can be represented as:

'ABCDE' → 'A', 'B', 'C', 'D', 'E'(^1S_k)

'ABCDE' → 'AB', 'BC', 'CD', 'DE'(^2S_k)

'ABCDE' → 'ABC', 'BCD', 'CDE'(^3S_k).

Preferable predictability for all random splits is achieved if $T = 2$.

3. Results and discussion

Table 3 contains statistical characteristics of the pEC50 models for six random splits (Table 3). Fig. 1 represents graphically the best models for each split. There is significant difference of the predictability for these models, however, all models can be considered as satisfactory. The model based on quantum chemical calculations described in Puzyn et al. (2011) is characterized by $n = 17$, $r^2 = 0.8619$. No external validation has been carried out for this model (Puzyn et al., 2011). Thus the model obtained in this work is probably characterized by predictability at least of the same degree as the recently published one. In fact the distribution of the $r^2_{(test)}$ values predicted here remains within the 0.83–0.96 range.

One can see (Table 3) robust models for six splits based on three SMILES attributes related to situation where threshold = 2. These are '[', '=', and 'O'. In other words, the model is calculated with solely $W(1S_k)$: $W(2S_k)$ and $W(3S_k)$ were not involved in the model. Table 4 demonstrates calculation of the descriptor. Table 5 contains the model calculated for split 1. Y-randomization (Mitra et al., 2010) of the CORAL models (Table 6) also confirms their predictive potential.

4. Conclusions

An application of the CORAL software for prediction of toxicity of series of nanomaterials was tested. We concluded that CORAL

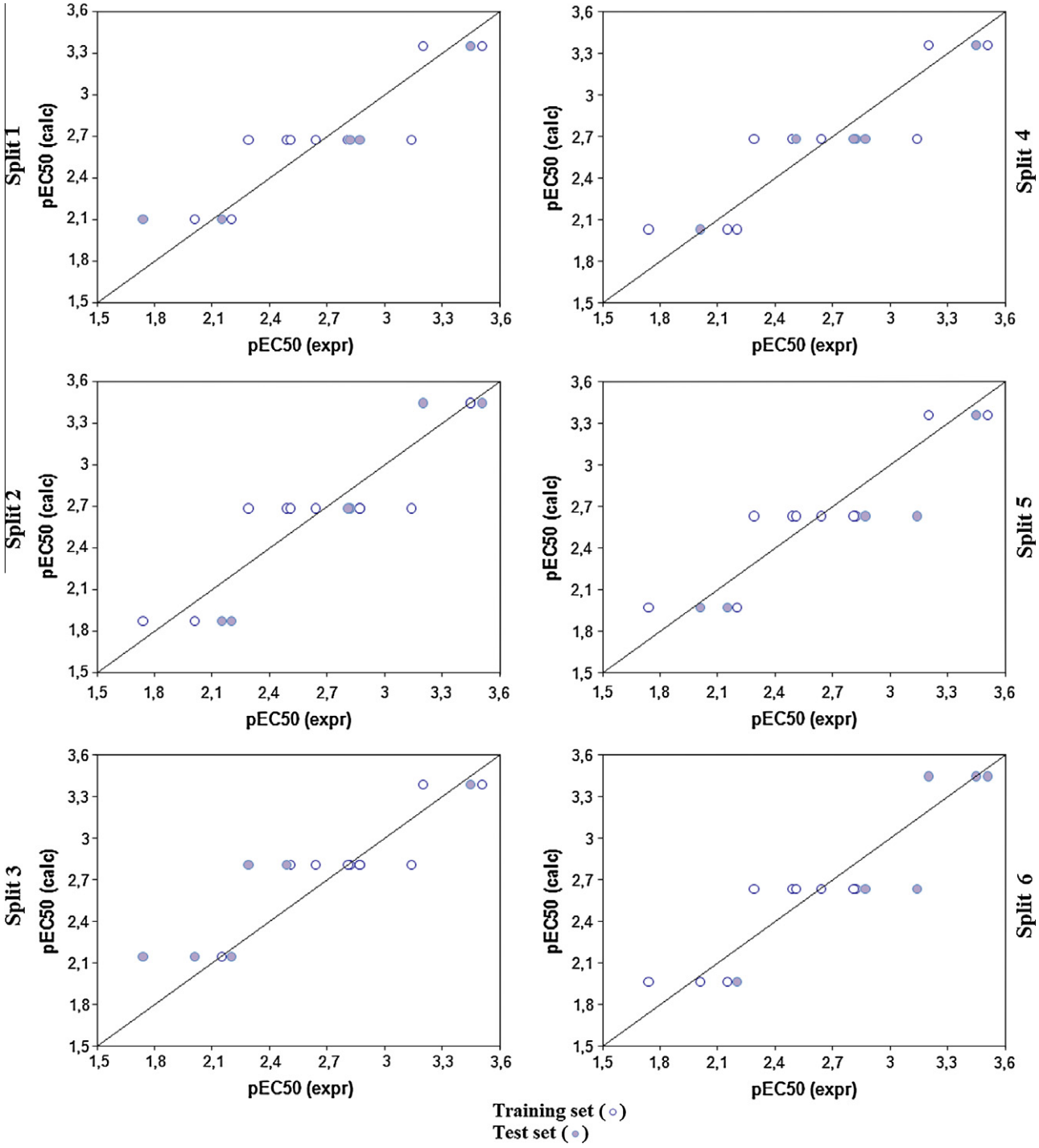

Fig. 1. QSAR models for the cytotoxicity to bacteria *Escherichia coli* of metal oxide particles for six random splits of the data into the training and test sets.

Table 4
An example of DCW(2) calculation: split 1; SMILES is 'O = [Zn]'; DCW(2) = -0.1268810.

^1S_k	$W(^1S_k)$	The number of ^1S_k in training set	The number of ^1S_k in test set
O	−1.5636269	11	6
=	−1.5615161	11	6
[	1.4991310	11	6
Zn	0.0	0	1
[	1.4991310	11	6

Table 5
Calculation of the metal oxide nanoparticles cytotoxicity to bacteria *Escherichia coli* by means of the following model pEC50 = 3.4056 + 0.4000 * DCW(2). The model has been built up for split 1.

Set	SMILES	DCW(2)	pEC50 (mol/L)		
			Expr	Calc	Expr–Calc
Training	[Cu] = O	−0.1268810	3.200	3.355	−0.155
Training	O = [V]O[V] = O	−1.8173889	3.140	2.679	0.461
Training	O = [Y]O[Y] = O	−1.8173889	2.870	2.679	0.191
Training	O = [In]O[In] = O	−1.8173889	2.810	2.679	0.131
Training	O = [Sb]O[Sb] = O	−1.8173889	2.640	2.679	−0.039
Training	O = [Al]O[Al] = O	−1.8173889	2.490	2.679	−0.189
Training	O = [Fe]O[Fe] = O	−1.8173889	2.290	2.679	−0.389
Training	O = [Si] = O	−3.2520240	2.200	2.105	0.095
Training	O = [Sn] = O	−3.2520240	2.010	2.105	−0.095
Training	[Co] = O	−0.1268810	3.510	3.355	0.155
Training	O = [Cr]O[Cr] = O	−1.8173889	2.510	2.679	−0.169
Test	O = [Zn]	−0.1268810	3.450	3.355	0.095
Test	O = [Bi]O[Bi] = O	−1.8173889	2.820	2.679	0.141
Test	O = [Zr] = O	−3.2520240	2.150	2.105	0.045
Test	O = [Ti] = O	−3.2520240	1.740	2.105	−0.365
Test	[Ni] = O	−0.1268810	3.450	3.355	0.095
Test	O = [La]O[La] = O	−1.8173889	2.870	2.679	0.191

Table 6
The checking of the models with Y-randomization.

Probe of Y-scrambling	Split 1 $n = 6$	Split 2 $n = 6$	Split 3 $n = 6$	Split 4 $n = 7$	Split 5 $n = 6$	Split 6 $n = 6$
R^2	0.9402	0.9648	0.8363	0.9468	0.9260	0.8487
R_r^2 1	0.2270	0.3446	0.4729	0.0243	0.5491	0.0230
R_r^2 2	0.2763	0.0539	0.3030	0.1060	0.0030	0.1746
R_r^2 3	0.3762	0.4356	0.3670	0.0720	0.0012	0.1954
R_r^2 4	0.0701	0.6815	0.0272	0.7143	0.0030	0.1014
R_r^2 5	0.0161	0.0810	0.0178	0.0243	0.2331	0.5864
R_r^2 6	0.0161	0.3798	0.0860	0.0047	0.5491	0.1837
R_r^2 7	0.7544	0.0307	0.0077	0.3846	0.5390	0.0169
R_r^2 8	0.4363	0.6747	0.0733	0.1407	0.0140	0.3394
R_r^2 9	0.1470	0.0457	0.1165	0.0797	0.0303	0.1229
R_r^2 10	0.1049	0.0022	0.0756	0.1589	0.7550	0.1954
$\overline{R^2}_r$	0.2424	0.2730	0.1547	0.1709	0.2677	0.1939
$^cR_p^2$	0.807	0.816	0.753	0.857	0.779	0.742

The $\overline{R^2}_r$ is average for ten probes of the Y-scrambling. The $^cR_p^2$ (calculated by $^cR_p^2 = R\sqrt{R^2 - \overline{R}_r^2}$) should be larger than 0.5 (Mitra et al., 2010).

represents a suitable tool for QSAR modelling of the pEC50 of the metal oxide nanoparticles cytotoxicity to bacteria *E. coli*. The robust model (i.e. model without overtraining) is based on data about the presence of oxygen and double bonds.

Acknowledgements

The authors thank for support of the OSIRIS, the NSF CREST Interdisciplinary Nanotoxicity Center NSF-CREST – Grant # HRD-0833178; NSF-EPSCoR Award #: 362492-190200-01/NSFEPS-0903787; and the Department of Defense through the U.S. Army Engineer Research and Development Center, Vicksburg, MS. for generous support through the contract: High Performance Computational Design of Novel Materials (HPCDNM) – Contract #W912 HZ-06-C-0057.

References

Afantitis, A., Melagraki, G., Koutentis, P.A., Srimveis, H., Kollias, G., 2011. Ligand-based virtual screening procedure for the prediction and the identification of novel β-amyloid aggregation inhibitors using Kohonen maps and Counterpropagation Artificial Networks. Eur. J. Med. Chem. 46, 497–508.

Balaban, A.T., Khadikar, P.V., Supuran, G.T., Thakur, A., Thakur, M., 2005. Study on supramolecular complexing ability vis-à-vis estimation of pKa of substituted sulfonamides: dominating role of Balaban index (J). Bioorg. Med. Chem. 15, 3966–3973.

Basak, S.C., 2011. Role of mathematical chemodescriptors and proteomics-based biodescriptors in drug discovery. Drug Dev. Res. 72, 225–233.

Bhhatarai, B., Gang, R., Gramatica, P., 2010. Are mechanistic and statistical QSAR approaches really different? MLR studies on 158 cycloalkyl-pyranones. Mol. Inf. 29, 511–522.

Cosentino, U., Moro, G., Bonalumi, D., Bonati, L., Lasagni, M., Todeschini, R., Pitea, D., 2000. A combined use of global and local approaches in 3D-QSAR. Chemom. Intell. Lab. 52, 183–194.

Das, K.Ch., Trinajstic, N., 2010. Comparison between first geometric-arithmetic index and atom-bond connectivity index. Chem. Phys. Lett. 497, 140–151.

Duchowicz, P.R., Mirifico, M.V., Rozas, M.F., Caram, J.A., Fernandes, F.M., Castro, E.A., 2011. Quantitative structure – spectral property relationships for functional groups of novel 1,2,5-thiadiazole compounds. Chemom. Intell. Lab. 105, 27–37.

Enoch, S.J., Cronin, M.T.D., Schultz, T.W., Madden, J.C., 2008. An evaluation of global QSAR models for the prediction of the toxicity of phenols to Tetrahymena pyriformis. Chemosphere 71, 1225–1232.

Furtula, B., Gutman, I., 2011. Relation between second and third geometric-arithmetic indices of trees. J. Chemom. 25, 87–91.

Ivanciuc, T., Ivanciuc, O., Klein, D.J., 2006. Modeling the bioconcentration factors and bioaccumulation factors of polychlorinated biphenyls with posetic quantitative super-structure/activity relationships (QSSAR). Mol. Diversity 10, 133–145.

Katritzky, A.R., Girgis, A.S., Slavov, S., Tala, S.R., Stoyanova-Slavova, I., 2010. QSAR modeling, synthesis and bioassay of diverse leukemia RPMI-8226 cell line active agents. Eur. J. Med. Chem. 45, 5183–5199.

Leszczynski, J., 2010. Nano meets bio at the interface. Nat. Nanotechnol. 5, 633–634.

Mitra, I., Saha, A., Roy, K., 2010. Exploring quantitative structure–activity relationship studies of antioxidant phenolic compounds obtained from traditional Chinese medicinal plants. Mol. Simul. 13, 1067–1079.

Puzyn, T., Rasulev, B., Gajewicz, A., Hu, X., Dasari, T.P., Michalkova, A., Hwang, H.M., Toropov, A., Leszczynska, D., Leszczynski, J., 2011. Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. Nat. Nanotechnol. 6, 175–178.

Randic, M., 1991. Novel graph theoretical approach to heteroatoms in quantitative structure–activity relationships. Chemom. Intell. Lab. 10, 213–227.

Tetko, I.V., Jaroszewicz, I., Platts, J.A., Kuduk-Jaworska, J., 2008. Calculation of lipophilicity for Pt(II) complexes: experimental comparison of several methods. J. Inorg. Biochem. 102, 1224–1237.

Toropov, A.A., Toropova, A.P., Benfenati, E., Leszczynska, D., Leszczynski, J., 2010a. SMILES-based optimal descriptors: QSAR analysis of fullerene-based HIV-1 PR inhibitors by means of balance of correlations. J. Comput. Chem. 31, 381–392.

Toropov, A.A., Toropova, A.P., Benfenati, E., Leszczynska, D., Leszczynski, J., 2010b. InChI-based optimal descriptors: QSAR analysis of fullerene[C60]-based HIV-1 PR inhibitors by correlation balance. Eur. J. Med. Chem. 45, 1387–1394.

Toropova, A.P., Toropov, A.A., Benfenati, E., Leszczynska, D., Leszczynski, J., 2010. QSAR modeling of measured binding affinity for fullerene-based HIV-1 PR inhibitors by CORAL. J. Math. Chem. 47, 959–987.

Toropova, A.P., Toropov, A.A., Benfenati, E., Gini, G., Leszczynska, D., Leszczynski, J., 2011. CORAL: QSPR models for solubility of [C_{60}] and [C_{70}] fullerene derivatives. Mol. Diversity 105, 249–256.