1 Extrapolating between toxicity endpoints of metal oxide
2 nanoparticles: Predicting toxicity to Escherichia coli and
3 human keratinocyte cell line (HaCaT) with Nano-QTTR
4

Supratik Kar ^{a, b, c}, Agnieszka Gajewicz ^a, Kunal Roy ^b, Jerzy Leszczynski ^c,
Tomasz Puzyn ^{a*}

^aLaboratory of Environmental Chemometrics,
Institute for Environmental and Human Health Protection, Faculty of Chemistry,
University of Gdańsk, Gdańsk, Poland

^bDrug Theoretics and Cheminformatics Laboratory,
Department of Pharmaceutical Technology,
Jadavpur University, Kolkata 700032, India

^cInterdisciplinary Nanotoxicity Center, Department of Chemistry and Biochemistry, Jackson
State University, Jackson MS, USA

5
6
7
8
9
10
*Corresponding author: Tomasz Puzyn; University of Gdańsk, Faculty of Chemistry, Wita
Stwosza 63, 80-308 Gdańsk, Poland; e-mail: t.puzyn@qsar.eu.org; phone: +48 58 523 5248

Abstract

Synthesis of novel nanoparticles should always be accompanied by a comprehensive assessment of risk to human health and to ecosystem. Application of *in silico* models is encouraged by regulatory authorities to fill the data gaps related to the properties of nanoparticles affecting the environment and human health. Interspecies toxicity correlations provide a tool for estimation of contaminant's sensitivity with known levels of uncertainty for a diverse pool of species. We propose here first interspecies cytotoxicity correlation models between *Escherichia coli* (prokaryotic system) and human keratinocyte cell line (HaCaT) (eukaryotic system) to assess the discriminatory features for cytotoxicity of metal oxide nanoparticles. The nano-QTTR models can be employed for extrapolating cytotoxicity to *Escherichia coli* and human keratinocyte cell line (HaCaT) for metal nanoparticles when the data for the other species are available. Informative illustrations of the contributing mechanisms of toxic action of the metal oxide nanoparticles to the HaCaT cell line as well as to the *E. coli* are identified from the developed nano quantitative toxicity-toxicity relationship (nano-QTTR) models.

Keywords: nano-QSAR, nanoparticles, metal oxides, cytotoxicity, HaCaT, *E. coli*

1. Introduction

Commercial application of nanotechnology is escalating rapidly with the development of a wide range of novel nanoparticles (NPs) for various areas of human life (Gajewicz et al., 2012). Ample amount of NPs may also pose anonymous risks to human health and to the environment due to their complex life-cycle. It is often critical to predict probable hazardous impact of new NPs before their mass production and application in the daily life. Regulatory agencies are directly looking for ways of assessing these risks as fast as possible (Zhang et al., 2012). One of the most important groups of engineered nanoparticles includes nano-sized metal oxides. Considering the complexity of nano-sized metal oxides and their abundant applications, it is essential to anticipate their toxicological properties in a fast and economical way in order to alleviate the adverse effects on human health and the environment. There is an urgent need for more efficient methods of testing both required and potentially unwanted properties of NPs (Puzyn et al., 2010). As the experimental evaluations of hazardous properties of NPs are expensive and time consuming, the application of reliable computational methods is one of the key alternatives before their mass production. On the other hand, increasing social and economic pressure to trim down the use of animal testing is an important reason to develop alternative methods which include *in silico* models. Quantitative structure-activity relationship (QSAR) model of nano-materials support the 3Rs

(replacement, refinement and reduction in animals in research) (Benigni and Giuliani, 2003) and serve as a promising approach for toxicity prediction of the nanomaterials. In fact, it is an attractive option to predict toxicological effects directly from the physicochemical properties and chemical structure of the NP entities (Kar and Roy, 2010a).

Encouraging results of investigations confirm that there is a strong need to broaden the traditional quantitative structure-activity/property/toxicity relationship (QSAR/QSPR/QSTR) paradigm to NPs and to develop “nano-QSAR” models which can accurately assess the toxicity of NPs (Puzyn et al., 2011). Notable nano-QSAR models have been established to predict the cellular uptake of NPs in pancreatic cancer cell (PaCa2) by various groups of authors (Epa et al., 2012; Fourches et al., 2010; Ghorbanzadehet al., 2012; Kar et al., 2014a). Rational nano quantitative structure-property relationship (nano-QSPR) models have been discussed based on the catastrophe theory by Carb -Dorca and Besalù (2011). Puzyn *et al.* (2011) employed quantum chemical calculations and Toropov *et al.* (2012) applied web based software CORAL to build QSAR models to describe the cytotoxicity of metal oxide NPs to bacteria *E. coli*. Xia *et al.* (2011) have mapped the surface adsorption forces of NPs for quantitative characterization in biological conditions. Gajewicz *et al.* (2014) developed a nano-QSAR model describing toxicity of 18 nano-metal oxides to a HaCaT cell line as a model for dermal exposure. Liu *et al.* (2013) developed a nano-QSAR model for toxicity of metal oxide nanoparticles (NPs) for human bronchial epithelial (BEAS-2B) and murine myeloid (RAW 264.7) cells. Shao *et al.* (2013) developed QSAR models using nanotoxicity endpoint of surface modified multi-walled carbon nanotubes.

Interspecies quantitative toxicity-toxicity relationship (QTTR) models involve a toxicity endpoint for a particular species acting as a predictor variable along with other descriptors to model same toxicity endpoint against another species (Kar and Roy, 2010b). The toxicity endpoint which acts as a predictor variable can highlight the mechanism of action of a particular compound to some extent as it is derived by standard experimental bioassay, while other descriptors as used in standard QSAR models are obtained purely from chemical structure or physicochemical experiments. This can increase the reliability of an interspecies quantitative correlation models compared to the simple QSAR models for further toxicity prediction. Another fascinating point regarding interspecies correlation is that occasionally toxicity of a compound to a particular biological species is reported but its quantitative toxicity to another biological species is not known. In those cases, interspecies toxicity correlations provide a tool for estimating sensitivity towards toxic chemical exposure with known levels of uncertainty for a diversity of different species (Raimondo et al., 2007).

The aim of the present work is as follows:

- 1) Development of interspecies nano-QTTR models for *E. coli* (prokaryotic) and HaCaT (eukaryotic) cell lines to assess the discriminatory features for the cytotoxicity of metal oxide NPs for both the species. It is probably the first report regarding interspecies nano-QTTR models according to our knowledge.
- 2) The models will be helpful to predict the cytotoxicity of the individual compounds for any one of species when the data for the other species are available,
- 3) Contributing mechanisms of action of the NPs to both species are identified and compared with previously established ones,
- 4) In the background of scarcity of data points for metal oxide toxicity, the present work establishes a new approach for the development and validation of nano-QTTR models. We believe that, this study will contribute for more possibilities in upcoming days for comprehensive risk assessment of nanoparticles when the cytotoxicity data is missing for a specific species.

2. Materials and methods

2.1 Dataset

The cytotoxicity data of 17 nano-metal oxides (Puzyn et al., 2011) to bacteria *E. coli* and 18 nano-metal oxides (Gajewicz et al., 2015a) to HaCaT cell line have been taken from previous experimental works. The results have been expressed in terms of the logarithmic values of molar $1/LC_{50}$ (the concentration of a given oxide causing reduction of bacteria viability of 50%, converted from weight to molar units). Combining the two datasets, common 16 metal oxides were used to develop nano-QTTR models. As CuO has only cytotoxicity value for *E. coli*, and, WO_3 and Mn_2O_3 had cytotoxicity values known for HaCaT cell line, we have used them as true external compounds for prediction of cytotoxicity value to HaCaT cell line and *E. coli*, respectively. It is interesting to point out that the QTTR models developed in this study is based on the experimental data for nano-sized metal oxides, having diameter size ranged from 15 to 90 nm. It is important to mention that the particle size does not influence the toxicity at the studied range of the analyzed oxides (Adams et al., 2006).

2.2 Descriptor calculation

We had calculated as well as collected a set of 34 descriptors quantitatively describing variability of the nanoparticles' structure. These included: 17 quantum-mechanical descriptors, 11 image descriptors (derived from Transmission Electron Microscopy images) and 6 descriptors were taken directly from the publically available periodic table (Supplementary section, Table S1).

2.2.1 Quantum-mechanical descriptors

Calculated 17 quantum-mechanical descriptors reflect electronic properties of the surface of nanoparticles (Supplementary section, **Table S2**). Here, we have calculated electronic properties based on small, stoichiometric clusters, reflecting all characteristics of fragments of crystal structures of particular oxides. All the clusters were of the same size: $0.5 \times 0.5 \times 0.5$ nm. Molecular geometries of each cluster reflecting all characteristics of fragments of crystal structures of particular oxides, were optimized at the level of semi-empirical PM6 method (Stewart, 2007) implemented in the MOPAC 2009 package (Stewart, 2009).

2.2.2 Image descriptors

A set of 11 image descriptors reflecting the size distribution, shape, porosity, and surface area for all studied nano-sized metal and semimetal oxides were calculated based on the TEM microscopic images at the stage of experimental characterization of the nanoparticles (Gajewicz et al., 2015a).

2.2.3 Periodic table descriptors

We have taken some basic information of metal oxides like metal electronegativity (χ), sum of metal electronegativity for individual metal oxide ($\sum \chi$), sum of metal electronegativity for individual metal oxide divided by the number of oxygen atoms present in a particular metal oxide ($\sum \chi / nO$), the charge of the metal cation corresponding to a given oxide (χ_{ox}), atomic number and valence electron of metal as simple molecular descriptors. These descriptors are obtained in virtually no time, which is very time effective as well as cost efficient too (Kar et al., 2014b).

2.3 Nano-QTTR model development and validation

As sixteen metal oxides have cytotoxicity values for both the species, they are considered for models development. Initially without taking any other descriptors, just simple linear regression has been performed between cytotoxicity data for all 16 metal oxides on the two species and it is interesting to point out that the encouraging initial results lead us for development of nano-QTTR models. Subsequently, identification of the most important descriptor from the total pool of 34 descriptors was done using stepwise multiple linear regression (stepwise-MLR) (Darlington, 1990). Along with the selected descriptors, cytotoxicity data to one of the species has been used as a predictor variable for the prediction of cytotoxicity to another species for all 16 metal oxides using partial least squares (PLS) method (Wold, 1995).

According to Hawkins *et al.* (2003), when the available sample size is small, holding a portion of it back for testing is wasteful, and it is much better to use cross-validation, but ensuring that this is done properly. Hawkins suggested that the full model selection procedure (including variable selection) has to be applied to each fold in the validation. This leads to generation of “true q^2 ” which is a better metric for validation than “naïve q^2 ”. Similar to the approach suggested by Hawkins, while calculating r_m^2 (LOO), variable selection process can be applied at each cycle of omission of data points and the resulting metric is called “true r_m^2 (LOO)” (Mitra *et al.*, 2010a). The “true r_m^2 (LOO)” statistic obtained from the model developed from the undivided data set after application of variable selection strategy at each cycle of validation may efficiently reflect the external validation characteristics of the developed model as demonstrated by Mitra *et al.* (2010a). The advantage of using this approach is that the developed model is more reliable due to use of larger number of data points in model development and validation for small dataset, and also that one need not worry about how to efficiently divide the data set into training and test sets.

In our present work, as the number of compounds is small, we have decided to calculate true Q^2 and true r_m^2 (LOO) without losing any information from the dataset as it is already established that true r_m^2 (LOO) efficiently reflects the external validation characteristics. As per the modifications made recently by Roy *et al.* (2010) regarding r_m^2 metric calculations, here we have calculated modified form of true r_m^2 (LOO) that is true $\overline{r_{m(LOO)}^2}$ and true $\Delta r_{m(LOO)}^2$. For the calculation of true $\overline{r_{m(LOO)}^2}$ and true $\Delta r_{m(LOO)}^2$, one compound is deleted from the total dataset of 16 compounds and models are developed with the remaining 15 compounds. Then, deleted compound is predicted with the developed model. This process is repeated until all the compounds are considered as test compounds once. The predicted values of individual compounds are taken for the calculation of true $\overline{r_{m(LOO)}^2}$, true $\Delta r_{m(LOO)}^2$ and true Q^2 . The developed models are also used to predict HaCaT cytotoxicity of CuO (for which *E. coli* cytotoxicity is known) and *E. coli* cytotoxicity of Mn₂O₃ and WO₃ (for which HaCaT cytotoxicity are known). The regression based nano-QTTR models have also been subjected to randomization study employing the metric $^cR_p^2$ (Mitra *et al.*, 2010b).

3. Results and discussion

3.1 Statistical aspects of nano-QTTR models

Linear regression has been performed between cytotoxicity data for all 16 metal oxides of two species and the regression model yielded a squared correlation coefficient (R^2) of 0.42

without using any kind of calculated descriptors. Taking HaCaT cytotoxicity as the response variable and *E. coli* cytotoxicity as predictor variable and *vice versa*, the values of leave-one-out cross-validation parameter (Q^2_{LOO}) are 0.22 and 0.28 respectively. In order to get more statistically significant and interpretable models, along with the thinned descriptors, one species cytotoxicity data is used as a predictor variable for the prediction of cytotoxicity for all 16 metal oxides for the other species using PLS method.

HaCaT and E. coli cytotoxicity correlation

Considering HaCaT cytotoxicity as the response variable and *E. coli* cytotoxicity along with calculated/selected descriptors as predictor variable, we have constructed the following PLS equation (Eq. 1):

$$\begin{aligned}
 pEC_{50}HaCaT &= 0.968 - (0.222 \times \mu) + (0.0003 \Delta H_f^c) + (0.166 \times pEC_{50}E.coli) \\
 n &= 16, LV = 2, F = 466.1, p = 1 \times 10^{-4}, R^2 = 0.88, RMSE_c = 0.14 \\
 Q^2_{LOO} &= 0.80, RMSE_{CV} = 0.22, Q^2_{Leave10\%-out} = 0.80, Q^2_{Leave20\%-out} = 0.81 \\
 Q^2_{Leave25\%-out} &= 0.82, Q^2_{Leave50\%-out} = 0.69, true Q^2 = 0.80 \\
 \overline{true^2_{mLOO}} &= 0.74, true \Delta r^2_{mLOO} = 0.04, R^2_p = 0.79
 \end{aligned} \tag{1}$$

In Eq. (1) H_f^c is the enthalpy of formation of metal oxide nano-cluster representing fragment of the surface and μ is electronic chemical potential of the cluster.

E. coli and HaCaT cytotoxicity correlation

Considering cytotoxicity to *E. coli* as the response variable and cytotoxicity to HaCaT along with calculated descriptors as predictor variable, we have constructed the following PLS equation (Eq. 2):

$$\begin{aligned}
 pEC_{50}E.coli &= 3.494 - (0.0006 \Delta H_{Me+}) - (0.308 \times \chi_{ox}) + (0.345 \times pEC_{50}HaCaT) \\
 n &= 16, LV = 1, F = 139.7, p = 1 \times 10^{-4}, R^2 = 0.91, RMSE_c = 0.19 \\
 Q^2_{LOO} &= 0.88, RMSE_{CV} = 0.17, Q^2_{Leave10\%-out} = 0.88, Q^2_{Leave20\%-out} = 0.87 \\
 Q^2_{Leave25\%-out} &= 0.85, Q^2_{Leave50\%-out} = 0.88, true Q^2 = 0.88 \\
 \overline{true^2_{mLOO}} &= 0.82, true \Delta r^2_{mLOO} = 0.09, R^2_p = 0.80
 \end{aligned} \tag{2}$$

In Eq. (2), H_{Me+} is enthalpy of formation of a gaseous cation and χ_{ox} is the charge of the metal cation corresponding to a particular oxide.

The values of all classical metrics are well within the acceptable limit for both equations. The

observed and predicted values for individual compounds for both species are presented in **Table 1**. It is interesting to point out that, the predicted values are in between minimum and maximum values of observed data for both species. In such a case, we firmly confirmed that the predictions might be considered as results of interpolation, not extrapolation. This result strongly supports that the predictions are reliable and under the confidence limit of applicability domain.

Based on the statistical aspects following observations are found:

(1) As mentioned earlier, due to the small number of compounds, true $\overline{r_{m(LOO)}^2}$ and true $\Delta r_{m(LOO)}^2$ were computed employing the observed and predicted values of individual compounds for the newly developed PLS models (See **Table 2**). Models M1 and M2 have identical values for true $\overline{r_{m(LOO)}^2}$ (0.74), although based on $\Delta r_{m(LOO)}^2$ metric the model M2 (0.04) is superior to the model M1 (0.13). Thus, in terms of true r_m^2 values, model M2 is better than model M1 though the reverse is observed when considering internal validation based on value of true Q^2 value. Hence, it can be summarized that when human cytotoxicity response is considered as the dependent variable, the predictability of the developed PLS models is almost same whether computed descriptors are considered alone or combined with *E. coli* cytotoxicity response as predictor variables. On the contrary, when *E. coli* cytotoxicity response is considered as the dependent variable, the predictability of the developed PLS model (M4) from both the computed descriptors and the human cytotoxicity response as predictor variables is superior to the PLS model (M3) developed with only computed descriptors only in respect of true $\overline{r_{m(LOO)}^2}$ and true Q^2 values.

(2) Again, when the sum of the residuals of the observed and predicted values for all the compounds are considered, the predictability is far better in case of model M2 and model M4 than those for model M1 and model M3, respectively (See **Table S3** and **Table S4** in supporting information).

(3) If the average values of R^2 and Q^2 for individual sixteen models are considered, (which are developed during the calculation of true $\overline{r_{m(LOO)}^2}$, true $\Delta r_{m(LOO)}^2$ and true Q^2), it can be inferred that the average values are almost comparable (when human cytotoxicity response is taken as the dependent variable) or superior (when *E. coli* cytotoxicity response is taken as the dependent variable) when one particular cytotoxicity response is considered as predictor variable along with the computed descriptors (See **Table S5** in supporting information).

A summary of the above results confirms that the model developed employing one of the cytotoxicity responses as predictor variables along with the computed descriptors is more

predictive than the model developed with only computed descriptors. The OECD principle 5 for QSAR model development demands mechanistic interpretation of the developed QSAR models, if possible (OECD, 2004). So, we have provided insight regarding the mechanistic interpretation for both species in **section 3.2**.

As CuO, Mn₂O₃ and WO₃ are not employed in the development of any model, they are considered as true external data points in the present study. The developed models are used to predict HaCaT toxicity of CuO (2.42) from eq. (1) and *E. coli* toxicity of Mn₂O₃ (2.87) and WO₃ (0.17) from eq. (2). Moreover, the visual inspection of the plotted relationship between the experimental and predicted toxicity of the considered metal oxides by equations 1 and 2 (**Figures 1 and 2**) additionally confirms the robustness of the models. For the reliability and acceptability of our prediction, we have tried to compare our predictions of the mentioned nanoparticles with available predictions made from reported models in another study. Gajewicz et al. (2015b) applied nano-read-across model to estimate the cytotoxicity to *E. coli* for untested metal oxide nanoparticles where they have predicted *E. coli* toxicity of Mn₂O₃ is 2.84 in log(EC₅₀)⁻¹ molar scale which is almost comparable with our predicted value of Mn₂O₃ (2.82 in molar scale). Again, they have applied their second nano-read-across model to estimate the toxicity to the HaCaT cell line for experimentally untested metal oxide nanoparticles where cytotoxicity to HaCaT cell lines of CuO is 2.39 (in molar scale) which is also comparable with our predicted value of 2.42 (in molar scale). It is interesting to point out that the models used by Gajewicz et al. (2015b) were single point QSTR model where as our present models are interspecies QSTR model and in both cases the prediction values are very close. Unfortunately we have not found any experimental or predicted cytotoxicity value of WO₃ to HaCaT cell lines. In our future work we will include experimental data of these metal oxide nanoparticles for the respective species once they become available.

3.2 Mechanistic aspects of cytotoxicity

3.2.1 Mechanism prevailing for HaCaT cell lines

Along with the response of *E. coli* toxicity, the descriptors μ and H_f^c are identified as the two most important features that collectively determine the toxicity of the metal oxide nanoparticles to the HaCaT cell line. The first significant descriptor (based on the standardized coefficient) μ is defined as the electronic chemical potential. The energy structure of metal oxides is calculated using chemical reactivity concepts developed within the density functional theory. In this perspective, any chemical system can be characterized by an electronic chemical potential (μ) and absolute hardness (χ). These quantities are reactivity descriptors for isolated chemical species that can be used to represent the electron transfer process of any combined systems (here, metal oxides) (Kohn et al., 1996). The chemical potential characterizes the escaping tendency of the electron cloud from the

equilibrium state, and is equivalent to the negative of Mulliken electronegativity (χ) of an atom:

$$\mu = \left(\frac{\partial E}{\partial N} \right)_Z = -\chi \quad (3)$$

In Eq. (3), E is the ground-state electronic energy, N is the number of electrons and Z is the nuclear charge. By inserting the energy parameters of the ionization potential and electron affinity into the Mulliken electronegativity, it is possible to show that the Mulliken chemical potential is a finite approximation of the difference of electronic energy with respect to the number of electrons (Mulliken, 1935), *i.e.*:

$$\mu(\text{Mulliken}) = -\chi(\text{Mulliken}) = \frac{(E_i + E_{ea})}{2} \quad (4)$$

In Eq. (4), E_i is the first ionization energy and E_{ea} is the electron affinity.

Thus, the electronic chemical potential can be apparently expressed using the Mulliken electronegativity which relates to properties of the cations themselves in the series of the studied oxides. The electronegativity mainly depends on the ionic radius and formal charge of the cation. Since electronegativity is a measure of the tendency of an atom to attract a bonding pair of electrons, it is clear in the context of the Haber-Weiss-Fenton cycle (Koppenol, 2011) that the increase of the cation electronegativity should result in the increase of catalytic properties of metal cations and consequently, it increases the toxicity of the metal oxide nanoparticle.

The second important descriptor H_f° defines the standard enthalpy of formation of metal oxide nanocluster. H_f° corresponds to the energy associated with a single metal-oxygen bond in the oxides (E_{H°) and the number of electrons involved in the formation reaction, since the standard enthalpy of formation H_f° for a given MeOx can be expressed as (Portier et al., 2004):

$$E_{\Delta H^\circ} = -\frac{2 \cdot \Delta H_f^\circ \cdot 2.612 \times 10^{19}}{N_A \cdot n_e} \Rightarrow \Delta H_f^\circ = -\frac{n_e \cdot E_{\Delta H^\circ}}{\text{const.}} \quad (5)$$

In Eq. (5), N_A is the Avogadro number and n_e is the number of electrons involved in the formation reaction. Thus, high absolute values of the enthalpy of formation of the cluster indicate metal oxide nanoparticles with strongly bound cations of large formal charge.

Summarizing characteristics of both descriptors one notices that the first one is closely related to the surface redox activity of the nanoparticle. In this process, the transfer of

electrons from the valence band to the conduction band is caused by the intracellular redox processes occurring in the biological media. On the other side, the second descriptor describes process involves the detachment of metal cations from the surface of metal oxides. Both processes are collectively responsible for inducing oxidative stress in the HaCaT cell lines. In addition, the response of *E. coli* toxicity has positive effects towards the HaCaT toxicity as explained by the developed nano-QTTR model. Though not directly related with the mechanistic interpretation for HaCaT toxicity, it is an important observation considering the model prediction.

3.2.2 Mechanism prevailing for *E. coli*

In terms of mechanistic interpretation of the model, both the computed descriptors along with the HaCaT toxicity response as predictor variables are responsible for the *E. coli* toxicity. The first descriptor H_{Me+} represents the enthalpy of formation of a gaseous cation having the same oxidation state as that in the metal oxide structure:

H_{Me+} descriptor combines both types of mechanism responsible for the cytotoxicity of the oxides which are closely related to the lattice energy, as H_L represents the enthalpy of the reaction:

A negative value of the lattice energy increases with increasing cationic charge (n). Similarly, a positive value of H_{Me+} increases with increasing charge of metal cations. As a result, the release of cations Me^{n+} having smaller charge is more energetically favourable than the release of cations with larger n. The above mechanism can rightly explain the following order of reduction of toxicity of the studied oxides: $Me^{2+} > Me^{3+} > Me^{4+}$. Obviously, much more energy is required to detach four electrons than three (or two) to form the appropriate cations.

The second descriptor χ_{ox} defines the charge of the metal cation corresponding to a given oxide, has a negative coefficient. The cytotoxicity of metal oxide is inversely proportional to the value of χ_{ox} . Those oxides having low χ_{ox} (like CoO, NiO, CuO and ZnO) usually exhibit strong reductive properties that define the easy detachment of the metal cation resulting in enhancement of their cytotoxicity to *E. coli*. As we have hypothesized in this study, for the most toxic metal oxides the small fragments release the electron much easily than similar

fragments build-in the crystal structure. Strongly reductive fragments initiate formation of reactive oxygen species (ROS) resulting in oxidative stress.

Thus, it can be inferred that metal oxides having the lower charge of the metal cation corresponding to a given oxide usually exhibit strong reductive properties. This implies easy detachment of the electrons from the metal oxides which enhances their cytotoxicity to the bacteria *E. coli*. In addition, the response of HaCaT toxicity has positive effect towards the *E. coli* toxicity as explained by the developed nano-QTTR model. It should be specifically mentioned here that for any interspecies model the toxicity of a specific species is not responsible for toxicity of the other species. However, it provides a proper direction and indication for prediction of toxicity for a particular compound, when the corresponding data for any other specific species is missing.

The reported models in a mathematical way strongly support the previously established mechanistic approaches, along with few new revealed hints, and this is the first reported nano-QTTR approach to link two species, providing novelty and substantial innovative applicability of the presented work. Our nano-QTTR work is the first step of developing possibility of predicting *in vivo* results based on *in vitro* studies. Such strategy might, in future, lead to a significant reduction of testing nanoparticles with laboratory animals. Here, it is important to mention that we have not investigated the influence of coating/ligand on toxicity. In our work we have studied only uncoated NPs. So, we do not comment on whether that will affect or not the mentioned parameters. This is a limitation of applicability of our models. But, theoretically, coatings/liagnds can modify the activity, since they are present on the surface. There is already an excellent study investigating the problem elsewhere (Melagraki & Afantitis, 2015).

Another aspect is that the applicability of the developed model is restricted to gram-negative bacteria only and specifically to *E. coli*. As hypothesized in Gajewicz et al. (2015a), differences in toxicity mechanisms between gram-negative bacteria (e.g. here *E. coli*) and eucaryotic cells (HaCaT) is mainly related to the presence of additional membrane and the periplasmatic space between both inner and outer layers. As such, the place where free radicals are formed is different in both cases. It is possible (but not experimentally confirmed yet) that since gram-positive bacteria do not have a double membrane, the mechanism of toxicity would be different, and it may be even more similar to that for eucaryotic cells than to that for gram-negative bacteria. However, it will be too early to comment that the developed models will also be applicable for cytotoxicity of metal oxide nanoparticles to gram positive bacteria without any experimental evidence. The developed models should be applicable with confidence for prediction of cytotoxicity in *E. coli*.

4. Conclusions

The nano-QTTR approach has been developed for the first time according to our knowledge to model very complex biological end point, cytotoxicity of metal oxides nanoparticles. Major concluding remarks of the present work can be summarized as follows:

1. The present work mainly proposes mathematical expressions to extrapolate the cytotoxicity data of nanoparticles for *E. coli* to that for HaCaT cell lines and *vice versa*. To predict HaCaT cytotoxicity, electronic chemical potential (μ) and standard enthalpy of formation of metal oxide nanocluster (H_f°) are proved to be the best encoding features in addition to the experimental cytotoxicity of *E. coli* in terms of both statistical significance and mechanistic interpretability. The first process involves the transfer of electrons from the valence band to the conduction band by the intracellular redox processes and the second process involves the detachment of metal cations from the surface of metal oxides. A consensus of both processes collectively form highly reactive and less specific hydroxyl radicals inducing oxidative stress in the cells. On the other hand, to predict the *E. coli* cytotoxicity, both enthalpy of formation of a gaseous cation (H_{Me+}) and the charge of the metal cation corresponding to a given oxide (χ_{ox}) serve as the most significant attributes along with the experimental cytotoxicity of human cell lines. The release of cations Me^{n+} having smaller charge is energetically more favourable than the release of cations with larger n. Again, metal oxides having lower charge on the metal cation corresponding to a given oxide usually exhibit strong reductive properties signifying easy detachment of the electrons from the metal oxides which in turn enhances their cytotoxicity to the bacteria *E. coli*.

2. As the number of available cytotoxicity data points for the metal oxide nanoparticles is very small, the models have been selected based on the true $\overline{r_{m(LOO)}^2}$, true $\Delta r_{m(LOO)}^2$ and true Q^2 without dividing the dataset into training and test set. The major advantage behind the calculation of these metrics is that there is a no chance of data loss and these parameters have the ability for the external predictability.

3. Extrapolation as well as prediction of cytotoxicity of metal oxide nanoparticles can be better done if the equations are expressed as follows:

$$E.coli cytotoxicity = f(Human cytotoxicity, \Delta H_{Me+}, \chi_{OX}) \quad (8)$$

and

$$Human cytotoxicity = g(E.coli cytotoxicity, \Delta H_f^\circ, \mu) \quad (9)$$

The above equations (correlating toxicity data of two different species) strongly suggest that the consideration of the species specific cytotoxicity of metal oxide nanoparticles as predictor variables along with other computed descriptors yields better models for prediction of the metal oxides to other species than the models developed using only the computed descriptors.

Acknowledgement

The authors are grateful for the financial support of the Polish National Science Center (grant no. UMO-2011/01/M/NZ7/01445). SK thanks the Department of Science and Technology (DST), Government of India for awarding him a Research fellowship under the INSPIRE scheme. AG thanks the European Social Fund for granting her with a fellowship in frame of the project “Development Program of the University of Gdansk in areas of Europe 2020 (UG 2020)” supported by Human Capital Operational Programme, Action 4.3, Strengthening of didactic potential of universities in key areas in the context of the goals of the Europe 2020 strategy (Grant No. UDA-POKL. 04.03.00-00-047/12). KR thanks the Council of Scientific and Industrial Research (CSIR), New Delhi for awarding a major research project. J.L. acknowledges support from the National Science Foundation ([CREST HRD-0833178](#) and EPSCoR 362492-190200-01\NSFEPS-0903787).

References

- Adams, L.K., Lyon, D.Y., Alvarez, P.J., 2006. Comparative eco-toxicity of nanoscale TiO₂, SiO₂, and ZnO water suspensions. *Water Res.* 40, 3527-3532.
- Benigni, R., Giuliani, A., 2003. Putting the Predictive Toxicology Challenge into Perspective: Reflections on the Results. *Bioinformatics* 19, 1194-1200.
- Carb -Dorca, R., Besalù, E., 2011. Construction of coherent nano quantitative structure-properties relationships (nano-QSPR) models and catastrophe theory. *SAR QSAR Environ. Res.* 22, 661-665.
- Darlington, R.B., 1990. *Regression and Linear Models*, McGraw Hill, New York.
- Epa, V.C., Burden, F.R., Tassa, C., Weissleder, R., Shaw, S., Winkler, D.A., 2012. Modeling Biological Activities of Nanoparticles. *Nano Lett.* 12, 5808-5812.
- Fourches, D., Pu, D., Tassa, C., Weissleder, R., Shaw, S.Y., Mumper, R.J., Tropsha, A., 2010. Quantitative nanostructure-activity relationship modeling. *ACS Nano* 4, 5703-5712.
- Gajewicz, A., Rasulev, B., Dinadayalane, T., Urbaszek, P., Puzyn, T., Leszczynska, D., Leszczynski, J., 2012. Advancing risk assessment of engineered nanomaterials: Application of computational approaches. *Adv. Drug Delivery Rev.* 64, 1663-1693.

Gajewicz, A., Schaeublin, N., Rasulev, B., Hussain, S., Leszczynska, D., Puzyn, T., Leszczynski, J., 2015a. Towards Understanding Mechanisms Governing Cytotoxicity of Metal Oxides Nanoparticles: Hints from Nano-QSAR Studies. *Nanotoxicology* 9, 313-325.

Gajewicz, A., Cronin, M.T.D., Rasulev, B., Leszczynski, J., Puzyn, t., 2015b. Novel approach for efficient predictions properties of large pool of nanomaterials based on limited set of species: nano-readacross. *Nanotechnology* 26, 015701.

Ghorbanzadeh, M., Fatemi, M.H., Karimpour, M., 2012. Modeling the Cellular Uptake of Magnetofluorescent Nanoparticles in Pancreatic Cancer Cells: A Quantitative Structure Activity Relationship Study. *Ind. Eng. Chem. Res.* 51, 10712-10718.

Hawkins, D.M., Basak, S.C., Mills, D., 2003. Assessing model fit by cross-validation, *J. Chem. Inf. Comput. Sci.* 43, 579-586.

Kar, S., Gajewicz, A., Puzyn, T., Roy, K., 2014a. Nano-quantitative structure–activity relationship modeling using easily computable and interpretable descriptors for uptake of magnetofluorescent engineered nanoparticles in pancreatic cancer cells. *Toxicol. in Vitro.* 28, 600-606.

Kar, S., Gajewicz, A., Puzyn, T., Roy, K., Leszczynski, J., 2014b. Periodic table-based descriptors to encode cytotoxicity profile of metal oxide nanoparticles: A mechanistic QSTR approach. *Ecotox. Environ. Safe.* 107, 162-169.

Kar, S., Roy, K., 2010a. Predictive toxicology using QSAR: a perspective, *J. Indian Chem. Soc.* 87, 1455-1515.

Kar, S., Roy, K., 2010b. First report on interspecies quantitative correlation of ecotoxicity of pharmaceuticals. *Chemosphere* 81, 738-747.

Kohn, W., Becke, A.D., Parr, R.G., 1996. Density Functional Theory of Electronic Structure, *J. Phys. Chem.* 100, 12974-12980.

Koppenol, W.H., 2011. The Haber-Weiss cycle-70 years later. *Redox Report* 6, 229-234.

Liu, R., Zhang, H.Y., Ji, Z.X., Rallo, R., Xia, T., Chang, C.H., Nel, A., Cohen, Y., 2013. Development of structure-activity relationship for metal oxide nanoparticles, *Nanoscale* 5, 5644-5653.

Melagraki, G., Afantitis, A., 2015. A Risk Assessment Tool for the Virtual Screening of Metal Oxide Nanoparticles through Enalos InSilicoNano Platform. *Curr. Top. Med. Chem.* 15, 1827-1836.

Mitra, I., Roy, P.P., Kar, S., Ojha, P.K., Roy, K., 2010a. On further application of rm2 as a metric for validation of QSAR models. *J. Chemometrics* 24, 22-33

Mitra, I., Saha, A., Roy, K., 2010b. Exploring quantitative structure-activity relationship (QSAR) studies of antioxidant phenolic compounds obtained from traditional Chinese medicinal plants. *Mol. Simul.* 36, 1067-1079.

Mulliken, R.S., 1935. Electronic Structures of Molecules XI. Electroaffinity, Molecular Orbitals and Dipole Moments. *J. Chem. Phys.* 3, 573-585.

OECD, 2004. 37th Joint Meeting of the Chemicals Committee and Working Party on Chemicals, Pesticides and Biotechnology, Organisation for Economic Co-Operation and Development, Paris, France.

Portier, J., Hilal, H.S., Saadeddin, I., Hwang, S.J., Subramanian, M.A., Campet, G., 2004. Thermodynamic correlations and band gap calculations in metal oxides. *Prog. Solid State Chem.* 32, 207-217.

Puzyn, T., Gajewicz, A., Leszczynska, D., Leszczynski, J., 2010. In: Puzyn, T., Leszczynski, J., Cronin, M., (Eds), *Recent Advances in QSAR Studies: Methods and Applications*, M.T. Springer: London, New York, 2010.

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.

Raimondo, S., Mineau, P., Barron, M.G., 2007. Estimation of chemical toxicity to wildlife species using interspecies correlation models. *Environ. Sci. Technol.* 41, 5888-5894.

Roy, K., Mitra, I., Kar, S., Ojha, P., Das, R.N., Kabir, H., 2012. Comparative studies on some metrics for external validation of QSPR models. *J. Chem. Inf. Model* 52, 396-408.

Shao, C.Y., Chen, S.Z., Su, B.H., Tseng, Y.J., Esposito, E.X., Hopfinger, A.J., 2013. Dependence of QSAR Models on the Selection of Trial Descriptor Sets: A Demonstration Using Nanotoxicity Endpoints of Decorated Nanotubes. *J. Chem. Inf. Model.* 53, 142-158.

Stewart, J.J.P., 2007. Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* 13, 1173-1213.

Stewart, J.J.P., 2009. MOPAC: Stewart Computational Chemistry. <http://openmopac.net/MOPAC2009.html> (accessed April 02, 2013).

Toropov, A.A., Toropova, A.P., Benfenati, E., Gini, G., Puzyn, T., Leszczynska, D., Leszczynski, J., 2012. Novel application of the CORAL software to model cytotoxicity of metal oxide nanoparticles to bacteria *Escherichia coli*. *Chemosphere* 89, 1098-1102.

Wold, S., 1995. PLS for multivariate linear modelling, In: van de Waterbeemd, H., (Ed.), *Chemometric Methods in Molecular Design*, VCH, Weinheim, Germany, pp. 195-218.

Xia, X.R., Monteiro-Riviere, N.A., Mathur, S., Song, X., Xiao, L., Oldenberg, S.J., Fadeel, B., Riviere, J.E., 2011. Mapping the Surface Adsorption Forces of Nanomaterials in Biological Systems. *ACS Nano* 5, 9074-9081.

Zhang, H., Ji, Z., Xia, T., Meng, H., Low-Kam, C., Liu, R., Pokhrel, S., Lin, S., Wang, X., Liao, Y.P., Wang, M., Li, L., Rallo, R., Damoiseaux, R., Telesca, D., Mädler, L., Cohen, Y.,

549 Zink, J.I., Nel, A.E., 2012. Use of metal oxide nanoparticle band gap to develop a predictive
550 paradigm for oxidative stress and acute pulmonary inflammation. ACS Nano 6, 4349-4368.

551
552 **Figure captions**

553 **Figure 1.** Plot of experimentally determined (observed) versus calculated (predicted) pEC_{50}
554 of HaCaT.

555 **Figure 2.** Plot of experimentally determined (observed) versus calculated (predicted) pEC_{50}
556 of *E. coli*.

TABLES

Table 1. Observed and predicted values for individual compounds for HaCaT cell lines and *E. coli*.

Metal	Observed pEC_{50} of HaCaT (in molar scale)	Predicted pEC_{50} of HaCaT (in molar scale)	Observed pEC_{50} of <i>E. coli</i> (in molar scale)	Predicted pEC_{50} of <i>E. coli</i> (in molar scale)
Al ₂ O ₃	1.85	1.98	2.49	2.50
Bi ₂ O ₃	2.5	2.58	2.82	2.75
CoO	2.83	2.97	3.51	3.49
Cr ₂ O ₃	2.3	2.28	2.51	2.60
Fe ₂ O ₃	2.05	2.17	2.29	2.43
In ₂ O ₃	2.92	2.92	2.81	2.81
La ₂ O ₃	2.87	2.83	2.87	2.95
NiO	2.49	2.55	3.45	3.38
Sb ₂ O ₃	2.31	2.33	2.64	2.63
SiO ₂	2.12	1.99	2.2	1.98
SnO ₂	2.67	2.24	2.01	2.15
TiO ₂	1.76	1.90	1.74	1.92
V ₂ O ₃	2.24	2.17	3.14	2.68
Y ₂ O ₃	2.21	2.15	2.87	2.83
ZnO	3.32	3.26	3.45	3.63
ZrO ₂	2.02	2.23	2.15	2.14

591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612

Table 2. Comparison of true $\overline{r^2_{m(LOO)}}$, true $\Delta r^2_{m(LOO)}$ and true Q^2 metrics among developed PLS models.

Response	Model	Descriptors	True $\overline{r^2_{m(LOO)}}$	True $\Delta r^2_{m(LOO)}$	True Q^2
<i>pEC₅₀</i> for HaCaT	M1	μ, Hfc (LV=1)	0.74	0.13	0.83
	M2	μ, Hfc <i>pEC₅₀</i> of <i>E. coli</i> (LV=2)	0.74	0.04	0.80
<i>pEC₅₀</i> for <i>E. coli</i>	M3	H _{Me+} , χ _{ox} (LV=1)	0.79	0.08	0.85
	M4	H _{Me+} , χ _{ox} , <i>pEC₅₀</i> of HaCaT (LV=1)	0.82	0.09	0.88

FIGURES

FIG 1

FIG 2