Environmental
Science
Nano

Second Generation Periodic Table Based Descriptors to Encode Toxicity of Metal Oxide Nanoparticles to Multiple Species: QSTR Modeling for Exploration of Toxicity Mechanisms

SCHOLARONE™
Manuscripts

Environmental Significance Statement

Followed by the success of our proposed periodic table based descriptors (PTDs) in the modeling study, we have reported a pool of second generation PTDs for better understanding the mechanism of toxicity of metal oxide nanoparticles (MNPs) towards three species namely *E. coli*, human keratinocyte cell line (HaCaT) and Zebrafish embryos employing QSTR models. The QSTR models can be efficiently employed for environmental risk screening tools for the mentioned species for any new/untested MNPs. The developed i-QSTTR models will allow in extrapolation of data from one species to another. Further a dataset of 42 MNPs was used as the true external dataset for prediction using the developed single line models for the very first time which will help in environmental risk assessment and data gap filling as these metal oxides are never tested for these mentioned species.

Second Generation Periodic Table Based Descriptors to Encode Toxicity of Metal Oxide Nanoparticles to Multiple Species: QSTR Modeling for Exploration of Toxicity Mechanisms

Priyanka De[a], Supratik Kar[b], Kunal Roy[a*], Jerzy Leszczynski[b*]

[a]Drug Theoretics and Cheminformatics Laboratory,
Department of Pharmaceutical Technology, Jadavpur University,
Kolkata 700032, India

[b]Interdisciplinary Center for Nanotoxicity,
Department of Chemistry, Physics and Atmospheric Sciences,
Jackson State University, Jackson, MS-39217, USA

*Corresponding author

Prof. Kunal Roy, Phone: +91 98315 94140; Fax: +91-33-2837-1078;

Email: kunalroy_in@yahoo.com; kunal.roy@jadavpuruniversity.in

Prof. Jerzy Leszczynski, Phone: +1 601 979 3723; fax: +1 601 979 7823;

E-mail: jerzy@icnanotox.org

1

Abstract

The application of *in silico* methods in the risk assessment of metal oxide nanoparticles (MNPs) and data gap filling has found profound usability. Followed by the success of periodic table-based descriptors in the modeling study, we have reported here a pool of second generation periodic table descriptors for better understanding the mechanism of toxicity of MNPs towards three species namely *E. coli*, human keratinocyte cell line (HaCaT) and Zebrafish embryos. These descriptors are easily derived from the molecular formula and periodic table in no time and can be used in models with prediction ability similar to or even better than those involving quantum chemical descriptors and physicochemical indices. Employing the developed 1st and 2nd generation periodic table-based descriptors, we have developed single species quantitative structure-toxicity relationship (QSTR) models and interspecies quantitative structure-toxicity-toxicity relationship (i-QSTTR) models to understand the relationship among the toxicities of metal oxide nanoparticles to different species along with the identification of the major mechanism(s) for such toxicities. These models further helped in extrapolating toxicity when the data for one species is available and the data for other species are unavailable. Further, we have made predictions for a set of 42 true external MNPs, employing all three QSTR models individually, the toxicity data for all three species using a two-stage prediction confidence check through applicability domain and "prediction reliability indicator". The developed models interpreted that the oxidation state of the metal, the electronegativity of the metal oxide and core count of the metal play substantial roles in the toxicity mediation of the MNPs irrespective of the species and response. Along with the first generation, the newly developed second generation periodic table-based descriptors can encode the toxicity features of MNPs efficiently as compared to classical quantum chemical descriptors involving time consuming computations and physicochemical descriptors involving experimental tests.

Keywords: *E. coli*, HaCaT, Metal Oxide, Nanoparticles, Periodic Table, QSTR, Zebrafish

1. Introduction

In the present days, nanotechnology has evolved in the frontline of modern science and technology with a vast range of applications of nanoparticles (NPs) in personal, commercial, medical, military, pharmacy, textiles, cosmetics, electronics, catalysts and self-cleaning coatings, to list a few.[1,2] Metal oxide nanoparticles (MNPs) have found immense use due to their novel optical, magnetic, and electronic properties.[3] In spite of the extensive beneficial aspects, nanomaterials pose a serious threat to the ecosystem due to their continuous release during their production process as well as from the nanomaterials containing products during use, recycling and disposal.[4] Nanomaterials possess unique characteristics having a diverse range of size, shape, chemical compositions and surface modifications which all directly and indirectly influence the toxicity profile.[5] Thus, the toxicity of nanomaterials to diverse organisms in different environmental compartments has become a noteworthy concern.

The MNPs have been shown to induce inflammation and oxidative stress, and changes in cell signaling and gene expression in mammalian cells. As the nanotechnology industry increases day-by-day, nanoscale products and by-products are penetrating the aquatic environment posing a serious threat to aquatic organisms.[6] Predictive toxicity models could form an integral component of an approach which can predict the types of nanomaterials responsible for creating environmental toxicity as a result of their physico-chemical characteristics. Different

computational approaches have been developed over the years, but Quantitative Structure-Activity/Property/Toxicity Relationship (QSAR/QSPR/QSTR) modeling is the most convenient, time effective and inexpensive one.[7-10] Additionally, predictive QSAR modeling circumvents the need to utilize animal models and has been proven to be an efficient tool for predicting the potentially adverse effects of chemical entities in terms of risk assessment, chemical screening, and priority setting.[11,12]

The QSPR/QSTR methodologies have been used to correlate diverse properties/toxicities of nanomaterials over the years.[13-20] A combination of basic physics with toxicity mechanism was utilized, where the "liquid drop" approach was applied to develop predictive classification models for toxicity of MNPs.[20] Nano-read-across methodology was proposed by Gajewicz *et al.*[21] which in the absence of adequate and relevant data has been used in predicting toxicities and properties. Boukhvalov and Yoon proposed first-principle calculations based theoretical descriptors to model cytotoxicity of metallic NPs.[22] Shin et al.[23] suggested spherical cluster and hydroxyl metal coordination complex to compute descriptors for the generation of classification models for NP cytotoxicity. A good number of articles deal with QSTR perturbation model to predict ecotoxicity and cytotoxicity of diverse NPs under changed experimental conditions as well as changing the chemical composition of nanoparticles size, shapes and conditions against several endpoints. These QSTR models employed moving average approach to compute descriptors to encode the toxicity responses.[24-26] Researchers have demonstrated the relationship between zeta potential and a series of intrinsic physico-chemical features of 15 MNPs revealed by a computational study.[27] Xia *et al.*[28] have mapped the surface adsorption forces of 16 MNPs for quantitative classification in biological systems.

Interspecies quantitative structure–toxicity–toxicity (i-QSTTR) modeling allows the prediction of toxicity of a species using experimental toxicity values of another species as a descriptor along with chemical structure derived/physicochemical descriptors. This type of modeling might reduce the use of higher level organisms and also helps in understanding the mechanism of action to some degree as it is derived by a standard experimental bioassay.[29] Many groups of researchers across the globe have carried out interspecies modeling on different endpoints.[30-33] Since the toxicity end point, which acts a predictor variable, is obtained through some standard experimental bioassay, it can to some extent describe the mechanism of action of a particular compound while the other predictor variables are solely obtained from the compounds' chemical structures and physicochemical behavior. Also it is not possible to carry out toxicity experiments on all species for every compounds, so interspecies toxicity correlations provide a tool for estimating sensitivity towards toxic chemical exposure with known levels of uncertainty for a diversity of different species and for bridging data gaps.[34] Periodic table-based descriptors were utilized by Kar *et al.* to find interspecies relationship between *E.coli* and HaCaT cell line towards MNPs toxicity for the very first time.[35] Basant and Gupta developed robust and reliable interspecies models to predict the cellular affinity of MNPs for multiple human cell types (PaCa2, HUVEC, RestMph, GMCSF_Mph, and U937).[36]

Scientists have developed several models towards toxicity of MNPs on various cell lines using quantum chemical descriptors[14] which provided considerable and reliable results. However, the use of quantum descriptors requires high computational time and efficient personnel. The introduction of periodic table based descriptors[35] helped in reducing time required for

computation because they were directly obtained from the periodic table or derived from it and also these descriptors were effective enough to produce similar or even better results as compared to quantum-based descriptors. In the present work, we have developed QSTR models on cytotoxicity profile of MNPs to bacteria *Escherichia coli* (prokaryotic system) and human keratinocyte cell line (HaCaT) (eukaryotic system) and also on enzyme inhibition of Zebrafish Hatching Enzyme (ZHE1) employing a new set of second generation of periodic table descriptors. Further we have demonstrated in the present work the use of periodic table-based descriptors for an external set of MNPs for which toxicity and enzyme inhibition data are not available (data gap filling) for the mentioned species. We have also developed i-QSTTR models that allow the prediction of a specific toxicity (for an endpoint) using the experimental toxicity values (for another endpoint) as a descriptor along with chemical descriptors. These models will be helpful to predict the cytotoxicity of the individual metal oxide nanoparticles for any one of species when the toxicity data for the other species are available. The outline of the study is demonstrated in **Figure 1**.
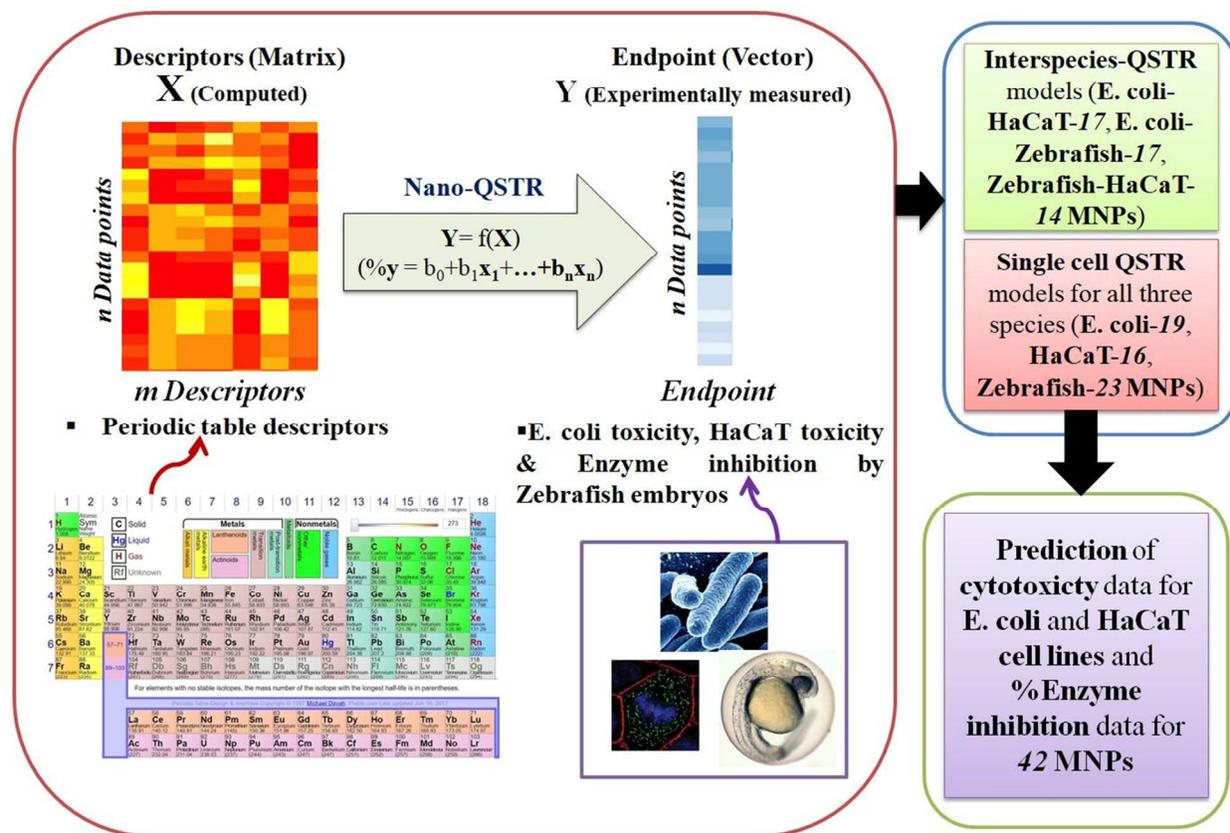

Figure 1 The complete scheme of the study.

2. Materials and methods

2.1. Dataset

The cytotoxicity data of 19 MNPs[14] to bacteria *E. coli*, 18 MNPs[37] to human keratinocyte cell line (HaCaT) and percentage decrease in enzymatic activity (enzyme inhibition to Zebrafish in % (%EI_Zebrafish)) of Zebrafish hatching enzyme (ZHE1) of 24 MNPs[38] were employed in the

4

present modeling studies. The data points are expressed in terms of pEC_{50} (negative logarithm of EC_{50} expressed in a molar unit) in case of toxicity against *E. coli* and HaCaT cell line while in form of percentage enzyme inhibition in case of Zebrafish embryo toxicity. In case of the *E. coli* cytotoxicity models, all the MNPs were utilized for model development. But for HaCaT cytotoxicity models, three MNPs (NiO, Mn_2O_3, SiO_2) and for ZHE1 enzyme inhibition models, one MNP (CoO) was removed for their outlier behavior based on preliminary modeling analysis and similarity measures. In case of interspecies QSTTR models, cytotoxicity data against *E. coli* and HaCaT cell lines had 19 MNPs in common and all these were used in the modeling analysis. For interspecies modeling considering *E. coli* cytotoxicity and ZHE1 enzyme inhibition in Zebrafish, 17 MNPs were in common and all of these data were used for interspecies toxicity modeling. For HaCaT cytotoxicity and Zebrafish embryo toxicity, 16 MNPs were common but out of these 3 compounds (CoO, Cr_2O_3, WO_3) had to be removed for i-QSTTR modeling based on the initial analysis.

2.2. Descriptor calculation

The QSTR and interspecies QSTTR models were developed by using fundamental information of the metal oxides obtained from the periodic table to encode their toxicity against *E. coli*, HaCaT and Zebrafish. A set of 23 descriptors were used among which 10 descriptors were directly collected from the periodic table and the rest were derived from the primary descriptors. The list of all collected and derived descriptors are given in the **Table 1**. The advantage of periodic table-based descriptors is that they are obtained quickly without any significant calculations and software usage unlike quantum chemical descriptors. First seven descriptors were used by us previously which are marked as first generation periodic table descriptors.[35] Newly introduced sixteen descriptors used in the present study are denoted as second generation periodic table descriptors. Computed descriptors for all MNPs are reported in the Supplementary Information excel file.

Table 1 List of descriptors used for model development.

No.	Generation of PT Descriptors	Mathematical Expression	Description
1	*1ˢᵗ Generation*	MW	Molecular weight of metal oxide
2		N_{metal}	Number of metal
3		N_{oxy}	Number of oxygen
4		χ	Metal electronegativity
5		$\sum\chi$	Total metal electronegativity in specific metal oxide
6		$\sum\chi/nO$	Total metal electronegativity in specific metal oxide relative to number of oxygen
7		χ_{ox}	Oxidation number of metal
8	*2ⁿᵈ Generation*	Z_{metal}	Atomic number of metal
9		Z^v_{metal}	Valence electron of metal
10		PN_{metal}	Period number of metal
11		$\lambda=(Z_{metal}-Z^v_{metal})/Z^v_{metal}$	Core environment of metal defined by the ratio of the

			number of core electrons to the number of valence electrons
12		$\mu = 1/(PN_{metal})$	-
13		V_{metal}	Valence of metal
14		$\alpha_{metal} = \lambda*\mu$	-
15		$\sum\alpha_{metal} = \alpha_{metal}*N_{metal}$	-
16		$\sum\alpha_{oxy} = N_{oxy}*0.33$	-
17		$\sum\alpha = \sum\alpha_{metal} + \sum\alpha_{oxy}$	The core count, gives a measure of the molecular bulk
18		$\varepsilon_{metal} = -\alpha_{metal} +(0.3*Z^{v}_{metal})$	Electronegativity count of metal
19		$\varepsilon_{oxy} = -\alpha_{oxy} + (0.3*Z^{v}_{oxy})$	Electronegativity count of oxygen
20		$\sum\varepsilon = \varepsilon_{metal}* N_{metal} + \varepsilon_{oxy}* N_{oxy}$	Electronegativity count of total metal oxide
21		$\sum\varepsilon/N$	Sum epsilon relative to number of atoms in the molecule
22		$(\sum\alpha)^2$	Square of summation of alpha, gives measure of molecular bulk
23		$(\sum\varepsilon/N)^2$	Squared sum epsilon by number of atoms

2.3. Splitting of the dataset

Selection of training and test sets plays a vital role in the development of QSTR models. The total number of data points as well as their combination of metal oxides are different for each cell lines. Not only that, the modeled responses are also different from each other. In the present work, the dataset for each response was divided into a training set and a test set based on the principle component analysis (PCA) score plot ensuring distribution and similarity measures of properties and response uniformly in both sets. This is why the composition of the training and test sets changes from model to model in this study. The PCA plots for single endpoint QSTR models as well as for interspecies models are placed in the Supplementary Information excel file as **Figures S1** and **S2**, respectively.

2.4. Chemometric tools

The QSTR models and interspecies QSTTR models were developed using Multiple Linear Regression (MLR)[7] method by Genetic Algorithm[39] technique of descriptor selection. In some cases, partial least squares (PLS)[40] regression was used from the descriptors selected in the stepwise approach due to lower number of data points. In case of PLS regression, to avoid overfitting, a strict test for the significance of each consecutive PLS component is necessary and then stopping when the new components are non-significant. For all the developed models, we have maintained the acceptable ration of number of data points and number of adjustable parameters.[7]

2.5. Validation metrics

The validation of the models was done by both internal and external validation metrics.[7] The determination (R^2) measures the fitting potential of the model whereas internal validation (which

deals with the predictive ability of the model based on training set compounds) is usually determined by a cross-validated squared correlation coefficient, $Q_{LOO}{}^2$ (leave-one-out or LOO). Although Q^2_{LOO} measures the model robustness, is not sufficient to determine the performance of the model when new sets of compounds are employed. The external validation deals with measuring the predictive ability of a model for the test set compounds. The external validation of the model was estimated using the $Q_{F1}{}^1$ and $Q_{F1}{}^2$ parameter.[41] The accepted threshold value for both internal (Q^2_{LOO}) and external predictive parameters (Q^2_{F1}, Q^2_{F2}) is 0.5.[41] Additionally, $r_m{}^2$ metric values were checked as these metrics were found more stringent than conventional metrics.[42] Additional parameters like Q^2F_3, RMSE, MAE values are also computed. [41]

2.6. Applicability domain
Applicability domain (AD) is a theoretical region in chemical space defined by the respective model descriptors and responses in which the predictions are reliable. In the present study, the AD is calculated by two different approaches: a) Leverage approach[43] and b) the standardization[44] approach to find out the influential observation for training set compounds and outliers for test set compounds whose predictions are not reliable.

2.7. True External Set Prediction
To verify the predictive power of the developed QSTR models, a set of MNPs were used as true external set.[21] The experimental data of these compounds were not available for these MNPs and hence used for predicting their toxicities on the three diverse species modeled (*viz. E.coli*, HaCaT and Zebrafish). Additionally, the domain of applicability and their prediction reliability were also checked by the standardization approach and with the '*prediction reliability indicator*' tool[45] respectively.

3. Result and discussion
Depending on the cytotoxicity and enzyme inhibition data and calculated periodic table-based descriptors, we have developed statistically significant QSTR models using GA-MLR and PLS methods.

3.1 Modeling of MNPs cytotoxicity to *E. coli*
The best equation is developed employing GA-MLR approach for the *E. coli* which is described as equation 1 mentioned below:

$$pEC_{50}Ecoli = 4.998(\pm0.349) - 0.757(\pm0.106)\chi_{ox} - 0.020(\pm0.015)(\textstyle\sum\varepsilon/N)^2$$

$$N_{train} = 12, R^2 = 0.88, R^2_{adj} = 0.85, Q^2_{(LOO)} = 0.78, \overline{r^2_{m(loo)}} = 0.69, \Delta r^2_{m(loo)} = 0.14$$

$$N_{test} = 7, Q^2_{F1} or R^2_{pred} = 0.83, Q^2_{F2} = 0.70, \overline{r^2_{m(test)}} = 0.67, \Delta r^2_{m(test)} = 0.16 \qquad (1)$$

The values within parentheses indicate standard error of respective regression coefficients. Although the second term $(\sum\varepsilon/N)^2$ is significant only at 78% confidence level, we have retained it in the equation as it shows its importance while predicting the test set data. From equation 1, we can infer that the model could explain 85% of the variance ($R_a{}^2$) while it could predict 78% of the cross-validated leave-one-out predicted variance (Q^2). The metrics Q^2_{F1} or R^2_{pred} and Q^2_{F2} are used for external validation purposes and their values are much higher than the acceptable

7

threshold value of 0.5. Since the number of modeled compounds is low, we have calculated $\overline{r^2_{m(loo)}}$ and $\Delta r^2_{m(loo)}$ employing the observed and LOO-predicted values of the training set followed by $r^2_{m(test)}$ and $\Delta r^2_{m(test)}$ for test set molecules. The statistical quality of the developed model is illustrated in **Table 2**. Additional validation metrics like Q^2F_3, RMSE, MAE values are placed in **Table S1** in Supplementary information file. The experimental and predicted toxicity values of MNPs against *E. coli* obtained from developed equation 1 are presented in **Table 3**. The scatter plot (**Figure 2,** Top left) showed that the points were limitedly scattered around the line of fit for both training and test sets. We have performed descriptor-descriptor intercorrelation study. Based on the results of the analysis, there is no significant intercorrelation among the descriptors for any single endpoint and/or interspecies models. The obtained result is reported in **Table S2** in Supplementary Information excel file. Similarly, the correlation among modeled descriptors and responses are computed to show the importance of specific descriptor to a particular model and illustrated in **Table S3** in Supplementary Information excel file.

The descriptors contributing to the toxicity of MNPs towards *E. coli* are χ_{ox} and $(\sum \varepsilon/N)^2$. The descriptor χ_{ox} is the oxidation number of the metal atom and has a negative coefficient towards the toxicity. If we closely study the variation of the cytotoxicity data with the numerical value of χ_{ox}, we can easily infer that with an increasing value for the descriptor the cytotoxicity decreases. This is well observed in case of copper oxide (CuO) where the oxidation state of Cu is 2 and it has a high cytotoxicity. In case of ZrO_2, the oxidation state is high ($\chi_{ox} = 4$) and its cytotoxicity is low ($pEC_{50} = 2.15$) which proves the reliability of the model. The values of χ_{ox} for the four most (CoO, NiO, CuO and ZnO) and least cytotoxic (SiO_2, ZrO_2, SnO_2 and TiO_2) compounds are two and four, respectively. All the metal oxides having χ_{ox} value of three showed moderate cytotoxicity which is no doubt a noteworthy observation.

Table 2 Statistical parameters showing the quality of the developed models.

Dependent variable	Statistical methods	Descriptors	N_{Train}	R^2	R_{adj}^2	$Q_{(LOO)}^2$	$\overline{r_{m(loo)}^2}$	$\Delta r_{m(loo)}^2$	N_{Test}	Q_{F1}^2	Q_{F2}^2	$\overline{r_{m(test)}^2}$	$\Delta r_{m(test)}^2$
Single line QSTR models													
pEC_{50}E.coli	GA-MLR	χ_{ox}, $(\sum\varepsilon/N)^2$	12	0.88	0.85	0.78	0.69	0.14	7	0.83	0.70	0.67	0.16
pEC_{50}HaCaT	PLS (1 LV)	$\sum\chi/nO$, λ, V_{metal}	13	0.78	0.70	0.61	0.51	0.16	3	0.83	0.83	0.67	0.12
%EI_Zebrafish	GA-MLR	N_{oxy}, $\sum\chi/nO$, $\sum\alpha_{metal}$	16	0.83	0.79	0.68	0.61	0.6	7	0.74	0.70	0.54	0.22
Interspecies-QSTTR models													
pEC_{50}E.coli	PLS	χ_{ox}, pEC_{50}HaCaT (1 LV)	12	0.78	0.73	0.67	0.58	0.11	5	0.91	0.82	0.70	0.10
pEC_{50}E.coli	GA-MLR	χ_{ox}, %EI_Zebrafish	11	0.83	0.79	0.74	0.65	0.12	6	0.91	0.90	0.84	0.08
%EI_Zebrafish	GA-MLR	$(\Sigma\varepsilon/N)^2$, pEC_{50}HaCaT	10	0.91	0.88	0.84	0.76	0.10	4	0.79	0.68	0.51	0.22

9

Table 3 Observed and predicted response values for individual single line models.

ID	Metal oxide	E. coli (pEC$_{50}$Ecoli)		HaCaT Cell line (pEC$_{50}$HaCaT)		Zebrafish (%EI_Zebrafish)	
		Observed	Predicted	Observed	Predicted	Observed	Predicted
1	AL$_2$O$_3$	2.49	2.73	1.85	1.80	3.44	3.01
2	CeO$_2$$^\$$	-	-	-	-	10.8	8.83
3	Co$_3$O$_4$$^\$$	3.00	2.98	-	-	-1.04	14.34
4	CoO*	3.51	3.46	2.83	7.82	4.00	X
5	Cr$_2$O$_3$$^\$$	2.51	2.70	2.3	13.93	44.72	24.55
6	CuO$^\$$	3.20	3.20	-	-	50.00	47.23
7	Fe$_2$O$_3$*	2.29	2.73	2.05	6.52	11.04	13.55
8	Fe$_3$O$_4$	-	-	-	-	13.04	12.86
9	Gd$_2$O$_3$	-	-	-	-	11.36	14.45
10	HfO$_2$	-	-	-	-	11.04	12.73
11	In$_2$O$_3$	2.81	2.73	2.92	10.08	7.12	12.43
12	La$_2$O$_3$$^\#$	2.87	2.72	2.87	18.02	13.28	10.92
13	Mn$_2$O$_3$*	3.08	2.72	2.64	X	17.20	9.94
14	NiO*	3.45	3.46	2.49	X	34.56	38.64
15	Ni$_2$O$_3$$^\$$	-	-	-	-	18.32	15.58
16	Sb$_2$O$_3$$^\$$	2.64	2.73	2.31	5.16	9.04	9.98
17	SiO$_2$*	2.20	1.97	2.12	X	10.48	7.53
18	SnO$_2$	2.01	1.97	2.67	6.86	2.56	10.86
19	TiO$_2$	1.74	1.97	1.76	5.68	9.92	8.26
20	WO$_3$	-	-	2.64	22.55	9.04	5.13
21	Y$_2$O$_3$	2.87	2.72	2.21	11.95	9.36	9.15
22	Yb$_2$O$_3$$^\$$	-	-	-	-	11.84	15.47
23	ZnO*	3.45	3.45	3.32	9.97	42.72	35.01
24	ZrO$_2$$^\#$	2.15	1.95	2.02	11.73	7.36	9.04
25	Bi$_2$O$_3$$^\#$	2.82	2.73	2.5	9.51	-	-
26	V$_2$O$_3$*	3.14	2.73	2.24	5.83	-	-

*, #, $ are the compounds used in test set for *E. coli*, HaCaT cell line and Zebrafish respectively.

X: Not used in the modeling

The descriptor $(\sum \varepsilon / N)^2$ is a measure of sum of electronegativity of the atoms of the metal oxide scaled by the number of atoms present. It can be calculated from the electronegativity count of the metal oxide $(\sum \varepsilon)$ which is expressed as following equation:

$$\sum \varepsilon = \varepsilon_{metal} * N_{metal} + \varepsilon_{oxy} * N_{oxy} \qquad (2)$$

Here, ε_{metal} and ε_{oxy} are the electronegativity count of metal and oxygen atoms, respectively and N_{metal} and N_{oxy} are the number of metal and oxygen atoms. The expressions for ε_{metal} and ε_{oxy} can be obtained from **Table 1**. The descriptor shows a negative contribution to the cytotoxic behavior of the nanoparticles. Compounds ZnO and CuO (both having same oxidation number) can be used for demonstration of the contribution of this descriptor where we can clearly identify that with a low $(\sum \varepsilon / N)^2$ value as in ZnO ($(\sum \varepsilon / N)^2 = 1.686$) the toxicity is higher (pEC$_{50}$ = 3.45)

and the opposite is observed in case of CuO having a higher value for $(\sum \varepsilon/N)^2$ (14.301) with a lower toxicity value (pEC$_{50}$ = 3.2).
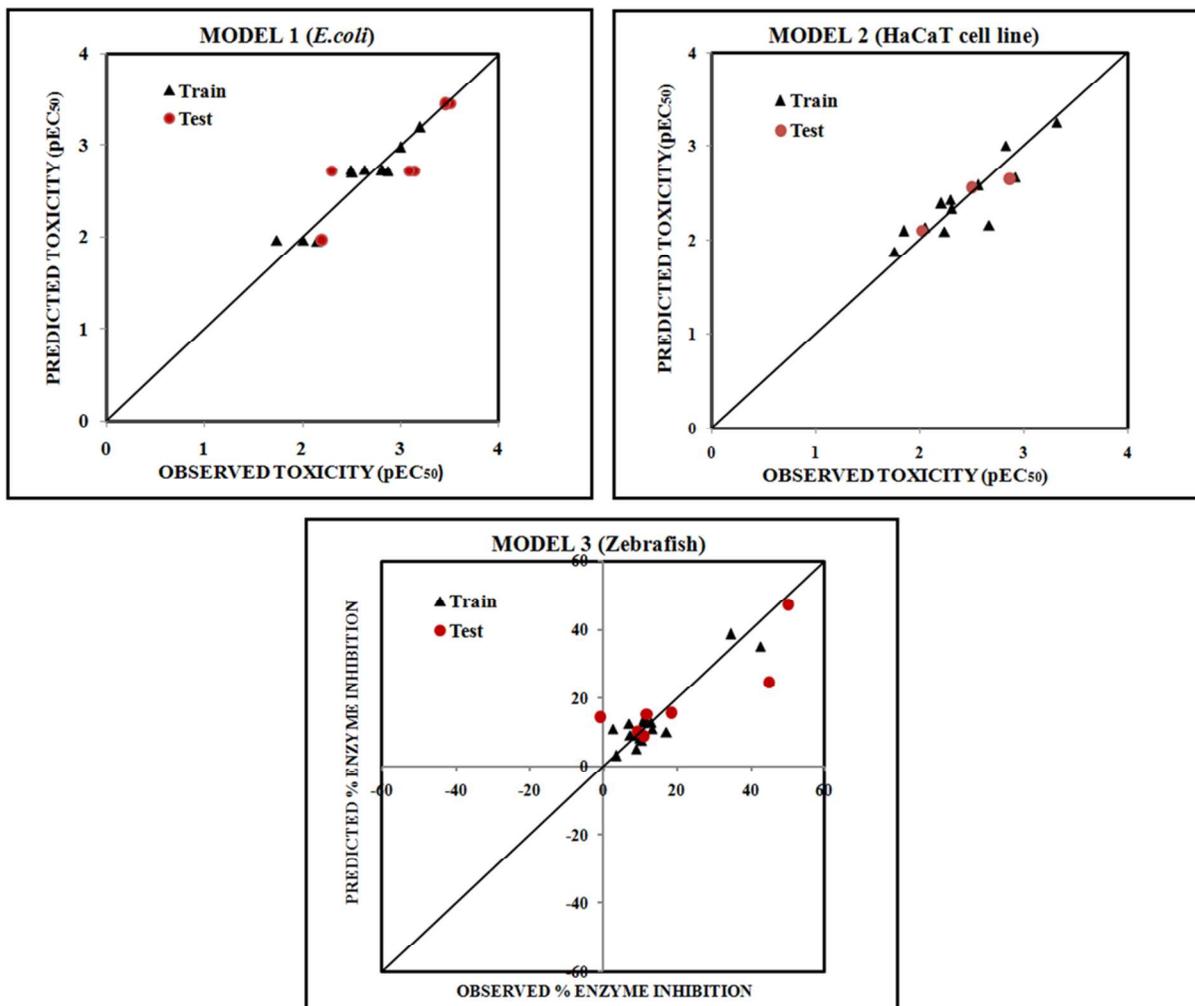

Figure 2 Scatter plot for the best QSTR model of *E. coli* (Top left), HaCaT (Top right) and Zebrafish (Bottom).
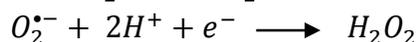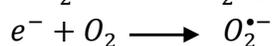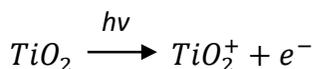
3.1.1 Mechanism for MNPs cytotoxicity to *E. coli*

The MNPs cause the modification and damage of cellular proteins, lipids, and DNA thus leading to disruption of membrane integrity, disturbance in cellular transport chains or induction of an oxidative stress which can subsequently lead to cell death.[47,48] The mechanism of cytotoxicity is thought to involve lipid peroxidation by reactive oxygen species (ROS) such as superoxide ($O_2^{\bullet-}$) and hydroxyl radicals ($\bullet OH$)[49,50] which is initiated by electron detachment from metal oxide nanoparticle and this energy is provided by solar radiation.[49,51] Studies suggest that mechanism of MNP toxicity depends on the release of metal ions from the surface of the nanoparticles.[20,52] Nanoparticles produce oxidative stress by generation of $O_2^{\bullet-}$ and $\bullet OH$ radicals as per following process.[37]

$$O_2 \longrightarrow O_2^{\bullet-}$$

$$O_2^{\bullet-} + H_2O_2 \longrightarrow \bullet OH + O_2$$
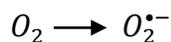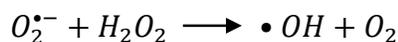
An example has been shown with TiO$_2$ nanoparticle that how in presence of sunlight and O$_2$, electrons are detached and free radicals generated.

$$TiO_2 \xrightarrow{h\nu} TiO_2^+ + e^-$$
$$e^- + O_2 \longrightarrow O_2^{\bullet-}$$
$$O_2^{\bullet-} + 2H^+ + e^- \longrightarrow H_2O_2$$
$$O_2^{\bullet-} + H_2O_2 \longrightarrow \bullet OH + OH^- + O_2$$
$$H^+ + H_2O \longrightarrow \bullet OH + H^+$$

Addition of two enzymes, *viz.*, superoxide dismutase (catalyzing the dismutation of O$_2^-$ into O$_2$ and H$_2$O$_2$) and catalase (catalyzing decomposition of H$_2$O$_2$ to H$_2$O and O$_2$) to Al$_2$O$_3$ suspension significantly reduce the damage caused to *E.coli* membranes[53,54] and confirmed that nanoparticles of TiO$_2$ presented as a coating on cellulose fibers still showed toxicity (but reduced one) in the absence of light and supports the above mechanism. Thus, the latter survey supports and explains the importance of the χ_{ox} descriptor as described by Kar *et al.*[35] Nanoparticles are available in the size range below or equal to 100nm, but NPs of size less than 15nm are toxic towards bacteria/micro-organism/human. This toxicity arises due to the reductive potential, *i.e.*, the detachment of the electron from the metal oxides which can be explained by the descriptor $(\sum \varepsilon / N)^2$. The generated strong reductive fragments instigate the formation of reactive oxygen species (ROS) and the response in the bacteria, the so-called the oxidative stress, which leads to cytotoxicity. The toxicity on MNPs decreases in the following order of the oxidation number of the metal cations:[55] Me^{2+}> Me^{3+}> Me^{4+} which is well explained by the χ_{ox} descriptor. Probable mechanism for cytotoxicity for MNPs against *E. coli* is illustrated in **Figure 3**.
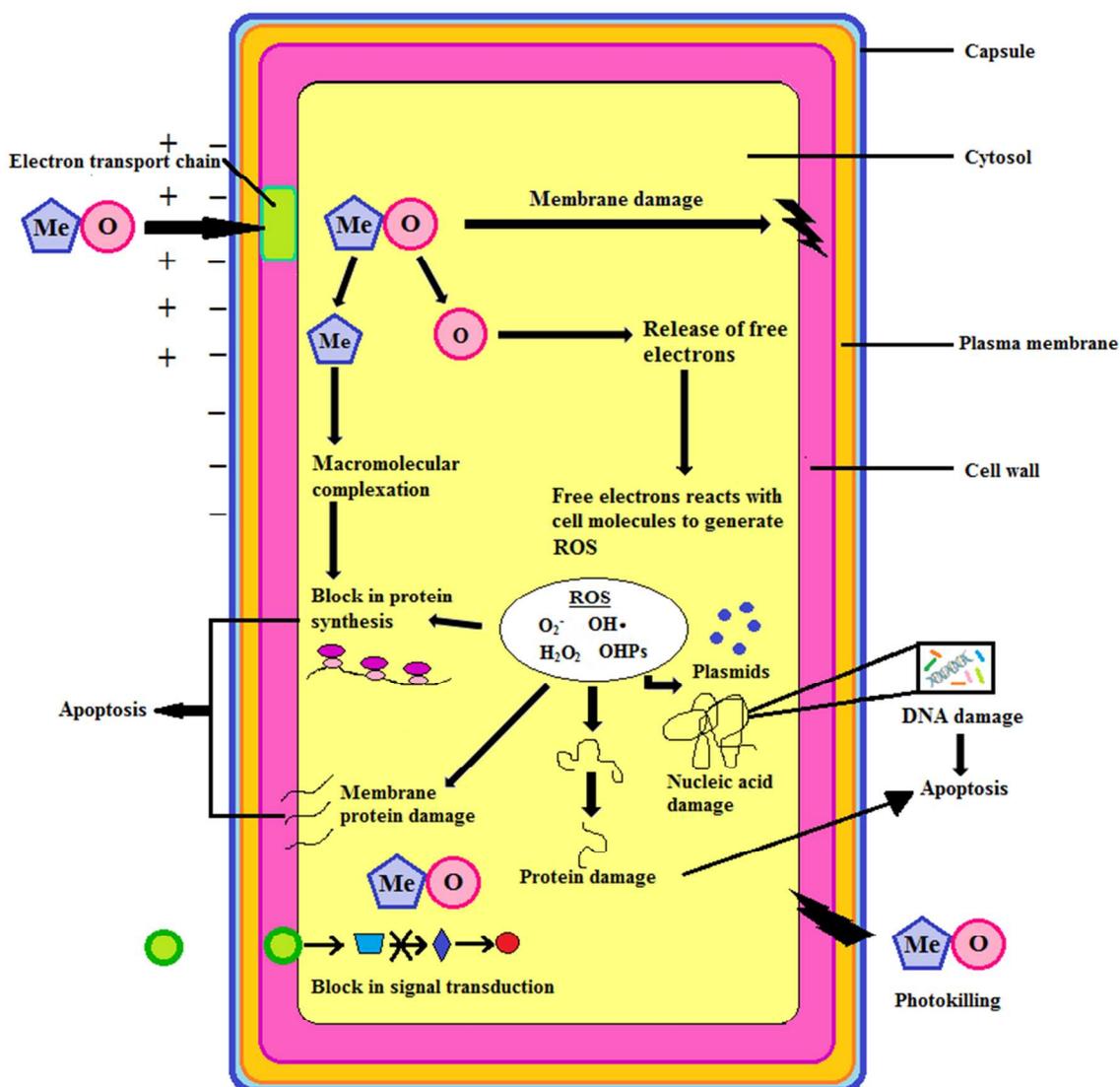
Figure 3 Proposed mechanism of *E. coli* cytotoxicity due to MNPs.

3.2. Modeling of toxicity of MNPs against HaCaT cell line

The PLS equation with one latent variable (LV) evolved as the best model for HaCaT cell line is described as follows:

$$pEC_{50} \ HaCaT = 1.26948 + 1.06163(\sum\chi)/nO + 0.03763\lambda - 0.14632V_{metal}$$

$$N_{train} = 12, R^2 = 0.78, Q^2_{(LOO)} = 0.61, \overline{r^2_{m(loo)}} = 0.509, \Delta r^2_{m(loo)} = 0.160$$

$$N_{test} = 3, Q^2_{F1} = 0.83, Q^2_{F2} = 0.83, \overline{r^2_{m(test)}} = 0.67, \Delta r^2_{m(test)} = 0.12 \tag{3}$$

The model showed acceptable values of the determination coefficient (R^2=0.78) and cross-validated correlation coefficient (Q^2_{LOO}=0.61), signifying the statistical reliability of the model. The predictivity of the model was judged by means of predictive R^2 (R_{pred}^2) or Q_{F1}^2 (0.83) and Q_{F2}^2 (0.83) which show good predictive ability of the model. We have also calculated $\overline{r^2_{m(loo)}}$ and

$\Delta r^2_{m(loo)}$ employing the observed and LOO-predicted values of the developed model. The statistical quality of the developed model is illustrated in **Table 2**. The experimental and predicted toxicity values of MNPs against HaCaT obtained from developed equation 3 are presented in **Table 3**. The scatter plot (**Figure 2,** top right) showed that the points were limitedly scattered around the line of fit for both training and test sets accounting for the robustness of the model.

The regression coefficient plot[40] (**Figure 4,** top left) provides the knowledge about the positive or negative contribution of descriptors towards the toxicity of the MNPs. Descriptors like $(\sum\chi)/nO$ and λ with positivecoefficients imply that as the descriptor values increase, the toxicity of MNPs also increases whereas the descriptor with a negative coefficient (V_{metal}) decreases the toxicity of MNPs with their increasing numerical values. From the variable importance plot (VIP) (**Figure 4,** top right), the order of contribution of each descriptor is obtained. The most and the least important descriptors contributing to the toxicity can be identified with the help of this plot. A variable with a VIP score >1 shows higher statistical significance as compared with one with a low VIP value.[56] The descriptors are arranged in the plot according to their importance (maximum contribution to minimum contribution) and their significance level is found to be in the following order: $(\sum\chi)/nO$, λ and V_{metal}.
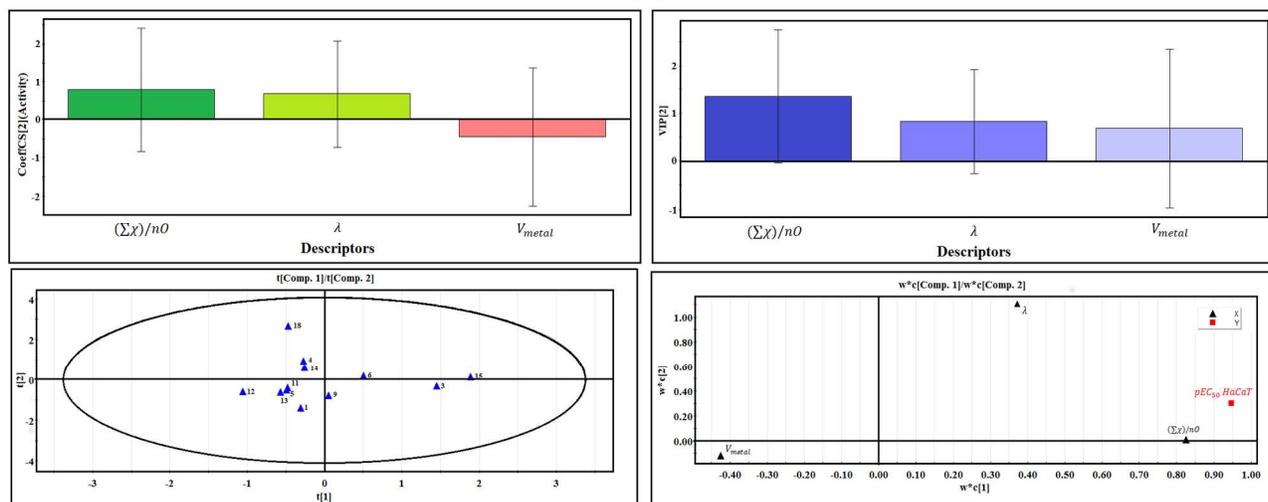

Figure 4 Regression coefficient plot (Top right), variable importance plot (Top left), score plot (Bottom left) and loading plot (Bottom right) of the best QSTR model for HaCaT cell line

The descriptor contributing most to the response is $(\sum\chi)/nO$, which is the totalmetal electronegativity in a specific metal oxide relative to the number of oxygens. The descriptor shows a positive coefficient thus indicating that with increase in its value, the toxicity of the metal oxide increases. Highly toxic compounds like ZnO (pEC_{50} = 3.32) and CoO (pEC_{50} = 2.83) have higher values of $(\sum\chi)/nO$ (1.65 and 1.88 respectively) while in case of low toxic compound like TiO_2 (pEC_{50} =1.76), the value of the descriptor is also less (0.77).

The next most important descriptor is λ which signify the core environment of metal defined by the ratio of the number of core electrons to the number of valence electrons and can be expressed as follows:

$$\lambda = \frac{Z_{metal} - Z_{metal}^v}{Z_{metal}^v} \qquad (4)$$

Here, Z_{metal} is the atomic number of the metal atom and Z_{metal}^v is the valence electron of the metal. The positive regression coefficient denotes that an increased value for λ will lead to an increased toxicity of metal oxide as seen in WO_3 (pEC_{50} = 2.56; λ = 36) and *vice versa* in case of Al_2O_3 (pEC_{50} = 1.85; λ = 3.33). The least important descriptor for this model is V_{metal} which is the valency of the specific metal *i.e.*, the number of electrons involved or available for chemical bond formation. The descriptor has a negative contribution towards metal oxide toxicity which is evident in MNPs like Cr_2O_3 and ZnO. In case of Cr_2O_3, a high valency (V_{metal} = 6) results in lower toxicity (pEC_{50} = 2.3) whereas in case of ZnO, V_{metal} is 2 and toxicity is high (pEC_{50} = 3.32).

The distribution of the compounds in the latent variable space as defined by the scores is expressed in a score plot, as given in **Figure 4 (**Bottom left**)**. Here, we have plotted the scores of the first two components t^1 and t^2. The applicability domain of the model is indicated by the ellipse, as defined by Hotelling's t^2. Hotelling's t^2 is a multivariate generalization of Student's t-test. It provides a check for compounds adhering to multivariate normality.[57] In this plot, compounds which are situated near each other have similar characteristics or properties, whereas compounds which are far from each other have dissimilar properties with respect to their toxic properties towards HaCaT cell line. For example, compounds which are located in the upper left-hand corner like **4** (Cr_2O_3) and **14** (Y_2O_3) have some similarity in their properties whereas compounds which are far from each other like **15** (ZnO) and **12** (TiO2) represent heterogeneity in the property space. The compounds which are close to the center of the plane have average properties. Since there are no compounds present outside the ellipse, we can conclude that there are no outliers according to this method.

The relationship between the X-variables and Y-variables can be understood by the loading plot (**Figure 4,** bottom right) where three X-variables and one Y-variable are shown. The loading plot was developed using the first two components. The loading plot helps in understanding the impact of different variables to the model. For interpretation, we should consider the distance of the descriptors from the plot origin. Similar type of descriptors contributing similar meaning will be grouped together, whereas the descriptors with different meaning will be far from each other. If a descriptor is far from the plot origin, it is considered to give greater impact on the response value. For example, the X-variable $\Sigma\chi/nO$ influences the Y-variable most because of its closeness to Y-variable, *i.e.*, with an increase in this descriptor's value, the $pEC_{50}HaCaT$ will increase. The descriptor V_{metal} which is present in the opposite side of the plot origin with respect to

pEC$_{50}$HaCaT, suggests that an increase in V$_{metal}$ value will result in a decrease in the cytotoxicity value for HaCaT cells.

3.2.1 Mechanism behind MNPs cytotoxicity to HaCaT

The generation of intracellular ROS levels could induce oxidative damages to cellular components, finally leading to apoptosis.[58] Intracellular ROS causes lipid peroxidation due to metal oxide induced stress generation as discussed by Lee *et al*.[59] A number of ROS have been found to be responsible for toxicity of MNPs, such as OH radicals[60], superoxide ions[61], hydrogen peroxide in cooperation with other ROS[62] *etc*. The descriptor $(\sum \chi)/nO$ is a measure of electronegativity and according to Portier *et al*.[63] the value of electronegativity of a given metal oxide (χ) is strongly related to the electronegativity of the corresponding cation (χ^+). The cation electronegativity depends on the ionic radius and formal charge of the cation, *i.e.*, higher values of χ^+ will characterize those cations that have a relatively large charge distributed along a relatively small atomic radius. Since electronegativity is an assessment 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[64] 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 MNP.

For example, metal cations such as Cu^{2+} can detach from the surface of MNPs and may catalyze the formation of hydroxyl radicals ($\bullet OH$) via the so-called Haber-Weiss-Fenton cycle through following steps[65]

$$O_2^{\bullet-} + Cu^{2+} \longrightarrow O_2 + Cu^+$$
$$Cu^+ + H_2O_2 \longrightarrow Cu^{2+} + OH^- + \bullet OH$$
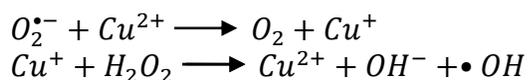
The ROS can be produced at any time as byproducts during cellular respiration in all aerobic organisms because they use molecular oxygen to obtain energy.[66] Problem arises when the cell is unable to maintain a balance between the levels of oxidized and reduced species through various antioxidants and enzymes that scavenge the free radicals resulting increased ROS production followed by toxicity.

Another mechanism through which metal oxides work is related to the ability of transferring electrons between the surface of MeOx and intracellular redox couples. This occurs due to the detachment of electrons from the valence band to the conduction band and takes place due to the impact of intracellular redox processes occurring in the biological media.[37] All these studies further highlight the aspects such as morphology of the cell, size, and shape of the NP, and the NPs' solubility, which mainly control the mechanism predominantly.[67] The complete proposed mechanism for the toxicity of MNPs to HaCaT cell line is reported in **Figure 5**.
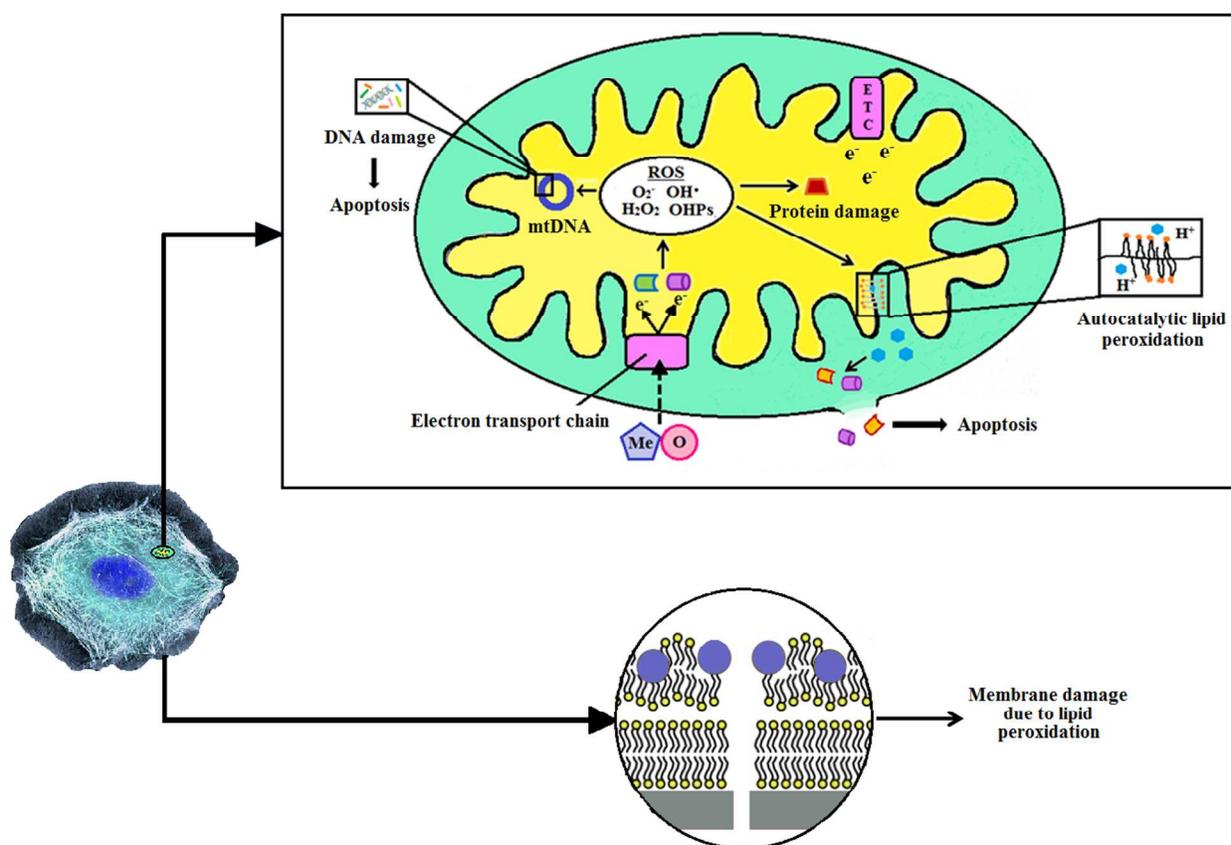
Figure 5 Proposed mechanism of HaCaT cytotoxicity due to MNPs.

3.3 Modeling of toxicity of MNPs against Zebrafish

The equation for the GA-MLR based model for inhibition of ZHE1 hatching enzyme activity is mentioned as following:

$$\%EI\_Zebrafish$$
$$= 10.286(\pm6.594) - 10.166(\pm2.217)\,N_{oxy} + 16.198(\pm3.712)\,\sum\chi/nO$$
$$+ 1.750(\pm0.512)\sum\alpha_{metal}$$
$$N_{train} = 16, R^2 = 0.83, R^2_{adj} = 0.79, Q^2_{(loo)} = 0.68, \overline{r^2_{m(loo)}} = 0.61, \Delta r^2_{m(loo)} = 0.06$$
$$N_{test} = 7, Q^2_{F1} = 0.74, Q^2_{F2} = 0.70, \overline{r^2_{m(test)}} = 0.54, \Delta r^2_{m(test)} = 0.22 \tag{5}$$

All regression coefficients in eq. (5) are significant at 95% confidence level. From equation 5, we can infer that the model could explain 83% of the variance while it could predict 68% of the variance (leave-one-out predicted variance). The metrics Q^2_{F1} or R^2_{pred} and Q^2_{F2} are used for external validation purposes, and their values are much higher than the stipulated threshold value. The statistical quality of the developed model is illustrated in **Table 2**. The experimental and predicted toxicity values of MNPs against HaCaT cell line obtained from developed

equation 5 are presented in **Table 3**. The scatter plot (**Figure 2**, bottom) showed that the points were limitedly scattered around the line of fit for both training and test sets.

The descriptor N_{oxy} is the number of oxygen atoms in the metal oxide, which shows a negative correlation towards the enzyme inhibition of ZHE1 hatching enzyme. Thus, with an increase in the number of oxygen the enzyme inhibition activity of the metal oxide would decrease as seen in Al_2O_3 (N_{oxy}= 3; $\%EI\_zebrafish$ = 3.33) and Fe_3O_4 (N_{oxy}= 4; $\%EI\_zebrafish$ = 13.04). The opposite occurs in metal oxides having less number of oxygen, *i.e.*, the enzyme inhibition activity increases as seen in NiO and ZnO both having one oxygen each and percentage enzyme inhibition is 34.56 and 42.72, respectively.

$\sum\chi/nO$ as described earlier is a measure of electronegativity and in case of zebrafish also, it shows a positive correlation to the enzyme inhibition activity. For MNPs like NiO and ZnO, where $\sum\chi/nO$ value is higher (1.91 and 1.65, respectively), the enzyme inhibition is also high (34.56 and 42.72 respectively). For compounds like ZrO_2, the $\sum\chi/nO$ value is less ($\sum\chi/nO$ = 0.665) and hence percentage enzyme inhibition is also less ($\%EI\_zebrafish$ = 7.36).

The $\sum\alpha_{metal}$ descriptor is the total core count of the metal which can be expressed through following equation:

$$\sum\alpha_{metal} = \alpha_{metal} * N_{metal} \qquad (6)$$

The detailed description for α_{metal} and N_{metal} is given in **Table 1**. The descriptor has a positive influence on the enzyme inhibition of metal oxides which means with an increase in the core count of metal, the inhibitory action will also increase and *vice versa*. This is evident from the compound Cr_2O_3 which has a high $\sum\alpha_{metal}$ value ($\sum\alpha_{metal}$= 15.33) and a high enzyme inhibition ability ($\%EI\_Zebrafish$= 44.72). On the other hand, CoO ($\sum\alpha_{metal}$ = 4.17) and Al_2O_3($\sum\alpha_{metal}$ = 3.33) having low descriptor values have low inhibition ability also (4.00 and 3.44 respectively).

3.3.1 Mechanism towards ZHE1 enzyme inhibition

The modelled data showed that while ZnO, CuO, Cr_2O_3 and NiO lead to ZHE1 enzyme inactivation and cause interference in hatching, rest of the nanoparticles did not have similar effects. This might be due to low solubility or the inability of the shed metal ions to interact with the metal-binding histidines in the enzyme center. There is a prominent correlation between the ability of CuO, ZnO, Cr_2O_3 and NiO to inhibit ZHE1 activity and exert hatching interference at the organism level. It was understood from ZHE1 activity that dissolution characteristics as well as chemistry of the shed metal cations influences histidine ligation at the enzyme center and could lead to toxicological consequences.[38] However, dissolution property is not only the determining factor for toxicity in case of moderate (CoO, Cr_2O_3, Fe_3O_4 and Sb_2O_3) or high (CuO, NiO, WO_3 and ZnO) dissolution potential MNPs, thus inferring that the chemistry of metal ions has an important role.

It has been already proved that nanoparticles have different affinities for proteins based on their charge, size and surface coating.[68] Zebrafish hatching enzyme (ZHE1) is a zinc metalloprotease enzyme in which the active Zn^{2+} binding site includes three histidine residues.[69,70] Many divalent metal ions can replace the zinc ion at the enzyme binding site. The binding affinity is not same for all metals and as demonstrated by Holt *et al.*[71] the binding affinity of metals to metallothionein is in the order of $Zn^{2+} < Cd^{2+} < Cu^{2+} < Hg^{2+}$. This ligand configuration also has the ability to bind to Cu^{2+}, Ni^{2+}, and Cr^{3+} thus inferring that the shed metal ions from CuO, NiO and Cr_2O_3 inhibit the hatching enzyme (ZHE1) by substituting Zn^{2+} ion in the active enzyme center. The enzymatic activity was increased when Zn^{2+} ion was replaced with Cu^{2+} or Co^{2+} in the catalytic center of astacin and inclusion of Ni^{2+} or Hg^{2+} inactivated the enzyme.[72] It was found that thermolysin activity increased when cobalt was present.[73] Interactions between the hatching enzyme and the nanoparticles could cause deformation and potential inactivation of the protein.[74,75] Any structural changes in the enzyme caused by interaction with MNPs can lead to inactivation or decreased selectivity for substrates thus leading to potential inactivation of the enzyme. Another probable mechanism might be photocatalytic degradation of ZHE1 enzyme due to ROS production which occurred in presence of nano-ZnO and nano-TiO_2.[76,77] In case of Fe_2O_3, it aggregates and adhere to the surface of embryo indirectly causing interference with the digestive function of the hatching enzyme.[78,79] The complete mechanism for the enzyme inhibition of MNPs to Zebrafish is reported in **Figure 6**.
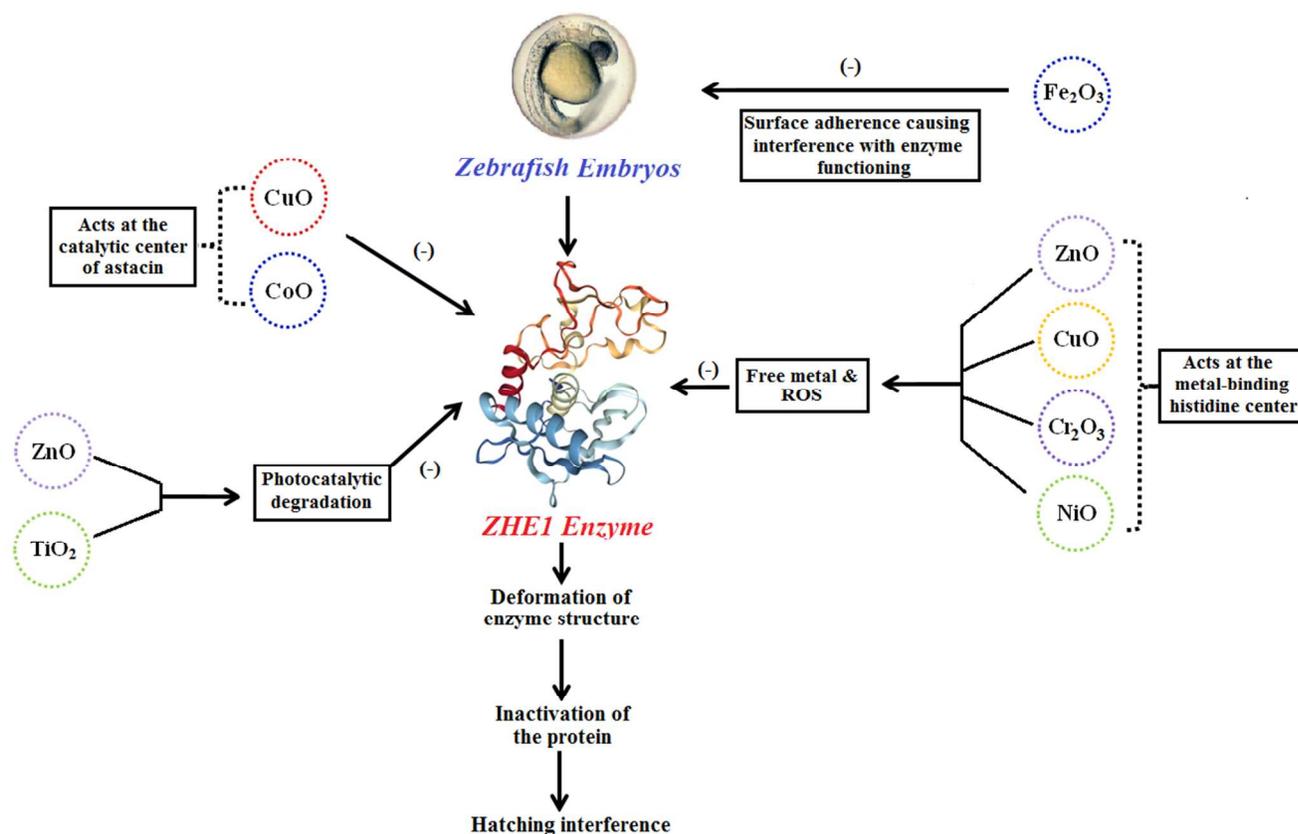
Figure 6 Proposed mechanism of Zebrafish hatching enzyme inhibition due to MNPs.

3.4. Interspecies quantitative structure toxicity-toxicity relationships (i-QSTTR)

Interspecies quantitative correlation is a mathematical relationship between two different biological endpoints measured different species. The i-QSTTR study deals with extrapolating response data for one species to another species for an explicit endpoint when the experimental data for the second species are not present for some data points. This advanced method can overcome the extra cost of manifold toxicity tests along with the competent understanding of the mechanism of toxic action (MOA) of MNPs for different species and endpoints.[29]

3.4.1. *E. coli* – HaCaT cell line interspecies toxicity correlation

On the basis of a good correlation coefficient found between the available experimental values for *E. coli* and HaCaT cell line ($r = 0.655$), we have decided to develop a quantitative linear relationship between the two biological endpoints employing i-QSSTR approach. The toxicity endpoint of *E. coli* is taken as the response variable and the HaCaT response as an independent variable along with other computed descriptors for the study. The interspecies PLS model with 2 descriptors and 1 latent variable (LV) for *E. coli* and HaCaT is given as follows:

$$pEC_{50}E.coli = 2.886 - 0.449\chi ox + 0.510 pEC_{50}HaCaT$$

$$N_{train} = 12, R^2 = 0.78, Q^2 = 0.67, PRESS = 1.06, \overline{r^2_{m(loo)}} = 0.58, \Delta r^2_{m(loo)} = 0.11,$$

$$N_{test} = 5, Q^2_{F1} = 0.91, Q^2_{F2} = 0.82, \overline{r^2_{m(test)}} = 0.70, \Delta r^2_{m(test)} = 0.10 \qquad (7)$$

Equation 7 could explain 78% of the variance and predict 67% of the variance of the response. The external predicted variance for Eq. (7) is 91% which is fairly high. The calculated (training set compounds) and predicted (test set compounds) toxicity values obtained from Eq. (7) are given in the **Supplementary Information** excel file. The statistical quality of the interspecies model is reported in **Table 2**. The closeness of experimental and predicted values for MNPs has been also ascertained from scatter plot shown in **Figure 7** (Top left).
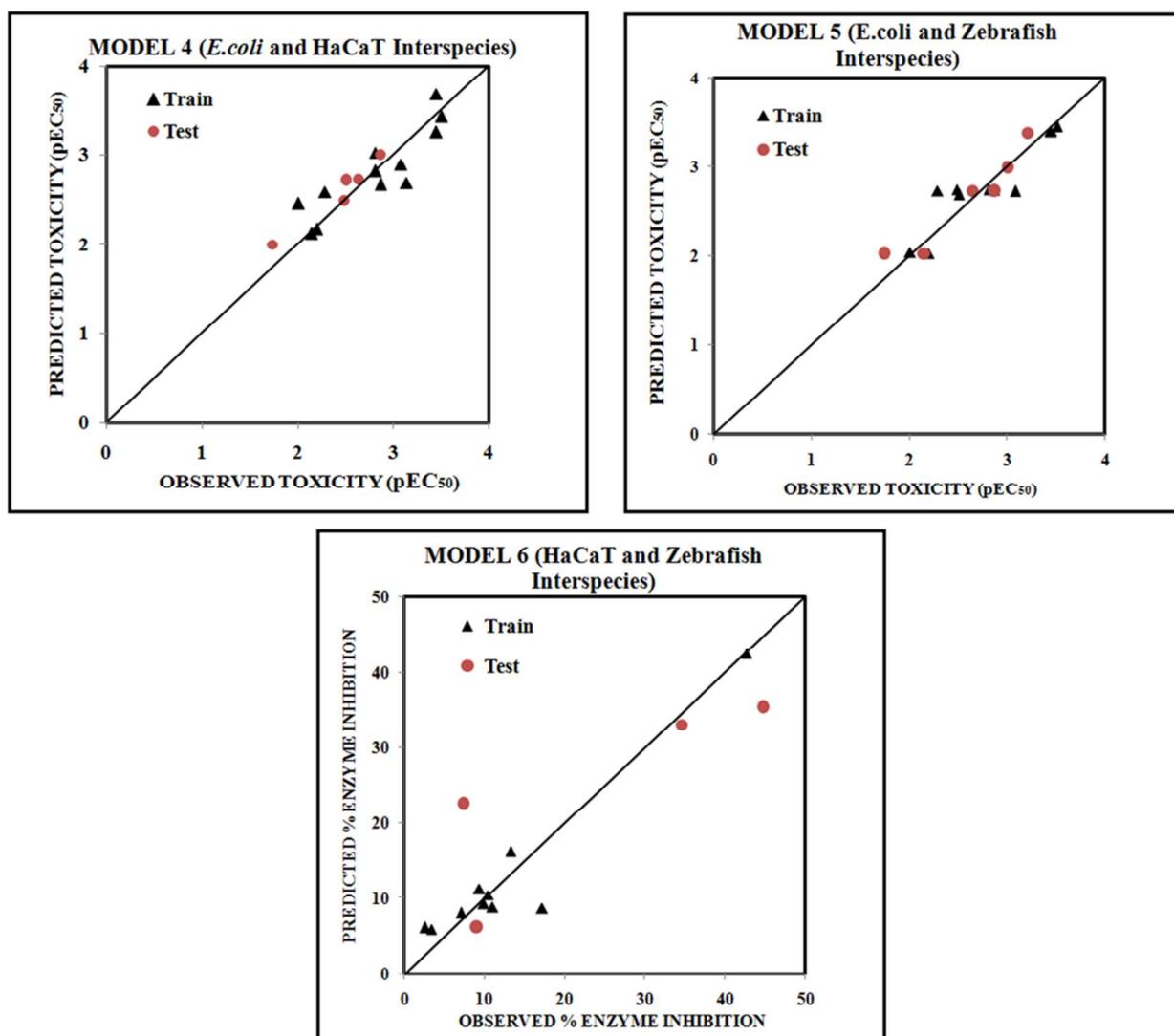

Figure 7 Scatter plot for the best i-QSTTR model of *E. coli*-HaCaT toxicity (Top left), *E. coli*-Zebrafish toxicity (Top right) and HaCaT-Zebrafish toxicity (Bottom).

The coefficient and VIP plots of the modeled descriptors are presented as histograms in **Figure 8** (Top and bottom, respectively). According to the descending VIP values, χ_{ox} has more contribution towards pEC$_{50}$*E.coli* than pEC$_{50}$HaCaT. According to the coefficient plot, χ_{ox} has a negative contribution towards *E. coli* toxicity of MNPs whereas the descriptor pEC$_{50}$HaCaT gives a positive contribution towards MNP toxicity to *E. coli*.
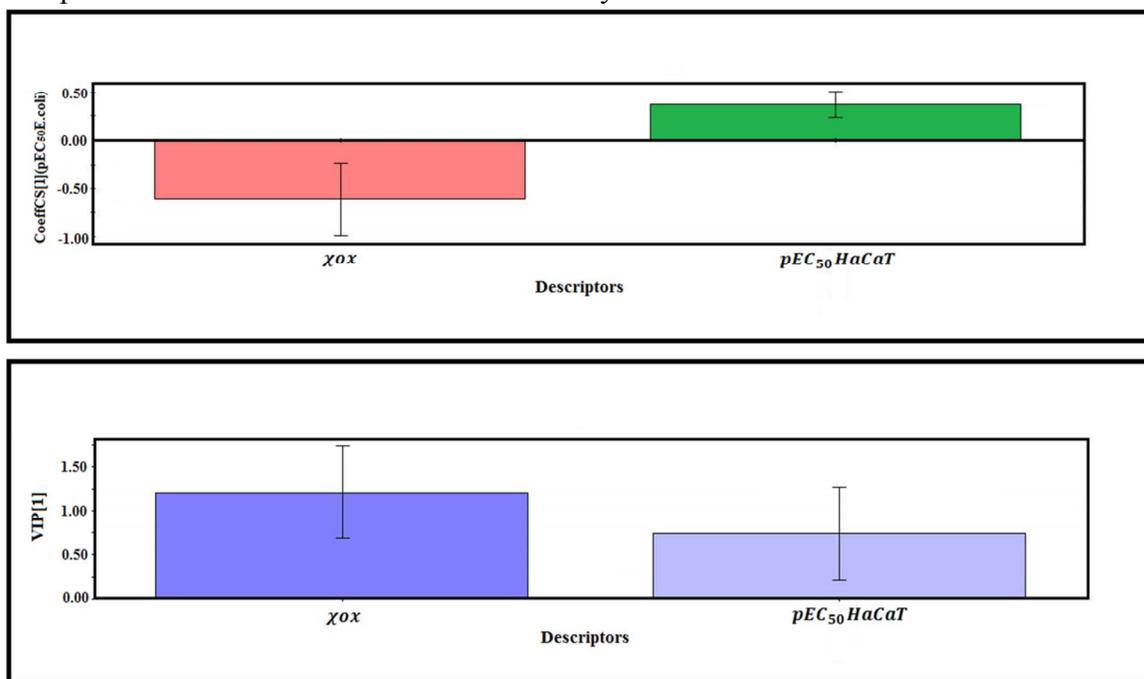

Figure 8 Regression coefficient plot (Top) and variable importance plot (Bottom) for *E. coli* – HaCaT cell line toxicity interspecies QSTTR model.

It's already discussed in the single line model for *E. coli* that χ_{ox} or the oxidation number of the metal atoms evolved as the most important descriptor with a negative correlation coefficient. The mechanism of toxic action of the MNPs related to this descriptor is already discussed previously. HaCaT toxicity (pEC$_{50}$HaCaT) evolved as a less important descriptor for *E. coli* toxicity as observed from equation 7 and the descriptor gives a positive correlation towards *E. coli* toxicity. MNPs having high values for HaCaT toxicity like CoO (2.83) and ZnO (3.32) have also high E. coli toxicity values (pEC$_{50}$E.coli), *i.e.*, 3.51 and 3.45, respectively. On the other hand, MNPs with low HaCaT toxicity value like Fe$_2$O$_3$ (2.05) and SiO$_2$ (2.12) have lower toxicity values (2.29 and 2.2, respectively) for *E. coli*.

3.4.2. *E. coli* - Zebrafish interspecies toxicity correlation

Interspecies relationship between *E. coli* toxicity and ZHE1 enzyme inhibition has been found out to understand the correlation between the two endpoints. Considering *E. coli* cytotoxicity as the response variable and Zebrafish hatching enzyme inhibition along with calculated/selected descriptors as predictor variables, we have constructed the following best equation employing GA-MLR statistical approach.

$$pEC_{50}E.coli = 4.875(\pm0.420) - 0.709(\pm0.125)\chi ox - 0.001(\pm0.005)\ \%EI\_zebrafish$$

$$N_{train} = 11, R^2 = 0.83, R^2_{adj} = 0.79, Q^2 = 0.74, PRESS = 0.48, \overline{r^2_{m(loo)}} = 0.65,$$
$$\Delta r^2_{m(loo)} = 0.12$$

$$N_{test} = 6, Q^2_{F1} = 0.91, Q^2_{F2} = 0.90, \overline{r^2_{m(test)}} = 0.84, \Delta r^2_{m(test)} = 0.08 \qquad\qquad (8)$$

Equation 8 could explain of the 83% variance and predict 74% of the variance of the response. External predicted variance for Eq. (8) is 91% which is fairly high. The calculated (training set compounds) and predicted (test set compounds) toxicity values obtained from Eq. (8) have been provided in the **Supplementary Information** excel file. The statistical quality of the interspecies model is reported in **Table 2**. The closeness of experimental and predicted values for MNPs has been also ascertained from scatter plot shown in **Figure 7** (Top right).

The developed equation with two independent variables consists of χ_{ox} and %EI_Zebrafish response as descriptors. The occurrence of χ_{ox} in all the developed single species and interspecies models for *E. coli* cytotoxicity proved that it is one of the important features for the cytotoxicity mechanism of *E. coli*. The descriptor %EI_Zebrafish imparts a negative contribution towards *E. coli* cytotoxicity. This probably indicates about the difference in the mechanism of toxicity for these two endpoints. For MNPs like CoO, %EI_Zebrafish value is low (4.0) and thus, according to the descriptor contribution, the response value (pEC$_{50}$E.coli = 3.51) is high, whereas when the %EI_Zebrafish is high (as in case of Cr$_2$O$_3$), the pEC$_{50}$E.coli is low (2.51).

3.4.3. HaCaT cell line - Zebrafish interspecies toxicity correlation
The interspecies correlation between HaCaT cell line cytotoxicity and Zebrafish hatching enzyme inhibition has been found out through i-QSSTR model and the resultant best equation derived using GA-MLR method is given below:

$$\%EI\_Zebrafish$$
$$= 4.276(\pm7.438) + 21.338(\pm3.376)(\textstyle\sum\varepsilon/N)^2 + 0.657(\pm3.256)pEC_{50}HaCaT$$

$$N_{train} = 10, R^2 = 0.91, R^2_{adj} = 0.88, Q^2 = 0.83, PRESS = 110.431, \overline{r^2_{m(loo)}} = 0.76,$$
$$\Delta r^2_{m(loo)} = 0.11$$

$$N_{test} = 4, Q^2_{F1} = 0.79, Q^2_{F2} = 0.68, \overline{r^2_{m(test)}} = 0.51, \Delta r^2_{m(test)} = 0.22 \qquad\qquad (9)$$

Equation 9 could explain of the 91% of the variance and predict 83% of the variance of the response. External predicted variance for Eq. (9) is 79% which is more than the acceptable threshold value. The calculated (training set compounds) and predicted (test set compounds) toxicity values obtained from Eq. (9) are reported in the **Supplementary Information** excel file. The statistical quality of the interspecies model is reported in **Table 2**. The closeness of experimental and predicted values for MNPs has been also ascertained from scatter plot shown in **Figure 7** (Bottom).

The descriptor $(\sum\varepsilon/N)^2$ is a measure of electronegativity of the metal oxide as already discussed earlier. It gives a positive coefficient for the response (percentage ZHE1 hatching enzyme inhibition). MNPs like ZnO (%EI_Zebrafish = 42.72) and Cr$_2$O$_3$ (%EI_Zebrafish = 44.72)

23

having high values for $(\sum \varepsilon/N)^2$(1.69 and 1.39 respectively) property also have high percentage enzyme inhibition values. On the contrary, MNPs like SnO$_2$ (%EI_Zebrafish = 2.56) and Al$_2$O$_3$ (%EI_Zebrafish = 3.44) having lower $(\sum \varepsilon/N)^2$(0.004 and 0.02 respectively) values also showed low percentage ZHE1 enzyme inhibition values.

HaCaT cytotoxicity ($pEC_{50}HaCaT$) also showed a positive correlation with percentage ZHE1 enzyme inhibition. For ZnO, the $pEC_{50}HaCaT$ value is high (3.32) and the percentage ZHE1 enzyme inhibition is also high (42.72) whereas for Al$_2$O$_3$, $pEC_{50}HaCaT$ is low (1.85) and the percentage ZHE1 hatching enzyme inhibition is also low (3.44). Thus, metal oxides with high cytotoxicity features for HaCaT cell lines tend to show high percentage ZHE1 hatching enzyme inhibition

3.5 AD study for Single species QSTR and interspecies-QSTTR models
The AD study was performed using the standardization approach and leverage method to detect the reliability of predictions of the models within a defined chemical space. Based on consensus of both approaches, it was found that there were no compounds staying outside the AD for training and test sets for HaCaT single endpoint model as well as all three interspecies models. In case of *E. coli* single cell line model, one compound (CuO) in the training set was an influential observation, but no compound in the test set was outside the AD. For the Zebrafish model, both approaches suggested that no compounds are outliers from the training and test set. However, according to the Williams plot, Cr$_2$O$_3$ is a response outlier (Y outlier) and it is wrongly predicted because the standardized residual value of this chemical is more than 3σ. All Williams plots are placed in the Supplementary Information excel file as **Figures S3** and **S4** for single endpoint and interspecies models, respectively.

3.6 Comparison with previously published models
A critical comparison of the statistical quality with previously reported models is not always possible due to differences in the composition of training and test sets; however, we have tried to discuss here advantages of our models over the previously reported models. In case of the *E.coli* single line model, we have used a higher number of compounds while modeling than that of the previously developed models.[35,80] The descriptor χ$_{ox}$ appeared in our model also corroborates the previous findings by Kar *et al*.[35] Apart from this, the use of simple periodic table descriptors has also provided with comparable results with that of the previously developed model[80] where quantum chemical descriptors were used. In case of HaCaT model, the results were comparable to that of previously developed model.[37] For zebrafish enzyme inhibition modeling, the previous authors[81] have employed enzyme activity as the endpoint which appears not be justified as an endpoint to reflect the toxicity of MNPs properly. We have utilized the percentage enzyme inhibition parameter as the endpoint which provides a better understanding of metal oxide toxicity. The linear model obtained in the present study for zebrafish is however comparable with that of the non-linear model as described in the previous paper.[81]

3.7 External dataset prediction
A dataset of 42 MNPs was used as the true external dataset for prediction using the developed single line models. The computed descriptors for all external set MNPs are illustrated in Supplementary Information excel file. The predicted values for the external dataset compounds are given in **Table 4**. The reliability of the prediction values employing the three single species

models was determined by 'Prediction Reliability Indicator' tool [http://dtclab.webs.com/software-tools] which indicates or categorizes the quality of predictions of the external set with confidence into three groups: *good* (composite score 3), *moderate* (composite score 2) and *bad* (with composite score 1). Also, the tool is designed to predict the AD of the external dataset through the standardization approach (details about AD can be found in **Table 4**). Therefore, we have checked the prediction of each MNP through two layered confidence checking (i.e., Prediction Reliability Score and AD criteria). MNPs like Ag_2O, Au_2O, Au_2O_3, PtO, PtO_2, Tl_2O and WO_3 having "bad quality" predictions should be relooked into for further studies. The prediction quality can be checked with color coding representation in **Figure 9**.

Figure 9 Prediction quality of external test set. **Green**: Composite Reliability Score 3 (Good prediction), **Yellow**: Composite Reliability Score 2 (Moderate prediction), **Red**: Composite Reliability Score 1 (Bad prediction), **Gray**: MNPs experimental data already present for the specific species.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Table 4 Predicted values for external dataset compounds for the individual single line models.

ID	Metal oxide	E. coli			HaCaT			Zebrafish		
		Predicted value in pEC_{50}(MLR model)	Prediction Reliability Score	AD	Predicted value in pEC_{50} (PLS model)	Prediction Reliability Score	AD	Predicted value in %EI_Zebrafish (MLR model)	Prediction Reliability Score	AD
1	Ag_2O	4.03	2	Out	6.51	1	Out	102.90	1	Out
2	Au_2O	3.82	1	Out	8.87	1	Out	137.01	1	Out
3	Au_2O_3	2.57	3	In	5.27	1	Out	61.82	1	Out
4	BaO	3.43	3	In	2.94	3	In	23.99	3	In
5	BeO	3.48	3	In	2.68	3	In	27.30	3	In
6	Bi_2O_3	EA	-	-	EA	-	-	12.52	3	In
7	CaO	3.48	3	In	2.38	3	In	21.57	3	In
8	CdO	3.45	3	In	3.64	3	In	37.56	3	In
9	CeO_2	1.94	3	In	2.33	3	In	EA	-	-
10	Co_2O_3	2.72	3	In	2.19	3	In	14.67	3	In
11	Co_3O_4	EA	-	-	2.51	3	In	EA	-	-
12	CuO	EA	-	-	3.75	3	In	EA	-	-
13	Fe_3O_4	2.98	3	In	2.30	3	In	EA	-	-
14	Ga_2O_3	2.75	3	In	2.46	3	In	10.22	3	In
15	Gd_2O_3	2.71	3	In	2.85	3	In	EA	-	-
16	GeO_2	1.97	3	In	2.01	3	In	10.32	3	In
17	HfO_2	1.92	3	In	2.69	3	In	12.73	3	In
18	HgO	3.32	3	In	4.27	3	In	46.17	3	In
19	IrO_2	1.90	3	In	2.68	3	In	20.90	3	In
20	MgO	3.48	3	In	2.56	3	In	25.71	3	In
21	MnO_2	1.96	3	In	1.50	3	In	9.22	3	In
22	Mo_2O_3	2.67	3	In	3.46	1	Out	38.99	2	Out
23	Nb_2O_3	2.67	3	In	3.18	1	Out	32.07	2	Out
24	Ni_2O_3	2.72	3	In	2.53	3	In	EA	-	-

26

25	OsO_2	1.91	3	In	2.66	3	In	20.72	3	In
26	PbO	3.48	3	In	3.89	1	Out	44.69	2	Out
27	PbO_2	1.97	3	In	2.65	3	In	15.65	3	In
28	PdO	3.27	3	In	3.08	1	Out	36.44	2	Out
29	PtO	2.55	1	Out	5.71	1	Out	64.01	2	Out
30	PtO_2	1.56	1	Out	4.50	1	Out	35.37	3	In
31	ReO_2	1.91	3	In	2.63	3	In	18.12	3	In
32	Rh_2O_3	2.66	3	In	3.66	1	Out	42.91	2	Out
33	RuO_2	1.79	3	In	2.88	1	Out	26.59	3	In
34	Sc_2O_3	2.73	3	In	2.15	3	In	5.56	3	In
35	SrO	3.46	3	In	2.66	3	In	23.38	3	In
36	Ta_2O_3	2.71	3	In	2.94	3	In	20.84	3	In
37	TcO_2	1.95	3	In	2.03	3	In	14.31	3	In
38	Tl_2O	4.22	2	Out	5.25	1	Out	70.80	1	Out
39	Tl_2O_3	2.72	3	In	2.96	3	In	15.48	3	In
40	WO_2	1.91	3	In	3.00	3	In	21.67	3	In
41	WO_3	0.42	1	Out	EA	-		EA	-	-
42	Yb_2O_3	2.71	3	In	2.89	3	In	EA	-	-

EA: Experimental value available that's why not used for prediction.

4. Overview and conclusion

The present work is divided in three major parts:

i) Development of single line QSTR models for three individual endpoints (pEC_{50}E.coli, pEC_{50}HaCaT and %EI_Zebrafish),

ii) Development of i-QSTTR models (*E.coli*-HaCaT, *E.coli*-Zebrafish and HaCaT-Zebrafish),

iii) External dataset prediction using single line models followed by two layered screening for confident and reliable predictions of a huge number of the true external dataset. The major obtained outcomes are discussed below:

1. Based on single endpoint QSTR models, we can infer that
   a. the oxidation state of the metal,
   b. the measure electronegativity of the metal oxide, and
   c. core count of the metal play significant roles on the toxicity of the MNPs irrespective of species and response.
2. Interspecies models helped in extrapolating data from one species to another species as well as in better understanding of the intercorrelation between two species.
3. Other than the response (toxicity) descriptors for the i-QSTTR models, we can also notice the predominance of same properties evolved from the single species QSTR models, *i.e.,* metal oxidation state and electronegativity. Thus, it can be concluded that these two properties of MNPs are vital for contributing to the toxicity.
4. The confident prediction of an external set of 42 MNPs can be used as data gap filling for respective endpoints for which no experimental data are available. The obtained toxicity as well as the predicted property of such a large number of MNPs will be a good source of data gap filling for regulatory agencies along with industries.
5. Finally, we can conclude that the second-generation periodic table-based descriptors along with the first-generation ones, which are derived quickly from the periodic table information, have the ability to encode the toxicity features of MNPs efficiently as compared to quantum chemical descriptors involving time consuming computations and physicochemical descriptors involving experiments. The developed descriptors also capable of developing statistically robust and acceptable as well as mechanistically interpretable QSTR models.

Conflicts of interest

The authors declare no conflict of interest.

References

1. A. Abdelhalim, A. Abdellah, G. Scarpa and P. Lugli, Metallic nanoparticles functionalizing carbon nanotube networks for gas sensing applications, Nanotechnology, 2014, 25, 055208.

2. T. Kim and T. Hyeon, Applications of inorganic nanoparticles as therapeutic agents, Nanotechnology, 2013, 25, 012001.

3. H. L. Karlsson, P. Cronholm, J. Gustafsson and L. Moller, Copper oxide nanoparticles are highly toxic: a comparison between metal oxide nanoparticles and carbon nanotubes, Chem. Res. Toxicol., 2008, 21, 1726-1732.

4. F. Gottschalk and B. Nowack, The release of engineered nanomaterials to the environment, J. Environ. Monitor., 2011, 13, 1145-1155.

5. K. A. Clark, R. H. White and E. K. Silbergeld, Predictive models for nanotoxicology: current challenges and future opportunities, Regul. Toxicol. Pharmacol., 2011, 59, 361-363.

6. L. Canesi, C. Ciacci, M. Betti, R. Fabbri, B. Canonico, A. Fantinati, A. Marcomini and G. Pojana, Immunotoxicity of carbon black nanoparticles to blue mussel hemocytes, Environ. Int., 2008, 34, 1114-1119.

7. K. Roy, S. Kar and R. N. Das, Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment, Academic press, 2015.

8. J. C. Dearden, The history and development of quantitative structure-activity relationships (QSARs), IJQSPR, 2016, 1, 1-44.

9. W. M. Berhanu, G. G. Pillai, A. A. Oliferenko and A. R. Katritzky, Quantitative structure–activity/property relationships: the ubiquitous links between cause and effect, ChemPlusChem, 2012, 77, 507-517.

10. A. Tropsha, Best practices for QSAR model development, validation, and exploitation, Mol. Inform., 2010, 29, 476-488.

11. K. Roy and R. N. Das, QSTR with extended topochemical atom (ETA) indices. 16. Development of predictive classification and regression models for toxicity of ionic liquids towards Daphnia magna, J. Hazard. Mater., 2013, 254, 166-178.

12. R. Gozalbes and J. V. de Julián-Ortiz, Applications of Chemoinformatics in Predictive Toxicology for Regulatory Purposes, Especially in the Context of the EU REACH Legislation, IJQSPR, 2018, 3, 1-24.

13. K. P. Singh and S. Gupta, Nano-QSAR modeling for predicting biological activity of diverse nanomaterials, RSC Advances, 2014, 4, 13215-13230

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

15. D. Fourches, D. Pu, C. Tassa, R. Weissleder, S. Y. Shaw, R. J. Mumper and A. Tropsha, Quantitative nanostructure– activity relationship modeling., ACS Nano, 2010, 4, 5703-5712.

16. V. C. Epa, F. R. Burden, C. Tassa, R. Weissleder, S. Shaw and D. A. Winkler, Modeling biological activities of nanoparticles, Nano Lett., 2012, 12, 5808-5812.

17. Y. T. Chau and C. W. Yap, Quantitative nanostructure–activity relationship modelling of nanoparticles, RSC Adv., 2012, 2, 8489-8496.

18. C.-Y. Shao, S.-Z. Chen, B.-H. Su, Y. J. Tseng, E. X. Esposito and A. J. Hopfinger, Dependence of QSAR models on the selection of trial descriptor sets: a demonstration using nanotoxicity endpoints of decorated nanotubes, J. Chem. Inf. Model., 2013, 53, 142-158.

19. D. Martin, U. Maran, S. Sild and M. Karelson, QSPR modeling of solubility of polyaromatic hydrocarbons and fullerene in 1-octanol and n-heptane, J. Phys. Chem. B 2007, 111, 9853-9857.

20. N. Sizochenko, B. Rasulev, A. Gajewicz, V. Kuz'min, T. Puzyn and J. Leszczynski, From basic physics to mechanisms of toxicity: The "liquid drop" approach applied to develop predictive classification models for toxicity of metal oxide nanoparticles., Nanoscale, 2014, 6, 13986-13993.

21. A. Gajewicz, M. T. D. Cronin, B. Rasulev, J. Leszczynski and T. Puzyn, Novel approach for efficient predictions properties of large pool of nanomaterials based on limited set of species: nano-read-across, Nanotechnology, 2014, 26, 015701.

22. D.W. Boukhvalov and T. H. Yoon, Development of Theoretical Descriptors for Cytotoxicity Evaluation of Metallic Nanoparticles, Chem. Res. Toxicol. 2017, 30, 1549-1555.

23. H. K. Shin, K. Y. Kim, J. W. Park and K. T. No, Use of metal/metal oxide spherical cluster and hydroxyl metal coordination complex for descriptor calculation in development of nanoparticle cytotoxicity classification model, SAR QSAR Environ. Res. 2017, 28, 875-888.

24. F. Luan, V.V. Kleandrova, H. González-Díaz, J.M. Ruso, A. Melo, A. Speck-Planche and M.N.D.S. Cordeiro. Computer-aided nanotoxicology: assessing cytotoxicity of nanoparticles under diverse experimental conditions by using a novel QSTR perturbation approach. Nanoscale 2014, 6, 10623-10630.

25. V.V. Kleandrova, F. Luan, H. González-Díaz, J.M. Ruso, A. Speck-Planche and M.N.D.S. Cordeiro, Computational Tool for Risk Assessment of Nanomaterials: Novel QSTR-Perturbation Model for Simultaneous Prediction of Ecotoxicity and Cytotoxicity of Uncoated and Coated Nanoparticles under Multiple Experimental Conditions, Environ. Sci. Technol. 2014, 48, 14686-14694

26. R. Concua, V.V. Kleandrova, A. Speck-Planche and M.N.D.S. Cordeiro. Probing the toxicity of nanoparticles: a unified in silico machine learning model based on perturbation theory, Nanotoxicology 2017, 11, 891-906.

27.   A. Mikolajczyk, A. Gajewicz, B. Rasulev, N. Schaeublin, E. Maurer-Gardner, S. Hussain, J. Leszczynski and T. Puzyn, Zeta potential for metal oxide nanoparticles: a predictive model developed by a nano-quantitative structure–property relationship approach., Chem. Mater., 2015, 27, 2400-2407.

28.   X. R. Xia, N. A. Monteiro-Riviere, S. Mathur, X. Song, L. Xiao, S. J. Oldenberg, B. Fadeel and J. E. Riviere, Mapping the surface adsorption forces of nanomaterials in biological systems., ACS Nano, 2011, 5, 9074-9081.

29. S. Kar, R. N. Das, K. Roy and J. Leszczynski, Can Toxicity for Different Species be Correlated?: The Concept and Emerging Applications of Interspecies Quantitative Structure-Toxicity Relationship (i-QSTR) Modeling, IJQSPR, 2016, 1, 23-51.

30. S. Cassani, S. Kovarich, E. Papa, P. P. Roy, L. van der Wal and P. Gramatica, Daphnia and fish toxicity of (benzo) triazoles: Validated QSAR models, and interspecies quantitative activity–activity modelling., J. Hazard. Mater., 2013, 258, 50-60.

31. J. Devillers and H. Devillers, Prediction of acute mammalian toxicity from QSARs and interspecies correlations, SAR QSAR Environ. Res., 2009, 20, 467-500.

32. S. Önlü and M. T. Saçan, Toxicity of contaminants of emerging concern to Dugesia japonica: QSTR modeling and toxicity relationship with Daphnia magna, J. Hazard. Mater., 2018, 351, 20-28.

33. S. Kar and K. Roy, First report on development of quantitative interspecies structure–carcinogenicity relationship models and exploring discriminatory features for rodent carcinogenicity of diverse organic chemicals using OECD guidelines, Chemosphere, 2012, 87, 339-355.

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

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

36. N. Basant and S. Gupta, Modeling uptake of nanoparticles in multiple human cells using structure-activity relationships and intercellular uptake correlations, Nanotoxicology, 2017, 11, 20-30.

37. A. Gajewicz, N. Schaeublin, B. Rasulev, S. Hussain, D. Leszczynska, T. Puzyn and J. Leszczynski, Towards understanding mechanisms governing cytotoxicity of metal oxides nanoparticles: Hints from nano-QSAR studies, Nanotoxicology, 2015, 9, 313-325.

38. S. Lin, Y. Zhao, Z. Ji, J. Ear, C. H. Chang, H. Zhang, C. Low-Kam, K. Yamada, H. Meng and X. Wang, Zebrafish high-throughput screening to study the impact of dissolvable metal oxide nanoparticles on the hatching enzyme, ZHE1., Small, 2013, 9, 1776-1785.

39. R. Leardi, Genetic algorithms in chemometrics and chemistry: a review, J. Chemometr., 2001, 15, 559-569.

40. S. Wold, M. Sjöström and L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometr. Intell. Lab., 2001, 58, 109-130.

41. R. Todeschini, D. Ballabio and F. Grisoni, Beware of Unreliable Q2! A Comparative Study of Regression Metrics for Predictivity Assessment of QSAR Models, J. Chem. Inf. Model. 2016, 56, 1905-1913.

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

43. P. Gramatica, Principles of QSAR models validation: internal and external, QSAR Comb. Sci. 2007, 26, 694-701.

44. K. Roy, S. Kar and P. Ambure, On a simple approach for determining applicability domain of QSAR models, Chemometr. Intell. Lab., 2015, 145, 22-29.

45. K. Roy, P. Ambure and S. Kar, How Precise Are Our Quantitative Structure-Activity Relationship Derived Predictions for New Query Chemicals?, ACS Omega, 2018, 3, 11392–11406, DOI: 10.1021/acsomega.8b01647

46. P. Khanna, C. Ong, B. H. Bay and G. H. Baeg, Nanotoxicity: an interplay of oxidative stress, inflammation and cell death, Nanomaterials, 2015, 5, 1163-1180.

47. M. Auffan, W. Achouak, J. Rose, M.-A. Roncato, C. Chanéac, D. T. Waite, A. Masion, J. C. Woicik, M. R. Wiesner and J.-Y. Bottero, Relation between the redox state of iron-based nanoparticles and their cytotoxicity toward Escherichia coli, Environ. Sci. Technol., 2008, 42, 6730-6735.

48. W. Lin, I. Stayton, Y.-w. Huang, X.-D. Zhou and Y. Ma, Cytotoxicity and cell membrane depolarization induced by aluminum oxide nanoparticles in human lung epithelial cells A549, Toxicol. Environ. Chem., 2008, 90, 983-996.

49. A. L. Neal, What can be inferred from bacterium–nanoparticle interactions about the potential consequences of environmental exposure to nanoparticles?, Ecotoxicology, 2008, 17, 362.

50. J. Lovrić, S. J. Cho, F. M. Winnik and D. Maysinger, Unmodified cadmium telluride quantum dots induce reactive oxygen species formation leading to multiple organelle damage and cell death, Chem. Biol., 2005, 12, 1227-1234.

51. E. Burello and A. P. Worth, A theoretical framework for predicting the oxidative stress potential of oxide nanoparticles, Nanotoxicology, 2011, 5, 228-235.

52. P. Somasundaran, X. Fang, S. Ponnurangam and B. Li, Nanoparticles: characteristics, mechanisms and modulation of biotoxicity, KONA Powder Part. J., 2010, 28, 38-49.

53. Q. Chang, L. Yan, M. Chen, H. He and J. Qu, Bactericidal mechanism of Ag/Al2O3 against Escherichia coli, Langmuir, 2007, 23, 11197-11199.

54. W. A. Daoud, J. H. Xin and Y.-H. Zhang, Surface functionalization of cellulose fibers with titanium dioxide nanoparticles and their combined bactericidal activities, Surf. Sci., 2005, 599, 69-75.

55. N. Sizochenko, A. Gajewicz, J. Leszczynski and T. Puzyn, Causation or only correlation? Application of causal inference graphs for evaluating causality in nano-QSAR models, Nanoscale, 2016, 8, 7203-7208.

56. N. Akarachantachote, S. Chadcham and K. Saithanu, Cutoff threshold of variable importance in projection for variable selection, Int. J. Pure Appl. Math., 2014, 94, 307-322.

57. J. E. Jackson, A user's guide to principal components.Journal, 2005, 587.

58. T. Ozben, Oxidative stress and apoptosis: impact on cancer therapy, J. Pharm. Sci., 2007, 96, 2181-2196.

59. S. H. Lee, H. R. Lee, Y.-R. Kim and M.-K. Kim, Toxic response of zinc oxide nanoparticles in human epidermal keratinocyte HaCaT cells, Toxicol. Environ. Health Sci., 2012, 4, 14-18.

60. M. Cho, H. Chung, W. Choi and J. Yoon, Linear correlation between inactivation of E. coli and OH radical concentration in TiO2 photocatalytic disinfection, Water Res., 2004, 38, 1069-1077.

61. J. Sawai, H. Kojima, H. Igarashi, A. Hashimoto, S. Shoji, T. Sawaki, A. Hakoda, E. Kawada, T. Kokugan and M. Shimizu, Antibacterial characteristics of magnesium oxide powder, World J. Microbiol. Biotechnol., 2000, 16, 187-194.

62. Y. Kikuchi, K. Sunada, T. Iyoda, K. Hashimoto and A. Fujishima, Photocatalytic bactericidal effect of TiO2 thin films: dynamic view of the active oxygen species responsible for the effect, J. Photochem. Photobiol., 1997, 106, 51-56.

63. J. Portier, G. Campet, A. Poquet, C. Marcel and M. A. Subramanian, Degenerate semiconductors in the light of electronegativity and chemical hardness, Int. J. Inorg. Mater., 2001, 3, 1039-1043.

64. W. H. Koppenol, The Haber-Weiss cycle–70 years later, Redox Rep., 2001, 6, 229-234.

65. S. J. Stohs and D. Bagghi, Oxidative Mechanisms in the Toxicity of Metal Ions, Free Radical Bio. Med., 2005, 39, 1267-1268.

66. B. Halliwell and J. M. C. Gutteridge, Free radicals in biology and medicine, Oxford University Press, USA, 2015.

67. N. Basant and S. Gupta, Multi-target QSTR modeling for simultaneous prediction of multiple toxicity endpoints of nano-metal oxides, Nanotoxicology, 2017, 11, 339-350.

68. M. Lundqvist, J. Stigler, G. Elia, I. Lynch, T. Cedervall and K. A. Dawson, Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts, P. Natl. A. of Sci., 2008.

69. K. Sano, K. Inohaya, M. Kawaguchi, N. Yoshizaki, I. Iuchi and S. Yasumasu, Purification and characterization of zebrafish hatching enzyme–an evolutionary aspect of the mechanism of egg envelope digestion, FEBS J., 2008, 275, 5934-5946.

70. A. Okada, K. Sano, K. Nagata, S. Yasumasu, J. Ohtsuka, A. Yamamura, K. Kubota, I. Iuchi and M. Tanokura, Crystal structure of zebrafish hatching enzyme 1 from the zebrafish Danio rerio, J. Mol. Biol., 2010, 402, 865-878.

71. D. Holt, L. Magos and M. Webb, The interaction of cadium-induced rat renal metallothionein with bivalent mercury in vitro, Chem-biol. Interact., 1980, 32, 125-135.

72. F. X. Gomis-Rüth, F. Grams, I. Yiallouros, H. Nar, U. Küsthardt, R. Zwilling, W. Bode and W. Stöcker, Crystal structures, spectroscopic features, and catalytic properties of cobalt (II), copper (II), nickel (II), and mercury (II) derivatives of the zinc endopeptidase astacin. A correlation of structure and proteolytic activity, J. Biol.Chem., 1994, 269, 17111-17117.

73. D. R. Holland, A. C. Hausrath, D. Juers and B. W. Matthews, Structural analysis of zinc substitutions in the active site of thermolysin, Protein Sci., 1995, 4, 1955-1965.

74. E. Casals, T. Pfaller, A. Duschl, G. J. Oostingh and V. Puntes, Time evolution of the nanoparticle protein corona, ACS Nano, 2010, 4, 3623-3632.

75. L. Fei and S. Perrett, Effect of nanoparticles on protein folding and fibrillogenesis, Int. J. Mol. Sci., 2009, 10, 646-655.

76. X. Zhu, L. Zhu, Z. Duan, R. Qi, Y. Li and Y. Lang, Comparative toxicity of several metal oxide nanoparticle aqueous suspensions to Zebrafish (Danio rerio) early developmental stage, J. Environ. Sci. Health A, 2008, 43, 278-284.

77. W. Bai, Z. Zhang, W. Tian, X. He, Y. Ma, Y. Zhao and Z. Chai, Toxicity of zinc oxide nanoparticles to zebrafish embryo: a physicochemical study of toxicity mechanism, J. Nanopart. Res., 2010, 12, 1645-1654.

78. X. Zhu, S. Tian and Z. Cai, Toxicity assessment of iron oxide nanoparticles in zebrafish (Danio rerio) early life stages, PLoS One, 20102, 7, e46286.

79. J. Cheng, E. Flahaut and S. H. Cheng, Effect of carbon nanotubes on developing zebrafish (Danio rerio) embryos, Environ. Toxicol. Chem., 2007, 26, 708-716.

80. Y. Mu, F. Wu, Q. Zhao, R. Ji, Y. Qie, Y. Zhou, Y. Hu, C. Pang, D. Hristozov and J. P. Giesy, Predicting toxic potencies of metal oxide nanoparticles by means of nano-QSARs, *Nanotoxicology*, 2016, **10**, 1207-1214.

81. N. Sizochenko, D. Leszczynska and J. Leszczynski, Modeling of Interactions between the Zebrafish Hatching Enzyme ZHE1 and A Series of Metal Oxide Nanoparticles: Nano-QSAR and Causal Analysis of Inactivation Mechanisms, Nanomaterials, 2017, 7, 330.