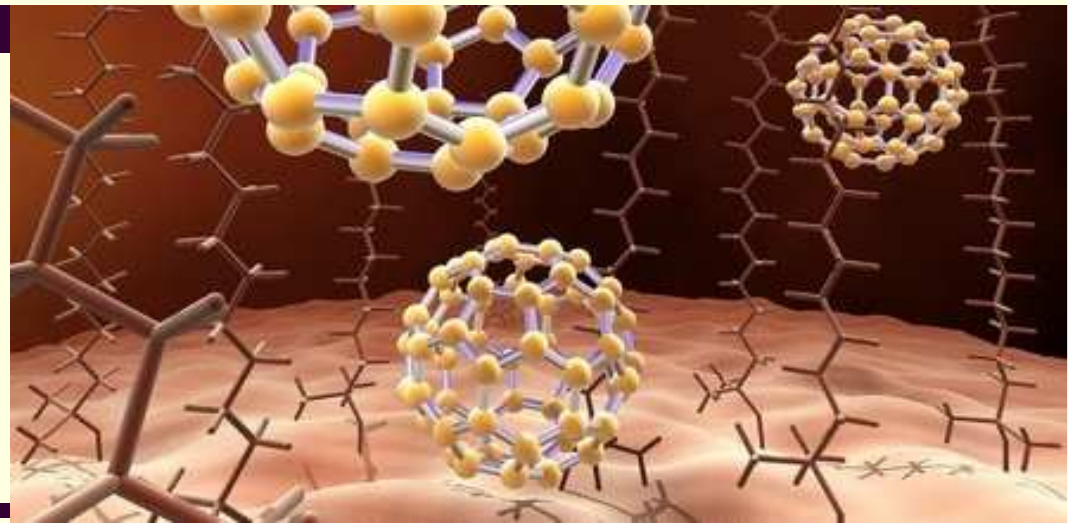


*EuroNanoForum 2009*

# **Challenges in Modeling Properties of Nanomaterials**

*Prof. Danuta Leszczynska*

**Interdisciplinary Center  
for  
Nanotoxicity**



*Jackson State University  
2009*

# Layout

---

- Why do we feel that modeling of nanoparticles is useful, and what do we expect to get out of our models...
- What is QSAR and how helpful it could be in modeling of nanoparticles
- Examples of our work...
- How good our models are...
- Future plans..

# What do we need for developing our models....

---

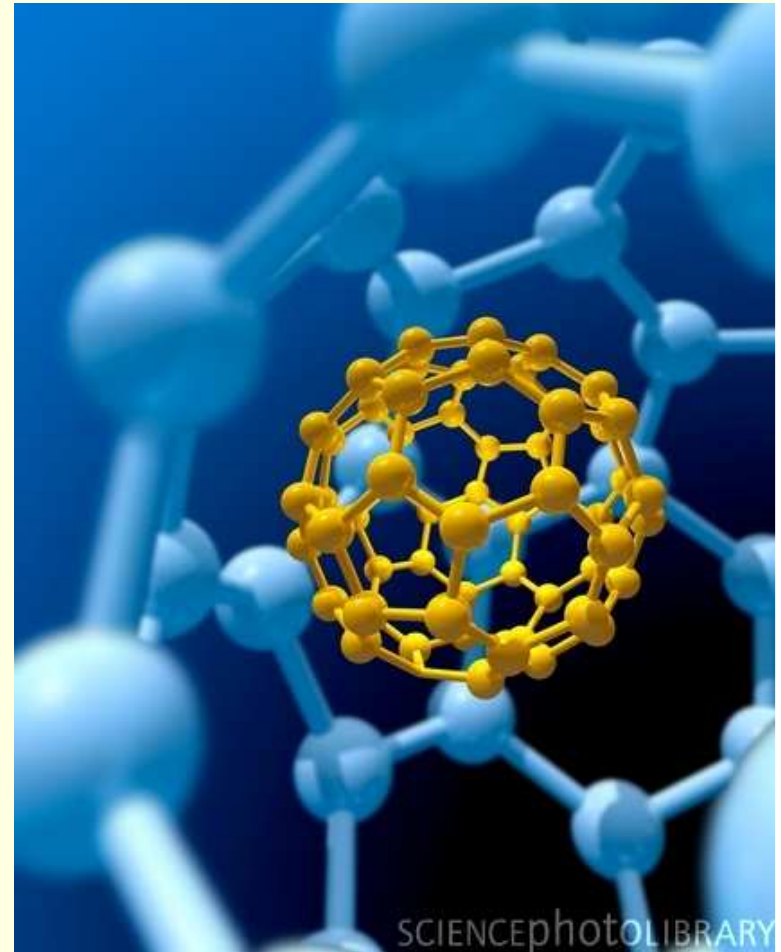
- Information about structures and basic properties of studied nanoparticles...they are scarce...
- Experimental data about nanoparticles... their possible toxicity, interactions, and other properties are not easily available either, and many times not compatible, if they are obtained from different studies..
- Experiments are expensive; countless experiments are more expensive, and they usually cover only one experimental set-up....

# Objectives behind our models

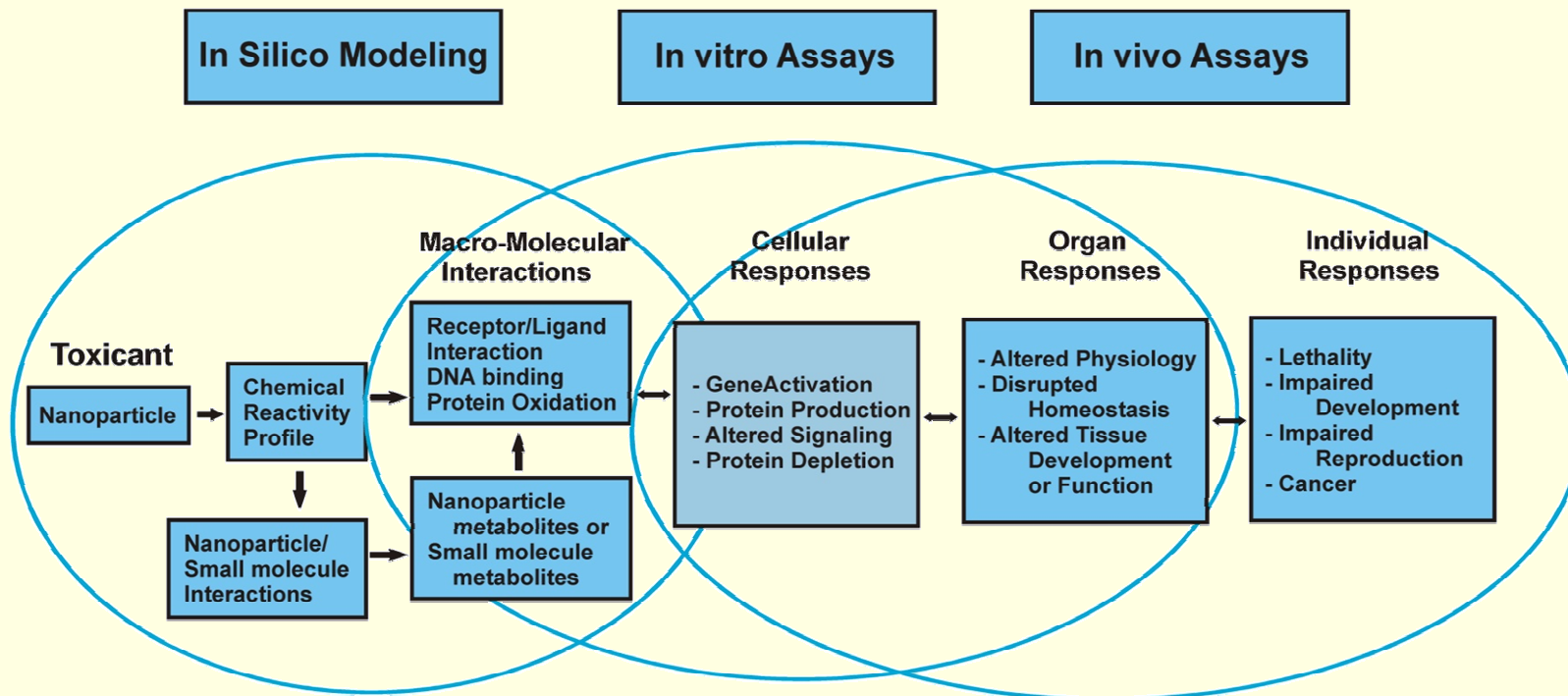
---

- Physical properties of nanomaterials
- Toxicological aspects of nanomaterials
- How to predict the environmental risk and toxic effects of nanomaterials

**This presentation is focused on metal oxides, fullerenes and carbon nanotubes**

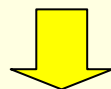


# An Example of Toxicity Pathways for Studied Nanoparticles

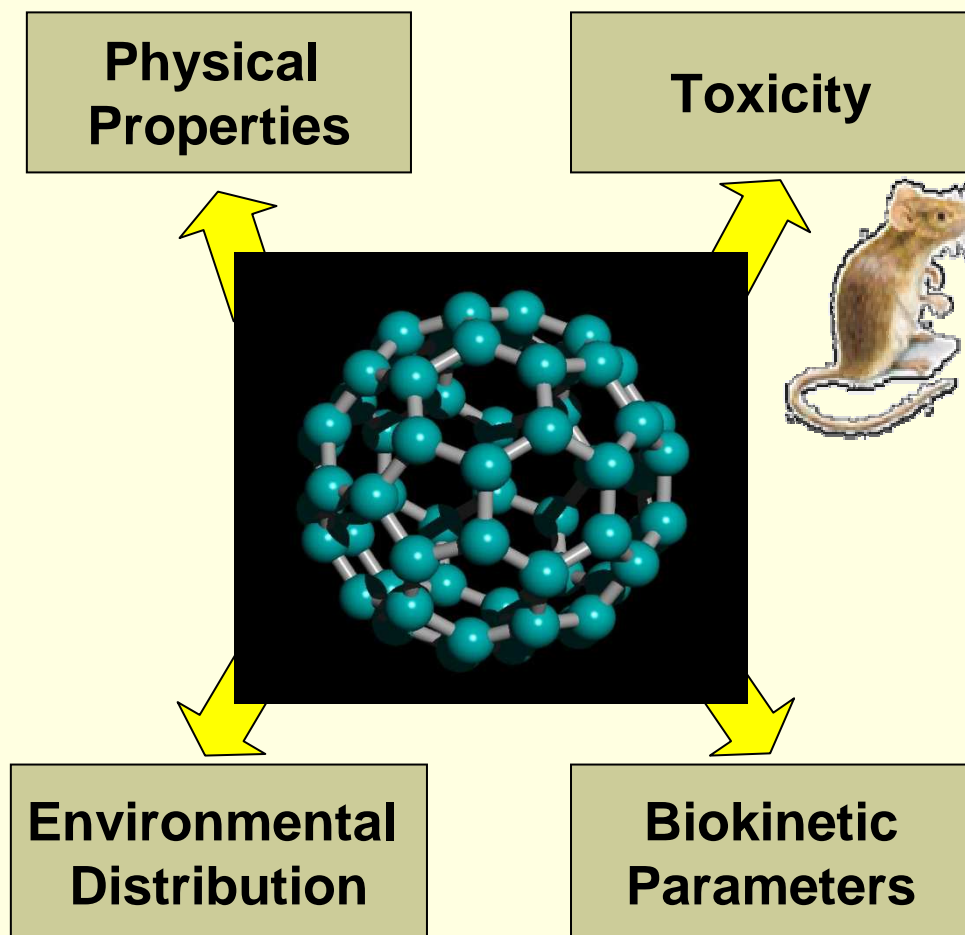


# *In silico* approaches

*In silico* methods can be used to predict the effects of chemicals on human health and the environment, as well as their distribution and fate within the environment and biological organisms



- **SARs: Structure-Activity Relationships**
- **AND**
- **QSARs: Quantitative Structure-Activity Relationships**
  
- ... are theoretical models that relate the structure of chemicals to their properties



---

**QSAR – what is this?**

*Quantitative Structure-Activity  
Relationship*

# Question

---

How to predict the properties (e.g. solubility) and activities (e.g. fate, toxicity) of unknown chemicals (especially organics) in the environment and in engineering systems?

*Precursor: knowing that polar molecules will dissolve in polar solvents, we can predict by studying structure of unknown molecule, if will be polar or not, and this way predict solubility in water.*

# Vocabulary

---

## ■ Structure

- Molecular characteristic
- (structure, bonds)

## Activity

Biological effect  
(toxicity, biotransformation)

## Property

Significant characteristics  
(solubility, volatility, partitioning, etc)

# Vocabulary

---

## ■ Structure:

Look at group of similar molecules

- Benzene
- Chloro benzene
- Bromobenzene
- o chlorobenzene

# Vocabulary

---

## **Activity:**

Find toxicity (or biotransformation, or any other biological effect) of each compound from the group

# Vocabulary

---

## **Property**

Find significant characteristics (solubility, partitioning, etc) for each compound from the group

# Developed relationships

---

- Structure-activity
- Property-activity
- Structure-property
- Property-property relationships

**All are referred as**  
**Quantitative-Structure-Activity Relationship (QSAR)**  
**or**  
**Quantitative-Structure-Property Relationship (QSPR)**

# QSAR

## *Quantitative Structure-Activity Relationship*

---

- A QSAR is a ***mathematical relationship between a biological activity of a molecular system and its geometric, chemical or physical characteristics.***
- When relationship is found, the QSAR model is determined, and can be then used to evaluate the activity of new compounds.
- Validation of QSAR model is needed before it could be used for **prediction** of the particular physical properties or biological activities of related compounds or drug candidates before they are put through expensive and time-consuming biological testing. In some cases, only computed values need to be known to make an assessment.

# ..therefore..

---

- QSAR's most general mathematical form is:

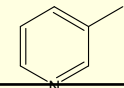
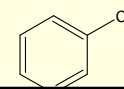
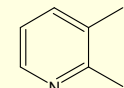
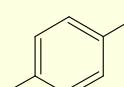
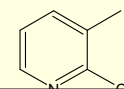
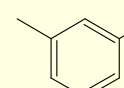
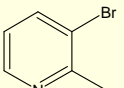
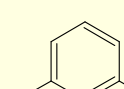
**Activity =  $f$  (physiochemical properties and/or structural properties)**

The problem of QSAR is to find coefficients  $C_0, C_1, \dots, C_n$  such that:

- **$Biological\ activity = C_0 + (C_1 * P_1) + \dots + (C_n * P_n)$**

and the prediction error is minimized for a list of given  $m$  compounds.

# QSAR/QSPR: General Approach

	A	Structure	Descriptors			
Training	-		-	-	-	-
	-		-	-	-	-
	-		-	-	-	-
	-		-	-	-	-
Test	-		-	-	-	-
	-		-	-	-	-
New	?		-	-	-	-
	?		-	-	-	-

Model  
 $F: A=F(S)$

Predictivity  
 $\Delta A$

Prediction



# Classical QSAR

*Properties that can be predicted by QSAR analysis:*

---

## *Physico-chemical properties:*

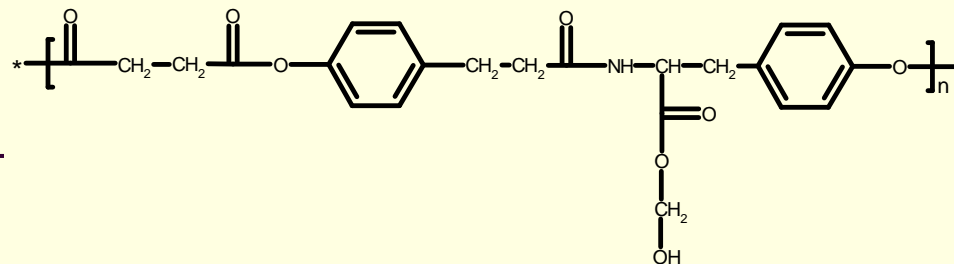
**Boiling points, melting points, density, viscosity, surface tension, solubility in various solvents, lipophilicity, magnetic susceptibility, retention indices, dipole moments, enthalpy of formation, etc.**

*Biological activity (toxicity):* **IC<sub>50</sub>, EC<sub>50</sub>, LD<sub>50</sub>, etc**

# Types of Molecular Descriptors

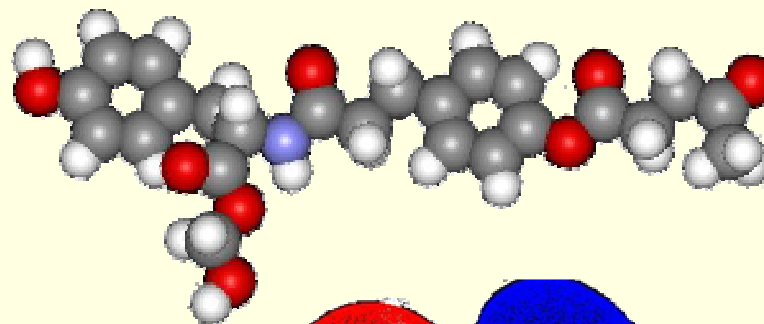
**Constitutional, Topological**

**2-D structural formula**

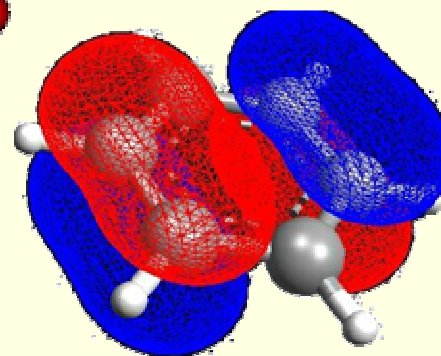


**Geometrical**

**3-D shape and structure**

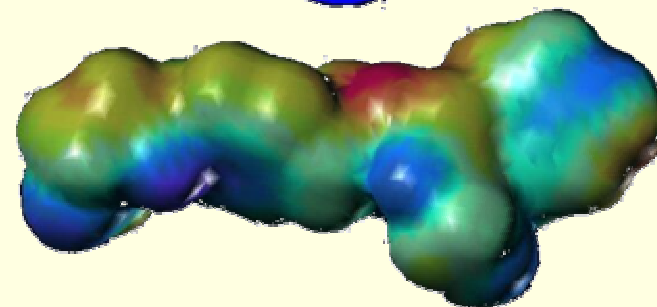


**Quantum Chemical**



**Electrostatic**

**Hybrid descriptors**



# Extension of the traditional (Q)SAR paradigm to nanoparticles

---

- **Aim:** To develop a rational structure-activity based paradigm for assessing the toxicological hazard of nanoparticles.
  
- **What do we need?**
- • **Data:** from the characterization of nanoparticles and on toxicological endpoints of interest (lethal dose toxicity, oxidative stress/inflammation, and etc)
- • **Structural descriptors:** physicochemical properties and calculated structural descriptors
- • **Methods:**
  - • Multivariate statistical methods that relate the structural descriptors to physicochemical properties and toxicological endpoints
  - • Molecular modeling methods

# Examples of QSAR modeling work done by our group

---

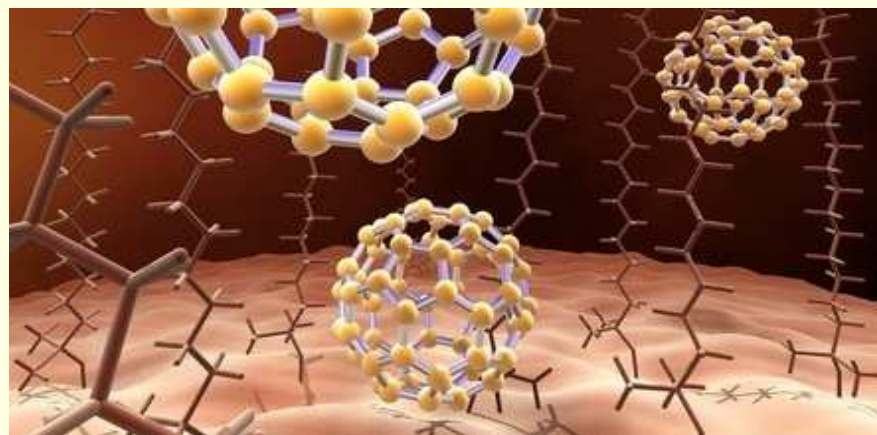
# I. Solubility of fullerene C<sub>60</sub> in various solvents

Fullerenes are sparingly soluble in many solvents.

The dependence of fullerene's solubility on molecular structure of the solvent must be understood in order to efficiently separate different members of the fullerene family from each other and from their precursors or derivatives.

The data on solubility of the fullerene C<sub>60</sub> in organic solvents could help to develop its various applications in chemistry and technology.

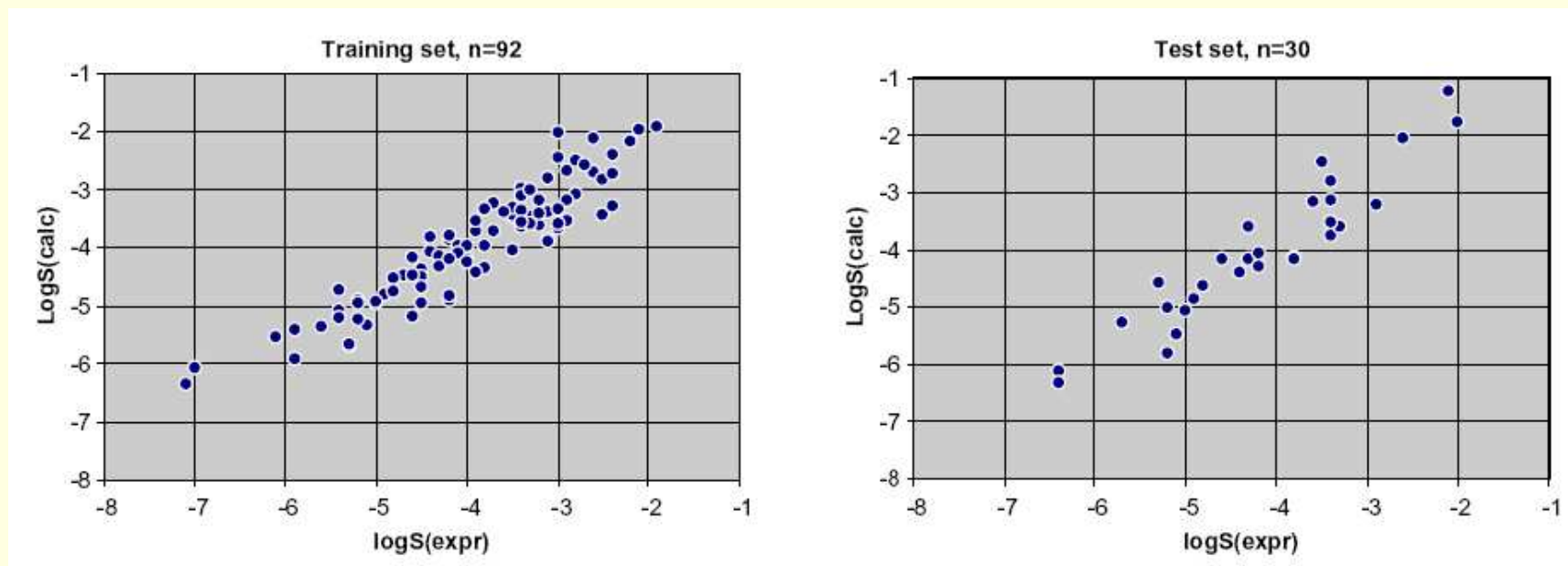
**This study was aimed to find better relationships, built simple and transparent models to understand a nature of solubility of C<sub>60</sub> in various solvents.**



# I. QSPR modeling of fullerene C<sub>60</sub> solubility in organic solvents using SMILES based optimal descriptors. (extended set, 122 solvents)

Optimal descriptors, calculated with Simplified Molecular Input Line Entry System, SMILES, have been used for modeling solubility of fullerene C<sub>60</sub> in organic solvents. Statistical characteristics of the model,  $n = 92$ ,  $R^2 = 0.861$ ,  $Q^2 = 0.854$ ,  $s = 0.401$ ,  $F = 558$  (training set)  
 $n = 30$ ,  $R^2 = 0.891$ ,  $R^2_{\text{pred}} = 0.875$ ,  $s = 0.435$ ,  $F = 228$  (test set).

The applied approach entirely based on topological data provides a reliable model for solubility of C<sub>60</sub>.

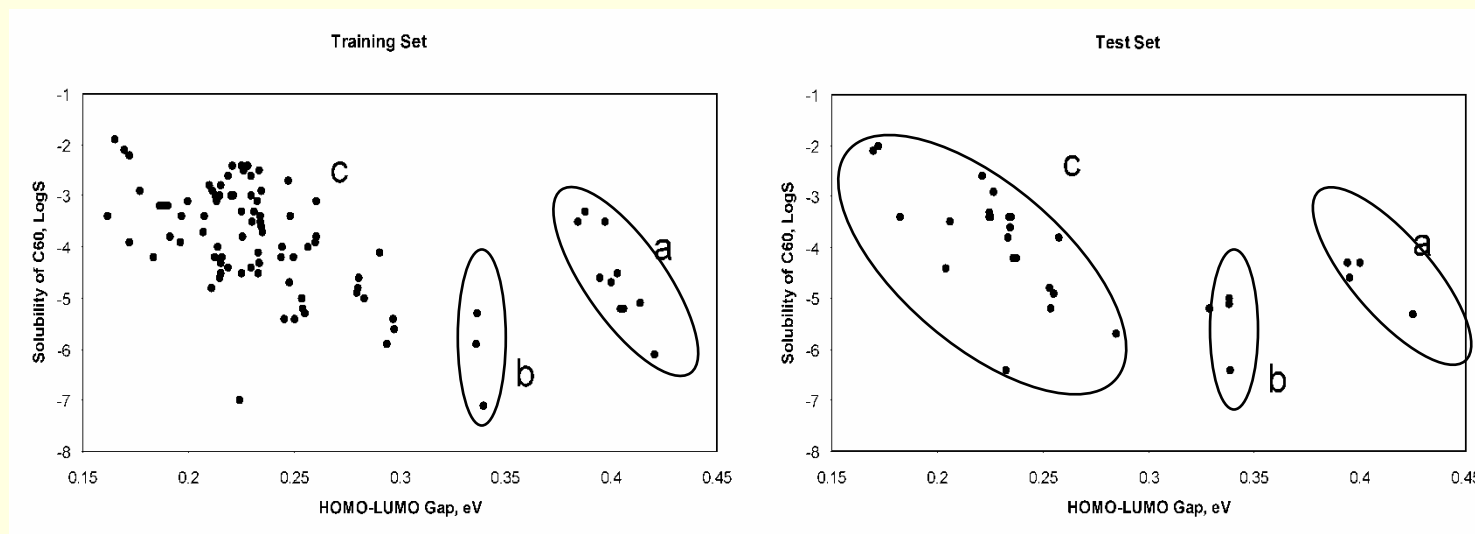


Plot of experimental versus calculated solubility of fullerene C<sub>60</sub> in organic solvents for the training and test sets (S, 10<sup>4</sup> molar fraction, T = 298 K).

## II. Quantum-chemical and QSPR modeling of fullerene C<sub>60</sub> solubility

### first way of solubility estimation – solvent's descriptor calculations and QSPR

Quantum-chemical calculations and QSPR modeling were used to model the solubility of fullerene C<sub>60</sub> in 122 organic solvents. A genetic algorithm and multiple regression analysis (GA-MLRA) were applied to generate correlation models. The best performance is accomplished by the four-variable MLRA model with prediction coefficient  $r^2_{\text{test}}=0.903$ . The present study reveals a correlation of highest occupied molecular orbital energy (HOMO), certain heteroatom fragments and geometrical parameters with solubility.



Plot of C<sub>60</sub> solubility data versus HOMO-LUMO gap for both sets.

(a)- alkanes, (b)-alcohols, (c)- other compounds (heteroatom-containing).

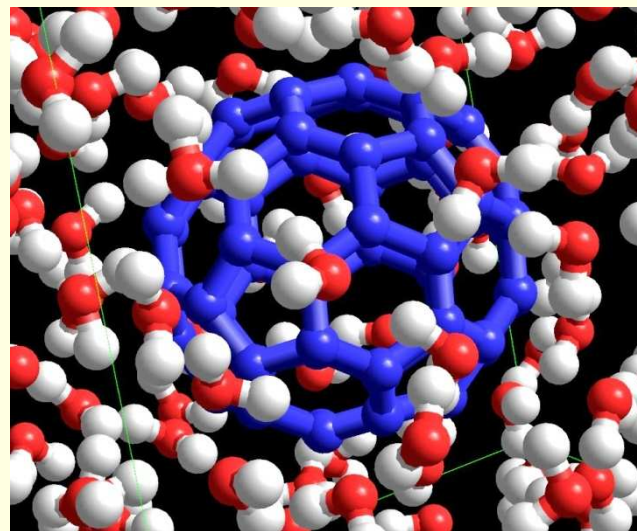
# Quantum chemical and QSPR modeling of fullerene C<sub>60</sub> solubility

another way of solubility estimation – direct quantum-chemical calculation of C<sub>60</sub>-solvent interactions and solubility

Direct quantum-chemical calculation has been performed for estimation of C<sub>60</sub> solubility in various organic solvents by using polarized continuum model's technique during *ab initio* calculations. For this purpose we used SCI-PCM and CPCM models. Another solubility model, Onsager solvent model, showed unreliable results and therefore fails to estimate solubility.

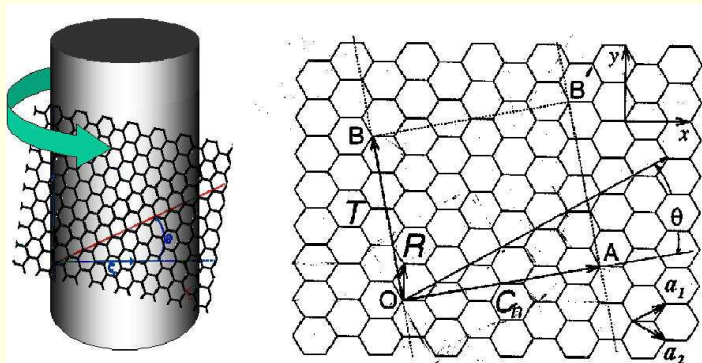
Because of these calculations a pretty time-consuming, the study is still in progress. We estimate to finish these calculations in summer 2009.

Preliminary calculations showed quite reliable results for PCM model allowing us to perform direct C<sub>60</sub> solubility estimation from quantum-chemical calculations in the future.



Tetyana Petrova, Bakhiyor Rasulev, Andrey Toropov, Danuta Leszczynska, Jerzy Leszczynski, Direct *ab initio* calculation of C<sub>60</sub>-solvent interactions and predicting the solubility in various organic solvents, 2009, (to be prepared)

### III. Predicting water solubility and octanol water partition coefficient for **carbon nanotubes** based on the chiral vector



**Nanotube**

**2D-projection**

If  $C_n$  is the chiral vector then it is defined as

$$C_n = n\hat{a}_1 + m\hat{a}_2$$

Obtained statistical results in a case of the **water solubility**:

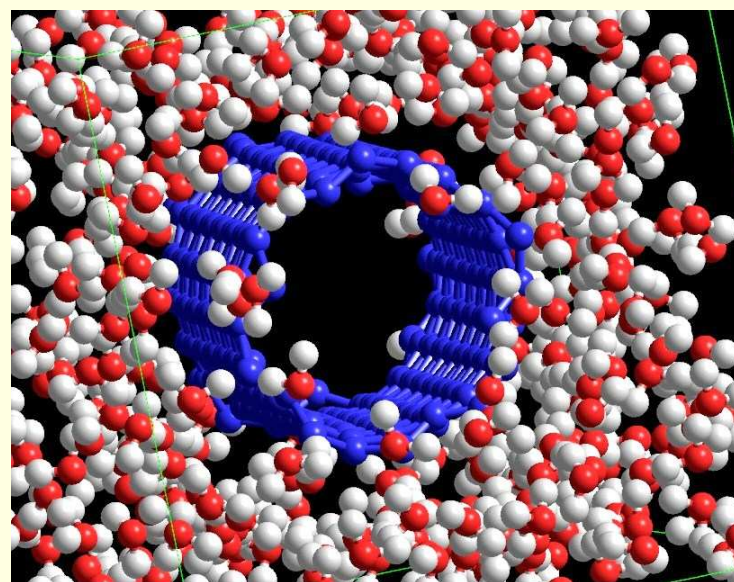
$n=8$ ,  $r^2=0.999$ ,  $s=0.053$ ,  $F=126611$  (training set);  
 $n=8$ ,  $r^2=0.999$ ,  $s=0.093$ ,  $F=674556$  (test set)

and in case of **octanol water partition coefficient**:

$n=8$ ,  $r^2=0.999$ ,  $s=0.366$ ,  $F=2927$  (training set);  
 $n=8$ ,  $r^2=0.999$ ,  $s=0.287$ ,  $F=5928$  (test set).

Components of the chiral vector of carbon nanotubes have been examined in role of structural descriptors.

Two-variable models of water solubility and octanol water partition coefficient calculated with components of the chiral vectors have quite good statistical characteristics.



## Results

Two-variable models of the water solubility and octanol-water partition coefficient have been calculated by the least squares method. These models are the following

$$\log S = -5.1041 - 3.5075n - 3.5941m \quad (1)$$

$$\log P = -3.9193 + 3.7703n - 3.6001m \quad (2)$$

The predictions of the water solubility and octanol water partition coefficient based on the components of the chiral vector are quite good. It indicates that water solubility and octanol water partition coefficient of CNTs are determined by features of the rolling up of graphite layer.

## Conclusions

Components of chiral vectors of nanotubes can be used as structural descriptors in multiple regression analysis to predict numerical values of water solubility and octanol water partition coefficient.

Components of chiral vector, experimental and calculated values of water solubility, and octanol water partition coefficient for CNTs under consideration

No.	<i>n</i>	<i>m</i>	Log $S_{\text{expr}}$	Log $S_{\text{calc}}$	Log $P_{\text{expr}}$	Log $P_{\text{calc}}$
1 <sup>a</sup>	9	0	-36.6	-36.6716	29.8	30.0134
2	10	0	-40.1	-40.1791	34.0	33.7837
3 <sup>a</sup>	11	0	-43.6	-43.6866	37.9	37.5540
4 <sup>a</sup>	12	0	-47.1	-47.1941	41.7	41.3243
5 <sup>a</sup>	13	0	-50.7	-50.7016	45.4	45.0946
6	14	0	-54.2	-54.2091	49.1	48.8649
7	15	0	-57.7	-57.7166	52.7	52.6352
8 <sup>a</sup>	16	0	-61.3	-61.2241	56.3	56.4055
9	17	0	-64.8	-64.7316	59.8	60.1758
10 <sup>a</sup>	18	0	-68.4	-68.2391	63.6	63.9461
11	5	5	-40.7	-40.6121	32.3	32.9327
12 <sup>a</sup>	6	6	-47.8	-47.7137	40.5	40.3031
13	7	7	-54.8	-54.8153	48.1	47.6735
14	8	8	-61.9	-61.9169	55.3	55.0439
15 <sup>a</sup>	9	9	-69.0	-69.0185	62.5	62.4143
16	10	10	-76.1	-76.1201	69.6	69.7847

<sup>a</sup> CNTs of the test set.

## 2. Study of metal oxides properties in nano-sized form

---

Metal oxides in nano-sized forms are very widespread among consumer products.

The dependence of metal-oxides behavior depending on molecular structure, metal atom types, agglomeration, size of particle and etc - must be understood in order to describe physico-chemical properties and then toxicity of different members of the metal oxides family.

The presented summary is aimed to find structure-property relationships to understand a nature of metal oxides' behavior depending on micro and macro parameters.



## QSAR model of toxicity towards *E.coli* bacteria for nanosized oxides by SMILES based optimal descriptors

Toxicity of fourteen metal oxides in nano-sized form towards *E.coli* bacteria has been measured. To find the structure-toxicity relationship between investigated nanoparticles and toxicity of them the QSAR technique has been applied.

Statistical quality of quantitative structure–activity relationship obtained by the optimal descriptors calculated from simplified molecular input line entry system (SMILES) notations for toxicity towards *E.coli* bacteria ( $-pLD_{50}$ ) of nanosized oxides are the following:  $n=7$ ,  $r^2=0.99$ ,  $s=0.053$ ,  $F=539$  (training set); and  $n=7$ ,  $r^2=0.82$ ,  $s=0.241$ ,  $F=23$  (test set).

First run of the Monte Carlo optimization gave the following model:

$$-pLD_{50} = 1.323(\pm 0.031) + 0.274(\pm 0.008) * DCW$$

$n=7$ ,  $r^2=0.991$ ,  $s=0.053$ ,  $F=539$  (training set);

$n=7$ ,  $r^2=0.823$ ,  $s=0.241$ ,  $F=23$  (test set).

Thus, the SMILES-based optimal descriptors can be used as a tool for prediction of the toxicity nanosized oxides. **This is a first try to model the toxicity relationship for the set of metal oxides in nanosized form, applying quantitative structure-activity methods.**

Nanooxide	SMILES	DCW	Expr, logLD50	Calc, LogLD50	Expr-Calc
Training set					
ZnO	O=[Zn]	7.4964167	3.450	3.379	0.071
TiO <sub>2</sub>	O=[Ti]=O	1.5785166	1.742	1.756	-0.014
Fe <sub>2</sub> O <sub>3</sub>	O=[Fe-]O[Fe-]=O	3.9376291	2.405	2.403	0.002
Y <sub>2</sub> O <sub>3</sub>	O=[Y]O[Y]=O	5.9178614	2.868	2.946	-0.078
ZrO <sub>2</sub>	O=[Zr]=O	2.8084706	2.151	2.093	0.058
In <sub>2</sub> O <sub>3</sub>	O=[In]O[In]=O	5.5731328	2.807	2.851	-0.044
Sb <sub>2</sub> O <sub>3</sub>	O=[Sb]O[Sb]=O	4.7980500	2.642	2.639	0.003
Test set					
SnO <sub>2</sub>	O=[Sn]=O	2.5685177	2.006	2.027	-0.021
CuO	[Cu]=O	5.5183250	3.203	2.836	0.367
La <sub>2</sub> O <sub>3</sub>	O=[La]O[La]=O	5.0833134	2.873	2.717	0.156
Al <sub>2</sub> O <sub>3</sub>	O=[Al]O[Al]=O	5.0875610	2.495	2.718	-0.223
Bi <sub>2</sub> O <sub>3</sub>	O=[Bi]O[Bi]=O	5.0827454	2.818	2.717	0.101
SiO <sub>2</sub>	O=[Si]=O	2.5672319	2.199	2.027	0.172
V <sub>2</sub> O <sub>3</sub>	O=[V-]O[V-]=O	5.4808394	3.142	2.826	0.316

# Cytotoxicity of 14 nano sized metal oxides: Experimental testing with bacteria *E. coli* and quantitative structure activity relationships (QSAR) modeling

The following metal oxides in nanosized form which were selected: ZnO, TiO<sub>2</sub>, SnO<sub>2</sub>, La<sub>2</sub>O<sub>3</sub>, Fe<sub>2</sub>O<sub>3</sub>, CuO, Al<sub>2</sub>O<sub>3</sub>, Sb<sub>2</sub>O<sub>3</sub>, V<sub>2</sub>O<sub>3</sub>, Y<sub>2</sub>O<sub>3</sub>, In<sub>2</sub>O<sub>3</sub>, Bi<sub>2</sub>O<sub>3</sub>, SiO<sub>2</sub> and ZrO<sub>2</sub>.

All of these nanosized metal oxides are widely used in many goods and details that present around us. And all of them are quite toxic to some extent, and for many of them the toxicity increasing accordingly to particle size reducing.

In this work we applied quantum-chemical calculations to find parameters that could be responsible for the toxicity properties of these metal oxides. Additionally to quantum-chemical parameters another calculated physico-chemical descriptors have been considered for relationships.

### Three preliminary models obtained:

$$-\text{Log}(\text{LD50}) = -0.7397 \cdot \text{D\_Ox} + 0.0039 \cdot \text{MSA} + 4.5739 \quad (1)$$

$$-\text{Log}(\text{LD50}) = -0.9300 \cdot \text{D\_Ox} + 0.0235 \cdot \text{PSA} + 4.8979 \quad (2)$$

$$-\text{Log}(\text{LD50}) = -0.6768 \cdot \text{D\_Ox} + 0.01369 \cdot \text{Polarizability2} + 4.6230 \quad (3)$$

N=13 (no ZrO<sub>2</sub>)

D\_Ox – oxidation degree

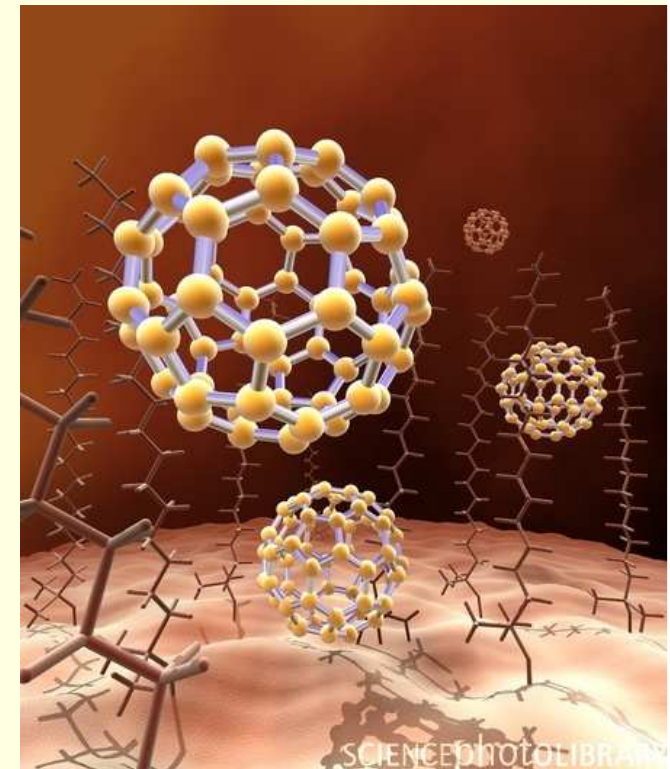
MSA – molar surface area (3D)

PSA – polarized surface area (2D)

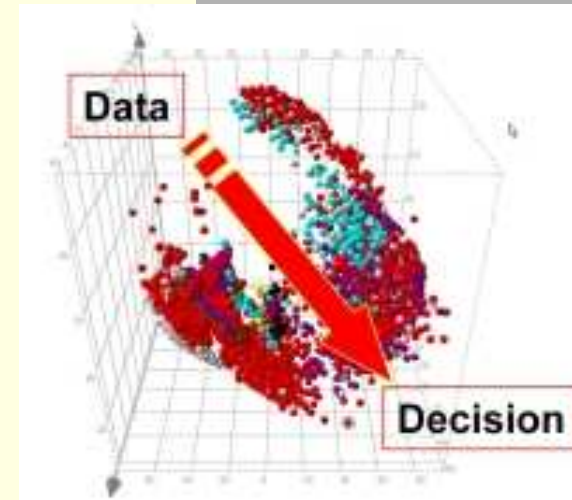
Polarizability2 - polarizability

Statistics	Equation 1	Equation 2	Equation 3
R-squared	0.8314	0.8281	0.8246
Adjusted R-squared	0.7976	0.7937	0.7895
q <sup>2</sup> -Cross validated R <sup>2</sup>	0.7039	0.7559	0.7308
Significant Regression	Yes	Yes	Yes
Significance-of-regression F-value	24.6487	24.0804	23.5000
Critical SOR F-value (95%)	5.3809	5.3809	5.3809

- Overall, the physical properties of nanomaterials can be predicted
- Some of the biological effects (including toxicity) of nanomaterials can be predicted with certain accuracy
- Property predictions for carbon nanomaterials more complicated in comparing to nanooxides, because of limited diversity.
- The good predictions are obtained for fullerene  $C_{60}$  and carbon nanotubes solubility relationships, including direct calculations of  $C_{60}$  solubility and the research in this direction to be continued.



# Acknowledgments



The authors would like to thank:

National Science Foundation for the NSF-CREST grant (2008) to establish Interdisciplinary Center for Nanotoxicity.

The Department of Defense through the U. S. Army Engineer Research and Development Center, Vicksburg, MS, Contract #W912HZ-06-C-0061

## *List of published and submitted papers*

### 2006-2007

1. Andrey A. Toropov, Danuta Leszczynska, Jerzy Leszczynski, QSPR study on solubility of fullerene C<sub>60</sub> in organic solvents using optimal descriptors calculated with SMILES, *Chemical Physics Letters*, 441 (2007) 119–122
2. Andrey Toropov, Bakhtiyor Rasulev, Danuta Leszczynska, Jerzy Leszczynski, Additive SMILES based optimal descriptors: QSPR modeling of fullerene C<sub>60</sub> solubility in organic solvents, *Chemical Physics Letters*, 444 (2007) 209–214
3. Andrey Toropov, Jerzy Leszczynski, A new approach to the characterization of nanomaterials: Predicting Young's modulus by correlation weighting of nanomaterials codes, *Chemical Physics Letters*, 433 (2006) 125–129
4. Andrey Toropov, Danuta Leszczynska, Jerzy Leszczynski, Predicting water solubility and octanol- water partition coefficient for carbon nanotubes based on the chiral vector, *Computational Biology and Chemistry*, 31 (2007) 127–128

### 2008-2009

5. Andrey Toropov, Bakhtiyor Rasulev, Danuta Leszczynska, Jerzy Leszczynski, Multiplicative SMILES Based Optimal Descriptors: QSPR Modeling of Fullerene C<sub>60</sub> Solubility in Organic Solvents, *Chemical Physics Letters*, 457 (2008), 332-336
6. Andrey Toropov, Alla Toropova, Emilio Benfenati, Danuta Leszczynska, , Jerzy Leszczynski, Additive InChI-based optimal descriptors: QSPR modeling of fullerene C<sub>60</sub> solubility in organic solvents, *Journal of Mathematical Chemistry*, 2008 (in press)
7. Tetyana Petrova, Bakhtiyor Rasulev, Andrey Toropov, Danuta Leszczynska, Jerzy Leszczynski, Modeling the fullerene C60 solubility in organic solvents: The use of quantum-chemical and topological descriptors in QSPR study, *Journal of Nanoparticle Research*, 2009, (submitted)
8. Andrey Toropov, Alla Toropova, Bakhtiyor Rasulev, Tomasz Puzyn, Xiaoke Hu, Huey-Min Hwang, Emilio Benfenati, D. Leszczynska, J. Leszczynski, QSAR model of toxicity towards *E.coli* bacteria for nanosized oxides by SMILES-based optimal descriptors, *Chemical Biology & Drug Design*, 2009 (submitted)
9. Tomasz Puzyn, Bakhtiyor Rasulev, Xiaoke Hu, Andrea Michalkova, Andrey Toropov, Huey-Min Hwang, Danuta Leszczynska and Jerzy Leszczynski, Cytotoxicity of 14 nano-sized metal oxides: Experimental testing with bacteria *E. coli* and quantitative structure-activity relationships (QSAR) modeling, 2009 (in preparation)

# THANK YOU

---

More information could be found on our new  
**Interdisciplinary Center for Nanotoxicity**  
funded by the  
**National Science Foundation**

[www.icnanotox.org](http://www.icnanotox.org)

