Research Article

Nanoquantitative Structure-Property Relationship Modeling on C_{42} Fullerene Isomers

Sorana D. Bolboacă¹ and Lorentz Jäntschi^{2,3}

¹Department of Medical Informatics and Biostatistics, Iuliu Hațieganu University of Medicine and Pharmacy,
6 Louis Pasteur Street, 400349 Cluj-Napoca, Romania

²Department of Physics and Chemistry, Technical University of Cluj-Napoca, 103-105 Muncii Bulevardul,
400641 Cluj-Napoca, Romania

³Doctoral Studies-Chemistry, Babeş-Bolyai University, 11 Arany Janos Street, 400028 Cluj-Napoca, Romania

Correspondence should be addressed to Lorentz Jäntschi; lorentz.jantschi@gmail.com

Received 5 October 2015; Revised 1 January 2016; Accepted 4 January 2016

Academic Editor: Teik-Cheng Lim

Copyright © 2016 S. D. Bolboacă and L. Jäntschi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The interest of scientists in nanostructures has been increased in the last years and proper methods for their assessment are needed. *In silico* methods found their usefulness in the replacement of experimental evaluation and are successfully used as efficient alternatives for estimation and prediction of compound's properties or activities. In this paper, it is shown that a Quantitative Structure-Property Relationship method is proper to be applied also on nanostructures. Based on computational experiment, several models to describe the total strain energy of C_{42} fullerene isomers were obtained and their characteristics are presented. Furthermore, the best performing model obtained on C_{42} fullerene isomers was validated on C_{40} fullerene isomers.

This paper is dedicated to Professor Mircea V. Diudea on the occasion of his 65th birthday.

1. Introduction

Since their discovery in 1985 [1], fullerenes attracted interest in different fields of science, including medical field (e.g., for potential use as antibiotics [2–4], as inhibitors of erythroid cells—fullerenol [5], as drug delivery system [6], or as inhibitors of inflammatory mediators [7]). Fullerene molecules are constructed from carbon atoms and take the shape of sphere (also known as buckyballs), ellipsoid, or tube [8]. First spherical fullerene, C_{60} , was discovered in 1985 [1]. Fullerenes have different properties and showed different number of associated isomers (Table 1) [9]. The smallest fullerene (C_{28}) was stabilized by metal encapsulation (with Ti, Zr, and U) by Dunk et al. [10]. Chen et al. showed that C_{32} fullerene has stronger aromaticity compared with C_{30} and C_{34} , respectively [11]. Fifteen distinct isomers with different energies were reported by Manna and Ghanty who encapsulate U into various C_{36} cages [12]. Muhammad et al.

showed that C_{20} is a closed-shell fullerene and fullerenes C_{26} and C_{30} are pure open-shell compounds, whereas C_{36} , C_{40} , and C_{42} are intermediate open-shell compounds [13].

The C_{42} fullerenes are small, not necessary spherical cages. The C_{42} cages enclosed high pentagon/hexagon ratio [14]. Fullerene C_{42} along with C_{60} showed highest values of the main peak on Matrix-Assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF) on mass spectrometric measurement [15].

Some activities of fullerenes have been modeled using quantitative structure-activity relationship (QSAR) approaches (such as anti-HIV protease inhibition activity [16], antiviral activity [17], and drug delivery system [18]). However, C_{60} received the main attention while other fullerenes were neglected in regard of QSAR/QSPR (Quantitative Structure-Property Relationship) modeling. The aim of our research was to model the total strain

TABLE 1: Several small fullerenes and their number of isomers.

Number	Fullerene	Number of isomers
1	C ₂₈	2
2	C ₃₀	3
3	C ₃₂	6
4	C ₃₄	6
5	C ₃₆	15
6	C ₃₈	17
7	C ₄₀	40
8	C ₄₂	45
9	C ₄₄	89
10	C ₄₆	116

Source: <http://www.nanotube.msu.edu/fullerene/fullerene-isomers.html> [accessed June 7, 2015].

SCHEME 1: Flowchart of the applied methods. The pool of filtered SMPI (Szeged Matrix Property Indices) descriptors contains those descriptors with absolute values between 10^{-7} and 10^7 .

energy of the isomers of C₄₂ fullerene using the structural information.

2. Materials and Methods

All C₄₂ fullerene isomers were included in the analysis. Data related to continuum elasticity expressed as total strain energy (TSE in eV) and the structures as *.xyz files of C₄₂ fullerene isomers were taken from [19] (Table 2).

The analysis was conducted on the downloaded file of the C₄₂ isomers without any modification on the available geometry. According to [19], the fullerene geometries were based on the geometry of the structures in Yoshida's Fullerene Library (UNIX files) and reoptimized using Dreiding-like force-field [20]. Here the obtained geometry is used.

The steps applied in the analysis are depicted in Scheme 1.

In the first step of the analysis the downloaded files were translated into *.mol file with Spartan software (<https://www.wavefun.com/products/spartan.html>). In the second step the *.mol file is transformed as *.hin file using

TABLE 2: C₄₂ fullerene isomers: identification number (IsoID) and total strain energy (TSE).

IsoID	TSE (eV)
#01	31.060
#02	30.537
#03	29.791
#04	29.805
#05	30.618
#06	29.850
#07	30.608
#08	29.782
#09	28.527
#10	29.393
#11	29.475
#12	28.340
#13	28.157
#14	27.147
#15	29.955
#16	28.175
#17	28.276
#18	29.474
#19	27.408
#20	28.175
#21	27.283
#22	29.140
#23	28.765
#24	27.743
#25	27.487
#26	28.353
#27	28.014
#28	29.051
#29	27.489
#30	28.972
#31	27.484
#32	26.657
#33	26.639
#34	27.371
#35	26.554
#36	27.973
#37	29.764
#38	31.101
#39	26.639
#40	27.501
#41	26.672
#42	28.665
#43	28.284
#44	26.737
#45	25.661

Babel software (<http://openbabel.org>). The partial charges were calculated in the third step using HyperChem software (<http://www.hyper.com/>) by applying PM3 (Parameterized Model number 3 [21]) single point (energy) semiempirical

TABLE 3: Characteristics of nano-QSPR models obtained on C₄₂ isomers.

Equation	R ²	R ² _{adj}	se	F(p)	t _{min} (p)	%PredErr	Q ²	se _{loo}	F _{loo} (p _{loo})
(1)	0.8883	0.8857	0.4577	342 (4.38 × 10 ⁻²²)	18.05 (1.09 × 10 ⁻²¹)	46.77	0.8656	0.5039	275 (1.56 × 10 ⁻²⁰)
(2)	0.9612	0.9593	0.2729	520 (2.32 × 10 ⁻³⁰)	6.69 (4.10 × 10 ⁻⁸)	31.92	0.9545	0.2960	439 (8.48 × 10 ⁻³⁰)
(3)	0.9836	0.9824	0.1796	820 (1.30 × 10 ⁻³⁶)	4.37 (8.40 × 10 ⁻⁵)	19.76	0.9809	0.1939	701 (3.80 × 10 ⁻³⁷)
(4)	0.9898	0.9888	0.1431	974 (2.87 × 10 ⁻³⁹)	3.55 (1.01 × 10 ⁻³)	15.95	0.9768	0.2171	418 (2.28 × 10 ⁻³⁴)

R²: determination coefficient; R²_{adj}: adjusted determination coefficient; se: standard error of estimate; F(p): Fisher's statistic (p-value); |t_{min}|: the minimum of absolute t-statistic associated with the intercept and coefficients of the model; %PredErr: percentage prediction error; Q²: determination coefficient in leave-one-out analysis; loo: leave-one-out analysis.

calculations. The structural features of the investigated nanoclass of compounds were extracted using unsymmetrical Szeged set, an extension of corresponding Szeged Matrix [22] (forth step). The calculated values of the structural descriptors and the collected values of total strain energy were included in nano-QSPR modeling in the fifth step of the analysis and models with the highest goodness-of-fit (defined as highest correlation coefficients) were analyzed and validated in leave-one-out and leave-many-out analyses [23, 24].

Leave-one-out analysis retrieves valid models if determination coefficient (Q²) takes values higher than 0.5. Leave-many-out analysis was conducted for the models with highest abilities in estimation expressed as the highest value of the correlation coefficient. The set was split using a simple random technique [25] in training and test with 2/3 of compounds in training set. The models obtained in training sets were used to predict the TSE in the test sets. The leave-many-out analysis was run five times for equations identified as being with highest estimation and internal prediction abilities in order to assess their prediction abilities.

The assessment of the prediction ability was done on an external dataset represented by C₄₀ isomers considering the same property. The TSE values and the structures for external validation were taken from the same source as C₄₂ isomers: <http://nanotube.msu.edu/fullerene/fullerene.php?C=40> (accessed December 20, 2015). Several metrics were used to assess the prediction ability of the model [23, 24]: determination coefficient on the external set (R²_{ext}), predictive square correlation coefficient on external set (Q²_{F2}, [26]), external prediction ability (Q²_{F3}), root mean square error of prediction (RMSEP), mean absolute error of prediction (MAEP), percentage predictive error (%PredErr), and concordance correlation coefficient (CCC [27]).

3. Results and Discussion

Structural information of the investigated C₄₂ isomers was obtained by calculation of the pool of descriptors given by Szeged Matrix Property Indices (SMPI) method [28]. Performing models in regard of goodness-of-fit (highest correlation coefficient) with 1, 2, 3, and 4 SMPI descriptors was obtained and is given in (1)–(4):

$$\hat{Y}_{\text{TSE}(1)} = -1176.25 + \text{IJUGE} \times 1.96 \quad (1)$$

$$\hat{Y}_{\text{TSE}(2)} = -542.87 - \text{IIUGF} \times 1.93 \times 10^{-3} + \text{IJUGE} \times 1.81 \quad (2)$$

$$\hat{Y}_{\text{TSE}(3)} = 838.80 - \text{IFEGE} \times 1.41 - \text{IIUGF} \times 3.66 \times 10^{-3} + \text{IJUGE} \times 2.16 \quad (3)$$

$$\hat{Y}_{\text{TSE}(4)} = -199.61 - \text{IFETB} \times 21.63 + \text{IFUGB} \times 40.90 - \text{IIUGF} \times 2.62 \times 10^{-3} + \text{IJUGE} \times 1.56, \quad (4)$$

where $\hat{Y}_{\text{TSE}}$ is total strain energy estimated by the model; IJUGE, IIUGF, IFEGE, IFETB, and IFUGB are SMPI descriptors. Two descriptors (IFETB and IFUGB) account for the atomic number as atomic property; the other two descriptors account for electronegativity (IJUGE and IFEGE), while one accounts for the first ionization energy (IIUGF). The investigated property is related to the geometry of compounds (fourth letter “G” in the name of descriptors) with one exception when it is related to topology (fourth letter “T” in the IFETB descriptor). The other letters reflect the linearization operator (first letter), matrix operation (second letter), and interaction descriptor (third letter).

As expected, the determination coefficient increases as the number of descriptors in the models increases, while the standard error of the estimate decreases (Table 3).

The distance between determination coefficient of the model and determination coefficient obtained in leave-one-out analysis varied from 0.0027 to 0.0227, the smallest distance being obtained by (3) (Table 3). On the other hand, the smallest difference between standard errors (estimation model and leave-one-out model) is obtained by the same model (3).

The analysis of the results presented in Table 3 showed that the model with four descriptors is the one with smallest percentage of prediction error. Furthermore, the data on the scatter closest to the straight line is observed for the model given by (4) (Figure 1). Figure 1 shows the absence of the differences between models from (3) and (4), with the dispersion of the point in the scatter closest to the line for model given by (4).

The main characteristics of the models given by (3) and (4) obtained in leave-many-out analysis (training versus test

TABLE 4: Characteristics of nano-QSPR models in leave-many-out analysis: C₄₂ isomers.

Id	Model				Training		Test			
	Equation (3)	Intercept	IFEGE	IIUGF $\times 10^{-3}$	IJUGE	R^2	F -stat	R^2	F -stat	
1		744.37	−1.27	−3.57	2.09	0.9797	360	0.9877	264	
2		682.4	−1.21	−3.46	2.07	0.9788	369	0.9935	361	
3		902.68	−1.48	−3.79	2.20	0.9794	376	0.9894	358	
4		678.39	−1.22	−3.53	2.12	0.9853	534	0.9851	171	
5		854.05	−1.49	−3.51	2.17	0.9828	458	0.9835	219	
Id	Model				Training		Test			
	Equation (4)	Intercept	IFETB	IFUGB	IIUGF $\times 10^{-3}$	IJUGE	R^2	F -stat	R^2	F -stat
6		−120.59	−19.88	37.19	−2.87	1.55	0.9901	568	0.9637	72
7		−277.84	−20.87	40.14	−2.30	1.54	0.9819	310	0.9814	154
8		−91.46	−20.86	39.01	−3.00	1.56	0.9878	459	0.9636	64
9		−225.97	−21.69	42.02	−2.35	1.48	0.9830	331	0.9794	139
10		−223.28	−18.71	36.37	−2.35	1.47	0.9887	497	0.9701	85

TABLE 5: Prediction power of nano-QSPR given by (3) and (4) on C₄₀ isomers.

Equation	R^2_{ext}	Q^2_{F2}	Q^2_{F3}	RMSEP	MAEP	$ t_{(Y-Y_{\text{pred}})} (p)$	%PredErr
(3)	0.6183	0.9501	NR	1.60	51.28	324 (8.37×10^{-69})	63.19
(4)	0.8462	0.5144	NR	1.60	5.27	52 (4.96×10^{-38})	6.49

R^2_{ext} : determination coefficient on the external set; Q^2_{F2} : predictive square correlation coefficient on external set; Q^2_{F3} : external prediction ability; RMSEP: root mean square error of prediction; MAEP: mean absolute error of prediction; %PredErr: percentage predictive error; NR: not reliable value.

analysis; 2/3 of compounds in training set run 5 times) are presented in Table 4.

The results presented in Table 4 showed the stability of the models, with internal prediction power (defined as determination coefficient in test sets) closed to the estimation power (determination coefficient in training set) from both investigated models. Therefore, the results obtained in training sets closely follow the results on the whole sample for (3) with R^2 in the same range when two decimals are of interest. The R^2 obtained in test set in all five runs of the leave-many-out analysis was equal to 0.99, so slightly higher than the R^2 obtained in training sets (0.98). In three cases out of five, the R^2 in training sets for (4) was in the same range for two decimals with the R^2 value given in Table 3. However, without any exception, the R^2 in test sets was smaller than the R^2 in training sets for (4), with values that varied from 0.0005 (id 7 in Table 4) to 0.0264 (id 6 in Table 4). These results showed that (3) performs slightly better in terms of determination coefficients in leave-many-out analysis.

The plots of the models obtained in the fourth run for (3) and fifth run for (4), as examples, are given in Figure 2.

The equations identified with estimation power and internal prediction abilities, namely, (3) and (4), were further applied on C₄₀ isomers to test the external prediction abilities. The prediction power of (4) proved to be better compared with prediction power of (3) (see Figure 3 and Table 5).

Despite the fact that the predictive square correlation coefficient on external set is higher for (3) compared with the value obtained with (4), all other calculated metrics sustain that the model given by (4) has better prediction abilities (highest determination coefficient on external set, lowest mean absolute error of prediction, and lowest percentage of predictive error; see Table 5). Furthermore, the analysis of the

overall spread of the points in the scatter-plot leads to the conclusion that (4) had better prediction abilities compared with (3). Nevertheless, the mean of residuals proved to be significantly different than the expected value (zero). It could be concluded that the model given by (4) better fit the data on which it was constructed compared with all other models. Nevertheless, *are the structural features extracted by SMPI descriptors on C₄₂ isomers able to predict the TSE on C₄₀ isomers?*

SMPI descriptors used by (3) and, respectively, (4) were used to predict the TSE on C₄₀ isomers. One out the three descriptors from (3) proved to have the slope not significantly different by zero and was not included in further analysis. The identified models obtained on C₄₀ isomers are given in

$$\hat{Y}_{\text{TSE}(5)} = -328.66 - \text{IIUGF} \times 2.43 \times 10^{-3} + \text{IJUGE} \times 1.70$$

$$R^2 = 0.8483,$$

$$R^2_{\text{adj}} = 0.8401,$$

$$\text{se} = 0.65,$$

$$F(p) = 103 (7.07 \times 10^{-16}),$$

$$|t_{\min}|(p) = 2.88 (0.0066),$$

$$n = 40,$$

$$Q^2_{F3} = 0.7834,$$

$$\text{RMSEP} = 1.60,$$

$$\text{MAEP} = 0.52,$$

FIGURE 1: Observed versus estimated TSE by (1)–(4).

FIGURE 2: Internal prediction versus estimation power in training and test analysis for (3) and (4).

FIGURE 3: Analysis of (3) and (4) on external dataset represented by C_{40} isomers.

$$\%PredErr = 0.64,$$

$$CCC = 0.9179$$

(5)

$$\begin{aligned} \hat{Y}_{TSE(6)} = & -IFETB \times 15.05 + IFUGB \times 31.49 \\ & - IIUGF \times 2.64 \times 10^{-3} + IJUGE \\ & \times 1.21 \end{aligned}$$

$$R^2 = 0.8853,$$

$$R^2_{adj} = 0.8479,$$

$$se = 0.57,$$

$$F(p) = 69 (3.76 \times 10^{-16}),$$

(6)

$$|t_{min}|(p) = 3.10 (0.0038),$$

$$n = 40,$$

$$Q^2_{F3} = 0.8362,$$

$$RMSEP = 1.60,$$

$$MAEP = 0.43,$$

$$\%PredErr = 0.52,$$

$$CCC = 0.9390,$$

where $\hat{Y}_{TSE}$ is total strain energy estimated by the model; IJUGE, IIUGF, IFETB, and IFUGB are SMPI descriptors. Two descriptors (IFETB and IFUGB) account for the atomic number as atomic property, one descriptor accounts for electronegativity (IJUGE), and one accounts for the first ionization energy (IIUGF). The investigated property is related to the geometry of compounds (fourth letter “G” in

the name of descriptors) with one exception that is related with compounds topology (IFETB descriptor). The other letters reflect the linearization operator (first letter), matrix operation (second letter), and interaction descriptor (third letter). Note that both models have the mean of residual not significantly different by zero ($p > 0.49$).

The analysis of the metrics associated with (5) and (6) leads to the conclusion that model given by (6) perform better than the model given by (5). The same conclusion is obtained by analyzing the plots of observed versus predicted TSE (Figure 4).

The results of our study showed that the identified nano-QSPR models fit the data based on which the model was identified (C_{42} isomers) but could be used for selection of those structural descriptors with fair abilities in prediction on external dataset (C_{40} isomers). To sum up, equations relating electronegativities, ionization potential, and energy have been identified on C_{42} isomers and proved to work also on C_{40} isomers. Note that electronegativities and ionization potential are atomic properties and since the investigated set contains just C and H atoms, the identified relation between the three properties could be assigned also to the topology and geometry of the investigated compounds.

To the best of our knowledge, structure-property relationship approaches were not applied on C_{42} or C_{40} fullerene isomers. The small-diameter fullerenes (C_{20} , C_{34} , C_{42} , and C_{60}) were mainly investigated in regard of properties (such as adsorption [29], distribution of CC distance [14], and Schlegel diagrams of molecular structures [30]). Therefore, this is the first report of a quantitative relationship between structure and property of C_{42} fullerene. Undoubtedly, the advancement from theoretical to experimental studies is desired.

4. Conclusions

The C_{42} fullerene isomers were successfully modeled and the total strain energy was characterized as function of information extracted from structure of the compounds. The models

FIGURE 4: Analysis of (5) and (6) on external dataset represented by C₄₀ isomers.

with goodness-of-fit in leave-one-out ($Q^2 = 0.9768$) and leave-many-out analyses proved also that prediction power is the one with four descriptors. The total strain reaction proved a function of electronegativity and first ionization energy, in relation to geometry of compounds. The structural descriptors able to fairly explain the total strain energy on C₄₂ isomers proved also able to explain the same property on C₄₀ fullerene isomers.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley, "C₆₀: Buckminsterfullerene," *Nature*, vol. 318, no. 6042, pp. 162–163, 1985.
- [2] R. Dinesh, M. Anandaraj, V. Srinivasan, and S. Hamza, "Engineered nanoparticles in the soil and their potential implications to microbial activity," *Geoderma*, vol. 173–174, pp. 19–27, 2012.
- [3] A. J. Huh and Y. J. Kwon, "Nanoantibiotics: a new paradigm for treating infectious diseases using nanomaterials in the antibiotics resistant era," *Journal of Controlled Release*, vol. 156, no. 2, pp. 128–145, 2011.
- [4] Y. S. Zhang, T. H. Dai, M. Wang, D. Vecchio, L. Y. Chiang, and M. R. Hamblin, "Potentiation of antimicrobial photodynamic inactivation mediated by a cationic fullerene by added iodide: in vitro and in vivo studies," *Nanomedicine*, vol. 10, no. 4, pp. 603–614, 2015.
- [5] N. V. Tishevskaya, Yu. M. Zakharov, E. V. Golubotovskii et al., "Effects of fullereneol C₆₀(OH)₂₄ on erythropoiesis *in vitro*," *Bulletin of Experimental Biology and Medicine*, vol. 157, no. 1, pp. 49–51, 2014.
- [6] S. Pacor, A. Grillo, L. Đorđević et al., "Effects of two fullerene derivatives on monocytes and macrophages," *BioMed Research International*, vol. 2015, Article ID 915130, 13 pages, 2015.
- [7] A. L. Dellinger, Z. Zhou, and C. L. Kepley, "A steroid-mimicking nanomaterial that mediates inhibition of human lung mast cell responses," *Nanomedicine: Nanotechnology, Biology, and Medicine*, vol. 10, no. 6, pp. 1185–1193, 2014.
- [8] A. Hirsch and M. Brettreich, *Fullerenes: Chemistry and Reactions*, John Wiley & Sons, New York, NY, USA, 2005.
- [9] D. Tománek, *Guide Through the Nanocarbon Jungle*, Morgan & Claypool, San Rafael, Calif, USA, 2014.
- [10] P. W. Dunk, N. K. Kaiser, M. Mulet-Gas et al., "The smallest stable fullerene, M@C₂₈ (M = Ti, Zr, U): stabilization and growth from carbon vapor," *Journal of the American Chemical Society*, vol. 134, no. 22, pp. 9380–9389, 2012.
- [11] Y.-M. Chen, J. Shi, L. Rui, and Q.-X. Guo, "Theoretical study on C₃₂ fullerenes and their endohedral complexes with noble gas atoms," *Journal of Molecular Structure: THEOCHEM*, vol. 907, no. 1–3, pp. 104–108, 2009.
- [12] D. Manna and T. K. Ghanty, "Enhancement in the stability of 36-atom fullerene through encapsulation of a uranium atom," *Journal of Physical Chemistry C*, vol. 117, no. 34, pp. 17859–17869, 2013.
- [13] S. Muhammad, K. Fukuda, T. Minami, R. Kishi, Y. Shigeta, and M. Nakano, "Interplay between the diradical character and third-order nonlinear optical properties in fullerene systems," *Chemistry—A European Journal*, vol. 19, no. 5, pp. 1677–1685, 2013.
- [14] E. Małolepsza, Y.-P. Lee, H. A. Witek, S. Irle, C.-F. Lin, and H.-M. Hsieh, "Comparison of geometric, electronic, and vibrational properties for all pentagon/hexagon-bearing isomers of fullerenes C₃₈, C₄₀, and C₄₂," *International Journal of Quantum Chemistry*, vol. 109, no. 9, pp. 1999–2011, 2009.
- [15] E. I. Kauppine, "Carbon Nanotubes and NanoBuds—Synthesis, Structure, Functionalisation and Dry Deposition for TCE and TFT Applications," July 2015, <http://www.jst.go.jp/sicp/ws2009-finland/abstract/wg2.02kau.pdf>.
- [16] M. Ibrahim, N. A. Saleh, W. M. Elshemey, and A. A. Elsayed, "Fullerene derivative as anti-HIV protease inhibitor: molecular modeling and QSAR approaches," *Mini-Reviews in Medicinal Chemistry*, vol. 12, no. 6, pp. 447–451, 2012.

- [17] L. Ahmed, B. Rasulev, M. Turabekova, D. Leszczynska, and J. Leszczynski, "Receptor- and ligand-based study of fullerene analogues: comprehensive computational approach including quantum-chemical, QSAR and molecular docking simulations," *Organic & Biomolecular Chemistry*, vol. 11, no. 35, pp. 5798–5808, 2013.
- [18] A. Trpkovic, B. Todorovic-Markovic, and V. Trajkovic, "Toxicity of pristine versus functionalized fullerenes: mechanisms of cell damage and the role of oxidative stress," *Archives of Toxicology*, vol. 86, no. 12, pp. 1809–1827, 2012.
- [19] D. Tománek, C42 Isomers. In: Guide through the Nanocarbon Jungle: Buckyballs, Nanotubes, Graphene, and Beyond, 2015, <http://www.nanotube.msu.edu/fullerene/fullerene.php?C=42>.
- [20] S. L. Mayo, B. D. Olafson, and W. A. Goddard, "DREIDING: a generic force field for molecular simulations," *Journal of Physical Chemistry*, vol. 94, no. 26, pp. 8897–8909, 1990.
- [21] J. J. P. Stewart, "PM3," in *Encyclopedia of Computational Chemistry*, P. von and R. Schleyer, Eds., John Wiley & Sons, New York, NY, USA, 1998.
- [22] M. V. Diudea, O. M. Minailiuc, G. Katona, and I. Gutman, "Szeged matrices and related numbers," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 35, pp. 129–143, 1997.
- [23] S. D. Bolboacă and L. Jäntschi, "Quantitative structure-activity relationships: linear regression modelling and validation strategies by example," *BIOMATH*, vol. 2, no. 1, Article ID 1309089, 11 pages, 2013.
- [24] S. D. Bolboacă, L. Jäntschi, and M. V. Diudea, "Molecular design and QSARs/QSPRs with molecular descriptors family," *Current Computer-Aided Drug Design*, vol. 9, no. 2, pp. 195–205, 2013.
- [25] S. D. Bolboacă, "Assessment of random assignment in training and test sets using generalized cluster analysis technique," *Applied Medical Informatics*, vol. 28, no. 2, pp. 9–14, 2010.
- [26] N. Chirico and P. Gramatica, "Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection," *Journal of Chemical Information and Modeling*, vol. 52, no. 8, pp. 2044–2058, 2012.
- [27] Lin's Concordance, December 2015, <http://services.niwa.co.nz/services/statistical/concordance>.
- [28] L. Jäntschi, "Szeged Matrix Property Indices," 2014, <http://l.academicdirect.org/Chemistry/SARs/SMPI>.
- [29] X. Liu, Y. Wen, Z. Chen et al., "Modulation of Dirac points and band-gaps in graphene via periodic fullerene adsorption," *AIP Advances*, vol. 3, no. 5, Article ID 052126, 2013.
- [30] Y.-N. Chiu, J. Xiao, C. D. Merritt et al., "Special geminals and Schlegel diagrams of molecular structures of fullerenes and metallofullerenes," *Journal of Molecular Structure: THEOCHEM*, vol. 530, no. 1-2, pp. 67–83, 2000.