

This article was downloaded by: [Daniel Bacelo]

On: 02 August 2015, At: 06:37

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London, SW1P 1WG



SAR and QSAR in Environmental Research

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/qsar20>

QSPR study on refractive indices of solvents commonly used in polymer chemistry using flexible molecular descriptors

S.E. Fioressi^a, D.E. Bacelo^a, W.P. Cui^b, L.M. Saavedra^c & P.R. Duchowicz^c

^a Departamento de Química, Facultad de Ciencias Exactas y Naturales, Universidad de Belgrano, Buenos Aires, Argentina

^b Qiannan Normal College Nationalities, Duyun, China

^c Instituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas INIFTA (CCT La Plata-CONICET, UNLP), La Plata, Argentina
Published online: 30 Jul 2015.



CrossMark

[Click for updates](#)

To cite this article: S.E. Fioressi, D.E. Bacelo, W.P. Cui, L.M. Saavedra & P.R. Duchowicz (2015): QSPR study on refractive indices of solvents commonly used in polymer chemistry using flexible molecular descriptors, SAR and QSAR in Environmental Research, DOI: [10.1080/1062936X.2015.1064472](https://doi.org/10.1080/1062936X.2015.1064472)

To link to this article: <http://dx.doi.org/10.1080/1062936X.2015.1064472>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

QSPR study on refractive indices of solvents commonly used in polymer chemistry using flexible molecular descriptors

S.E. Fioressi^{a*}, D.E. Baceo^a, W.P. Cui^b, L.M. Saavedra^c and P.R. Duchowicz^c

^aDepartamento de Química, Facultad de Ciencias Exactas y Naturales, Universidad de Belgrano, Buenos Aires, Argentina; ^bQiannan Normal College Nationalities, Duyun, China; ^cInstituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas INIFTA (CCT La Plata-CONICET, UNLP), La Plata, Argentina

(Received 4 April 2015; in final form 17 June 2015)

A predictive Quantitative Structure–Property Relationship (QSPR) for the refractive indices of 370 solvents commonly used in the processing and analysis of polymers is presented, using as chemical information descriptors the simplified molecular input line entry system (SMILES). The model employs a flexible molecular descriptor and a conformation-independent approach. Various well-known techniques, such as the use of an external test set of compounds, the cross-validation method, and Y-randomization were used to test and validate the established equations. The predicted values were finally compared with published results from the literature. The simple model proposed correlates the refractive index values with good accuracy, and it is not dependent on 3D-molecular geometries.

Keywords: QSPR theory; polymer; refractive index; graph theory; CORAL method

1. Introduction

It is well known that the solvent plays an important role in polymer synthesis, processing, and characterization. The structure and dynamics, as well as the physical and optical attributes, of polymers are partly determined by the solvent [1]. Some important properties of a polymer, such as average molecular weight and refractive index, needed to evaluate its potential usefulness in applications such as manufacture of waveguides, solar cells, semiconductors, optical and packaging materials, can be modulated by changing the solvent used during processing [2–5]. Light-scattering experiments in dilute solutions are often used to determine average molecular weights of polymers [6]. The method is based on the fact that a polymer dissolved in a solvent is an optical inhomogeneity that scatters light, and the scattering depends on the solution concentration and the refractive index of both solvent and polymer. The rate of change of the refractive index (n) of a dilute solution for different concentrations of the polymer (c) is key in the interpretation of light-scattering experiments. For such dilute solutions, the specific refractive index increment ($\partial n/\partial c$) is a useful constant whose value only depends on the solution components and temperature. The specific refractive index increment is proportional to the difference between the refractive indices of the polymer and the solvent. Therefore, the accuracy of the method can be improved if the refractive index increment between polymer and solvent is increased.

*Corresponding author. Email: sfioressi@yahoo.com

The measurement of the refractive index of a polymer also involves a solution. The n of organic polymers ranging from 1.3 to 1.7 can be measured with an accuracy of four decimal places at 23°C through ASTM D 542 - 00 and ISO 489:1999 standard methods, using a V-prism refractometer. A contact liquid is required to produce a planar contact between the sample and the instrument's prism. Because alterations in n due to chemical interaction between the solvent and the polymer must be avoided, the number of solvents that can be used is limited. Moreover, the solvent's n must be higher but not below one unit to the second decimal place when compared with the index of the polymer being measured.

It is clear that the refractive index of solvents used in the processing and analysis of polymers should be carefully considered when evaluating if a solvent is suitable for a given purpose. Moreover, the possibility of having a simple theoretical methodology to quickly and accurately predict the refractive indices of commonly used solvents, which constitute a large set of very diverse compounds, can have uses in many others areas. During recent decades, the application of Quantitative Structure–Property Relationship (QSPR) theory [7–9] has proved to provide excellent prediction of properties in a fast and inexpensive way.

In QSPR theory, the property of a chemical compound is determined solely by its molecular structure, which is quantified through a set of adequate molecular descriptors. The descriptors are numbers containing specific information on the constitutional, topological, geometrical, hydrophobic, and/or electronic attributes of the chemical structure [10–12]. The statistical correlation of an experimental property with a set of descriptors results in a model which can be used to discover practical relationships and trends. Within the QSPR framework there are several factors that are crucial; among them, the most important are: (a) the composition of the training and test sets; (b) the choice of representative molecular descriptors with low collinearities between them; (c) the amount of descriptors included in the model; (d) the use of suitable modelling methods; and (e) the employment of validation techniques to verify the predictive performance of the developed models [13–16].

Although many QSPR models have been developed for homologous series of compounds [17–19], general QSPR studies modelling the refractive index of an unrelated set of organic solvents with their chemical structure are scarce. Katritzky et al. [20] developed a five-parameter correlation equation for 125 different organic compounds using the Comprehensive Descriptors for Structural and Statistical Analysis (CODESSA) software. Using quantum chemical, topological, and constitutional descriptors, a good correlation with $r^2 = 0.945$ and $r = 0.0155$ was obtained. Models involving quantum-chemical descriptors usually imply a relatively difficult calculation of the optimum molecular geometry, involving high computational costs and much time. In this context, conformation-independent 0D, 1D and 2D-QSPR methods emerge as an alternative approach for developing models based on the constitutional and topological molecular features of compounds [21,22]. The exclusion of 3D structural aspects also avoids problems associated with ambiguities resulting from an incorrect computational geometry optimization due to the existence of compounds in various conformational states. These kinds of problems may also lead to the loss of predictive capability of the QSPR when applied to the prediction of an external test set of compounds. In this work, we propose a flexible descriptor-based QSPR model [23] for the prediction of refractive index values, in a molecular set composed of 370 organic solvents used for the synthesis and characterization of polymers. In the realms of the approach used, the calculated flexible descriptor is a molecular descriptor which depends both on the molecular structure and the property under analysis (n), but does not explicitly depend on the 3D-molecular geometry.

2. Materials and methods

2.1 The flexible molecular descriptor definition

The high-quality experimental refractive indices measured at 298 K on 370 common solvents were collected from a published compilation [24]. The n values range in the interval [1.280, 1.750], and the complete list of solvents studied here are included in Table S1 as Supplementary Material (available via the Supplementary Content tab on the article's online page at <http://dx.doi.org/10.1080/1062936X.2015.1064472>). Several kinds of flexible molecular descriptors can be readily calculated with CORAL freeware for Windows [25]. This software has been successfully applied previously in several QSPR studies, and also for Quantitative Structure–Activity Relationship (QSAR) analyses [18].

Solvents were represented with Simplified Molecular Input Line Entry System (SMILES) notation, the chemical format used by CORAL. Three different structural representation (SR) approaches are available in the CORAL program: (i) a chemical graph, such as a hydrogen-suppressed graph (HSG), hydrogen-filled graph (HFG) and graph of atomic orbitals (GAO); (ii) SMILES; and (iii) a hybrid representation which includes both graph and SMILES [23]. The most appropriate combination of structural attributes should be chosen for a modelling process, because the selected SR defines the number and types of local descriptors that participate in the QSPR analysis.

In the graph approach, the Morgan's extended connectivity indices of k^{th} order for vertex (atom) number j (kEC_j , $k = 0-3$) can be used as structural attributes. It should be noted that the index of zero-th order 0EC_j represents the vertex degree for j (number of neighbour atoms to j), while the higher order indices kEC_j are obtained through a recursive formula based on 0EC_j (see in Table S2, available online).

In the SMILES approach, the one-, two-, and three-elements SMILES attributes ${}^1s_k, {}^2s_k, {}^3s_k$, respectively, can be calculated. If a SMILES is a sequence of elements such as 'ABCDE', then such structural attributes can be represented with Equations (1)–(3):

$$\text{'ABCDE'} \rightarrow \text{'A'}, \text{'B'}, \text{'C'}, \text{'D'}, \text{'E'} ({}^1s_k) \quad (1)$$

$$\text{'ABCDE'} \rightarrow \text{'AB'}, \text{'BC'}, \text{'CD'}, \text{'DA'} ({}^2s_k) \quad (2)$$

$$\text{'ABCDE'} \rightarrow \text{'ABC'}, \text{'BCD'}, \text{'CDE'} ({}^3s_k) \quad (3)$$

In addition, the *NOSP* attribute represents an index calculated according to the presence or absence of the chemical elements nitrogen, oxygen, sulphur, and phosphorus; the *HALO* attribute represents the presence or absence of fluorine, chlorine, and bromine; *BOND* is a mathematical function representing the presence or absence of double (=), triple (#), or stereo chemical bonds (@ or @@), and the *ATOMPAIR* is a mathematical function for the presence of the seven chemical elements: F, Cl, Br, N, O, S, and P.

Within the CORAL framework, a QSPR model is obtained through a one-variable linear correlation between n and a properly defined flexible descriptor (*DCW*). The *DCW* descriptor is a linear combination of special coefficients, the so-called correlation weights (*CW*). A *CW* value is calculated for each type of structural attribute of the training set. The method for obtaining the *CW* values for all the structural attributes is based on the Monte Carlo (MC) simulation method, by searching for the highest correlation coefficient (r) between n and the *DCW* descriptor.

The *DCW* flexible descriptor depends upon the threshold value (T) and the number of epochs or iterations (N_{epochs}) used. These parameters are positive integers from the MC

method that should be correctly specified in order to calculate the *DCW* values. The T defines rare (noise) SMILES attributes that do not contribute to the predicted n , so that all SMILES attributes that take place in less than T SMILES notations of the training set are classified as rare instead of as active. N_{epochs} is the number of iterations used during the numerical optimization procedure [26]. In current study, T ranges from 0 to 5 and the maximum number of iterations used is 50.

2.2 Model validation

The validation of the QSPR consists of testing its ability to predict the property for molecular structures not considered during the model development. The theoretical validation of the linear regression models is based on the popular validation criteria based on Cross Validation using Leave-One-Out (loo) and Leave-More-Out (ln%, with n% being the percentile of molecules removed from the training set). The statistical parameters $r_{\text{ln}\%}$ and $s_{\text{ln}\%}$ (correlation coefficient and standard deviation of Leave-More-Out) measure the stability of the QSPR upon inclusion/exclusion of molecules. The number of cases for random data removal analysed in this study is 100,000. According to the specialized literature, the loo-explained variance (r_{loo}^2) should be greater than 0.5 for a validated model, although this is a necessary but not sufficient condition for its predictive power [13].

A more reliable validation is applied that consists of using an external test set of structures. The 370 solvents are ranked according to their n values, and every alternate compound is assigned to the training set (train), validation set (val) and test set (test). Each set thus includes 124, 123, and 123 compounds, respectively.

We use Y-Randomization [27] as a way of checking that the model does not result from happenstance and to avoid the development of fortuitous correlations. This technique consists of scrambling the experimental property values in such a way that they do not correspond to the respective compounds. After analysing 10,000 cases of Y-Randomization, the smallest standard deviation value obtained using this procedure (s^{rand}) has to be a higher (poorer) value than the one found by considering the true calibration (s).

3. Results and discussion

The QSPR analysis was performed by searching the best linear regression models on the training set of 370 compounds; in order to decide which structural attributes are the most efficient for each SR during the flexible descriptor design, the *DCW* flexible descriptor was optimized by increasing r_{train}^2 , until the model started to lose predictive capability in the validation and test sets. This is the same situation that appears when the most predictive model must be selected among several multivariable linear regressions, descriptors having being searched in a pool containing thousands of them [28]. Table 1 contains a summary of the statistical quality of the best QSPR models found by trying different possible CORAL methods. It reveals that the best choice of SR for each solvent structure is a hybrid approach that includes both graph and SMILES representations.

The statistics for the stepwise evolution of the solvent model are presented in Table 1, where the first selected structural attribute is 1S_k , then the following ones are 1EC_j , and 2S_k in that order. The common practice of keeping the model's size as small as possible (Ockham's razor) was followed; in order to avoid any fortuitous correlation, no more than three attributes in the *DCW* calculation were considered because no further improvement can be obtained

Table 1. The stepwise search for the best QSPR model in the hybrid approach as structure representation.

Structural attributes	r^2_{train}	s_{train}	r^2_{val}	s_{val}	r^2_{test}	s_{test}
1S_k	0.68	0.035	0.75	0.028	0.77	0.028
$^1EC_j, ^2S_k$	0.93	0.017	0.93	0.015	0.83	0.023
$^1EC_j, ^1S_k, ^2S_k$	0.93	0.017	0.94	0.014	0.87	0.021

beyond that value [29]. The model with two attributes was chosen due to its good performance and simplicity. More complete details for the QSPR model established are as follows:

$$n = 1.3611 (\pm 0.0002995) + 0.0120 (\pm 0.0000352) * \text{DCW} \quad (4)$$

$$N_{\text{train}} = 124, r^2_{\text{train}} = 0.93, s_{\text{train}} = 0.017, F = 1554$$

$$N_{\text{val}} = 123, r^2_{\text{val}} = 0.93, s_{\text{val}} = 0.015$$

$$N_{\text{test}} = 123, r^2_{\text{test}} = 0.83, s_{\text{test}} = 0.023, o(3s) = 4$$

Here, F is the Fisher parameter and $o(3s)$ [30] indicate the number of outlier compounds having a residual (difference between experimental and calculated n) greater than three times s_{train} and lower than three times s_{train} . Equation (4) (see Figure 1) presents only one outlier compound in the training set that has a residual higher than three standard deviations, compound **1**. We assume that this abnormal behaviour results from the fact that this compound has the lowest refractive index value in the set (extreme sample). The compounds in the validation and test sets that exceed the $o(3s)$ have similar residuals to compound **1**. After an exhaustive analysis of these compounds, we are confident that their molecular structures and

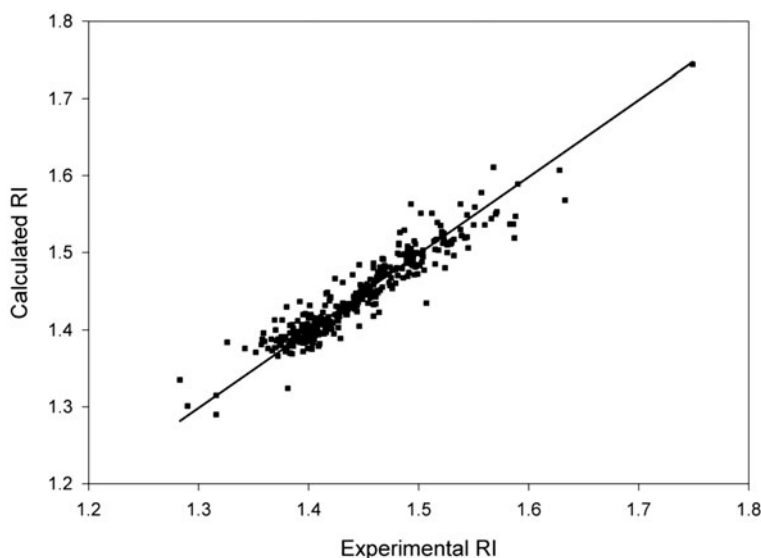


Figure 1. Calculated versus experimental refractive indices for Equation (4) ($N = 370$).

experimental refractive values are correct. Hence, we can assume that this irregular behaviour may be attributed to the wide structural diversity of the 370 molecules considered in this study. On the other hand, the dispersion plot of the residuals, presented in Figure 2, shows that the points follow a random pattern around the zero line in such a way that a linear regression model is fulfilled.

Table S3 (on the article's online page) includes the predicted refraction index values and the *DCW* descriptor, which considers mixed graph and SMILES approaches for the SR. We apply Y-randomization, demonstrating that $s_{\text{train}} < s^{\text{rand}}$ and thus a valid structure–property relationship is achieved. Equation (4) also satisfies the external validation conditions reported in [31] (Table S4, available online):

- $1 - r_0^2/r_{\text{test}}^2 < 0.1$ (0.02) and $1 - r_0'^2/r_{\text{test}}^2 < 0.1$ (0.00) and,
- $0.85 \leq k \leq 1.15$ (1.00) and $0.85 \leq k' \leq 1.15$ (1.00)
- $r_m^2 > 0.5$ (0.76)

The parameters used during model building were $T = 5$ and $N_{\text{epochs}} = 50$. The correlation weights produced by the MC simulation appear listed in Table S5, while Table S6 includes an example for calculating the *DCW* for **1**. Figure 1 shows that the predicted versus experimental refractive index obtained with Equation (4) follows a straight line.

The flexible descriptor of Equation (4) considers 1EC_j as local graph invariant, and the structural attributes that contribute to the *DCW* calculation are listed in Table S5. Furthermore, higher positive *CW* values tend to predict higher n values. Figure 2 shows the dispersion plot of residuals for the solvents studied. The QSPR given by Equation (4) predicts the refractive index of 370 structurally diverse solvents with a good accuracy, and compares favourably with previous published results. For instance, Equation (4) analyses a higher number of solvent compounds than a previous reported study [15], and only involves a single descriptor as compared with the five descriptors used in that research. Furthermore, the developed model was properly internally and externally validated.

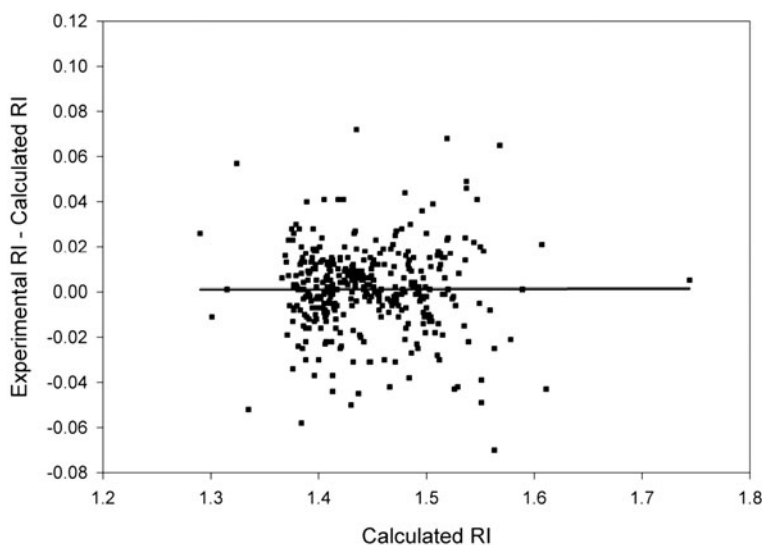


Figure 2. Dispersion plot of residuals for 370 solvents.

4. Conclusions

The refractive index is considered a fundamental physical property in the selection of an adequate solvent for processing and characterization of a polymer. In this work, a solvent structure model was proposed and successfully validated. The refractive index values correlate with good accuracy, demonstrating that such a model is predictive in the validation process. The calculated flexible descriptor does not require structural information on the molecular conformation of the solvents studied, which means that the method is able to model the physical property by representing the molecular structure aspects with a similar or better degree of detail as when using a 3D-geometry-dependent approach [15]. The procedure employed here can be readily applied to other solvent properties, which will be investigated in the near future.

Supplementary material

The Supplementary Material for this paper is available at <http://dx.doi.org/10.1080/1062936X.2015.1064472>.

Acknowledgements

We thank the financial support provided by the National Research Council of Argentina (CONICET) PIP11220100100151 project and to Ministerio de Ciencia, Tecnología e Innovación Productiva for the electronic library facilities. WC acknowledges the support of the AGMUS-QNCN Sunday Research Academy.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- [1] C. Scharsich, R.H. Lohwasser, M. Sommer, U. Asawapirom, U. Scherf, M. Thelakkat, D. Nieher, and A. Köhler, *Control of aggregate formation in poly (3-hexylthiophene) by solvent, molecular weight, and synthetic method*, J. Polym. Sci., Part B: Polym. Phys. 50 (2012), pp. 442–453.
- [2] H. Ma, A.K.Y. Jen, and L.R. Dalton, *Polymer-based optical waveguides: Materials, processing, and devices*, Adv. Mater. 14 (2002), pp. 1339–1365.
- [3] H. Zhou, L. Yang, and W. You, *Rational design of high performance conjugated polymers for organic solar cells*, Macromolecules 45 (2012), pp. 607–632.
- [4] P.D. Cunningham, N.N. Valdes, F.A. Vallejo, L.M. Hayden, B. Polishak, X.H. Zhou, and R.J. Twieg, *Broadband terahertz characterization of the refractive index and absorption of some important polymeric and organic electro-optic materials*, J. Appl. Phys. 109 (2011), pp. 043505–043505.
- [5] Z.B. Henson, K. Müllen, and G.C. Bazan, *Design strategies for organic semiconductors beyond the molecular formula*, Nat. Chem. 4 (2012), pp. 699–704.
- [6] D.W. van Krevelen and K. te Nijenhuis, *Properties of Polymers: Their Correlation with Chemical Structure; Their Numerical Estimation and Prediction from Additive Group Contributions*, Elsevier, New York, 2009.
- [7] A.R. Katritzky, V.S. Lobanov, and M. Karelson, *QSPR: The correlation and quantitative prediction of chemical and physical properties from structure*. Chem. Soc. Rev. 24 (1995), pp. 279–287.
- [8] A.R. Katritzky, M. Karelson, and V.S. Lobanov, *QSPR as a means of predicting and understanding chemical and physical properties in terms of structure*, Pure Appl. Chem. 69 (1997), pp. 245–248.
- [9] R. Kunal, S. Kar, and R. Narayan Das, *A Primer on QSAR/QSPR Modeling: Fundamental Concepts*, Springer, Berlin, 2015.

- [10] A.R. Katritzky and E.V. Goordeva, *Traditional topological indices vs. electronic, geometrical, and combined molecular descriptors in QSAR/QSPR research*, J. Chem. Inf. Comput. Sci. 33 (1993), pp. 835–857.
- [11] M.V.E. Diudea, *QSPR/QSAR Studies by Molecular Descriptors*, Nova Science Publishers, New York, 2001.
- [12] R. Todeschini and V. Consonni, *Molecular Descriptors for Chemoinformatics (Methods and Principles in Medicinal Chemistry)*, Wiley-VCH, Weinheim, 2009.
- [13] A. Golbraikh and A. Tropsha, *Beware of q^2 !*, J. Mol. Graphics Modell. 20 (2002), pp. 269–276.
- [14] D.M. Hawkins, S.C. Basak, and D. Mills, *Assessing model fit by cross validation*, J. Chem. Inf. Model. 43 (2003), pp. 579–586.
- [15] M.V. Putz, A.-M. Putz, M. Lazea, L. Ienciu, and A. Chiriac, *Quantum-SAR extension of the spectral-SAR algorithm. Application to polyphenolic anticancer bioactivity*, Int. J. Mol. Sci. 10 (2009), pp. 1193–1214.
- [16] M.V. Putz, C. Ionascu, A.M. Putz, and V. Ostafe, *Alert-QSAR. Implications for electrophilic theory of chemical carcinogenesis*, Int. J. Mol. Sci. 12 (2011), pp. 5098–5134.
- [17] Z. Ha, Z. Ring, and S. Liu, *Quantitative structure–property relationship (QSPR) models for boiling points, specific gravities, and refraction indices of hydrocarbons*, Energy Fuels 19 (2005), pp. 152–163.
- [18] W. Liu, P. Yi, and Z. Tang, *QSPR models for various properties of polymethacrylates based on quantum chemical descriptors*, QSAR Comb. Sci. 25 (2006), pp. 936–943.
- [19] A.A. Toropov and A.P. Toropova, *QSPR modeling of alkanes properties based on graph of atomic orbitals*, J. Mol. Struct.: THEOCHEM 637 (2003), pp. 1–10.
- [20] A.R. Katritzky, S. Sild, and M. Karelson, *General quantitative structure–property relationship treatment of the refractive index of organic compounds*, J. Chem. Inf. Comput. Sci. 38 (1998), pp. 840–844.
- [21] P.R. Duchowicz, N.C. Comelli, E.V. Ortiz, and E.A. Castro, *QSAR study for carcinogenicity in a large set of organic compounds*, Curr. Drug Saf. 7 (2012), pp. 282–288.
- [22] A. Talevi, C.L. Bellera, M.D. Ianni, P.R. Duchowicz, L.E. Bruno-Blanch, and E.A. Castro, *An integrated drug development approach applying topological descriptors*, Curr. Comput. Aided Drug Des. 8 (2012), pp. 172–181.
- [23] A.A. Toropov, A.P. Toropova, E. Benfenati, and G. Gini, *OCWLGI descriptors: Theory and praxis*, Curr. Comput. Aided Drug Des. 9 (2013), pp. 226–232.
- [24] J. Brandrup, E.H. Immergut, and E.A. Grulke (eds.), *Polymer Handbook*, 4th ed., Wiley-Interscience, New York, 1999.
- [25] Coral, <http://www.insilico.eu/coral>.
- [26] A.P. Toropova, A.A. Toropov, S.E. Martyanov, E. Benfenati, G. Gini, D. Leszczynska, and J. Leszczynski, *CORAL: QSAR modeling of toxicity of organic chemicals towards Daphnia magna*, Chemometr. Intell. Lab. Syst. 110 (2012), pp. 177–181.
- [27] C. Rücker, G. Rücker, and M. Meringer, *Y-Randomization and its variants in QSPR/QSAR*, J. Chem. Inf. Model. 47 (2007), pp. 2345–2357.
- [28] A.G. Mercader, P.R. Duchowicz, F.M. Fernández, and E.A. Castro, *Advances in the replacement and enhanced replacement method in QSAR and QSPR theories*, J. Chem. Inf. Model. 51 (2011), pp. 1575–1581.
- [29] A.A. Toropov, A.P. Toropova, E. Benfenati, O. Nicolotti, A. Carotti, K. Nesmerak, A.M. Veselinović, J.B. Veselinović, P.R. Duchowicz, D. Babelo, E.A. Castro, B.F. Rasulev, D. Leszczynska, and J. Leszczynski, *Quantitative structure–activity relationships in drug design, predictive toxicology, and risk assessment, in QSPR/QSAR Analyses by Means of the CORAL Software: Results, Challenges, Perspectives*, K. Roy, ed., Hershey, IGI Global, 2015, pp. 560–585.
- [30] R.P. Verma and C. Hansch, *An approach toward the problem of outliers in QSAR*, Bioorg. Med. Chem. 13 (2005), pp. 4597–4621.
- [31] K. Roy, *On some aspects of validation of predictive quantitative structure–activity relationship models*, Expert. Opin. Drug. Discov. 2 (2007), pp. 1567–1577.