General Quantitative Structure−Property Relationship Treatment of the Refractive Index of Organic Compounds

Alan R. Katritzky,*,‡ Sulev Sild,§ and Mati Karelson*,§

Center for Heterocyclic Compounds, University of Florida, Gainesville, Florida 32611-7200, and Institute of Chemical Physics, University of Tartu, 2 Jakobi Street, EE 2400 Tartu, Estonia

A general five parameter quantitative structure−property relationship (QSPR) model ($R^2 = 0.945$, $s = 0.0155$) is described for the refractive index of a structurally diverse data set of 125 organic compounds. The model, developed with the CODESSA program, involves quantum chemical, topological, and constitutional descriptors.

INTRODUCTION

The refractive index (n) is one of the most important optical properties that is frequently employed to characterize organic compounds.[1] The refractive index is defined as a ratio of the velocity of light in a vacuum to the velocity of light in the substance of interest. It has been used as an indicator of the purity of organic compounds, but the relationship of refractive index to other optical, electrical, and magnetic properties has more significance. The relationships between refractive index and polarizability, critical temperature, surface tension, density, boiling point, among other properties, have been reviewed by Partington.[2] Refractive index is also widely used in material science to evaluate materials.[3]

Most theoretical treatments for the prediction of refractive index (n) have been developed in terms of molar refraction, which quantifies the intrinsic refractive power of the basic structural units of a material. Alternative definitions of molar refraction (R) have been proposed by Lorentz and Lorentz[4,5] (eq 1), by Gladstone and Dale[6] (eq 2), and by Vogel[7] (eq 3).

$$R_{LL} = \frac{n^2 - 1}{n^2 + 2}V \tag{1}$$

$$R_{GD} = (n - 1)V \tag{2}$$

$$R_V = nM \tag{3}$$

From any of these equations, it is relatively easy to calculate the refractive index provided the molar refraction and the molecular volume (V) or molar weight (M) are known. Since the molar refraction and the molecular volume are usually measured experimentally, such a simple approach cannot be used for compounds that lack the necessary experimental information. Hence, it is of considerable importance to be able to estimate such missing values by theoretical methods. Two major approaches have been applied for this purpose.

The first is the use of group contributions for the estimation of the molar refraction, as proposed by several authors including Vogel et al.,[8] Huggins,[9] and Goedhart.[10] All these group contribution methods have been developed from experimental data measured in the liquid phase. Limitations of the group contribution techniques are (i) the method cannot be applied when the compound of interest contains structural functionality for which the group contributions are not available and (ii) possible interactions between different groups present in the molecule that lead to the nonadditivity of a property are not accounted for.

Second, the quantitative structure−property (or activity) relationship (QSPR/QSAR) approach has also been used to estimate the molar refraction. Kier and Hall developed a very good QSPR model ($R^2 = 0.998$, $s = 0.043$) for a data set of 55 alkanes using four topological indices.[11] The same authors developed another good QSPR model ($R^2 = 0.998$, $s = 0.047$) for 24 alkyl substituted benzenes containing six to ten carbon atoms with two topological indices.[11] Predictions by these models do not require the availability of experiment data; however, they were developed for a limited range of compounds and thus have limited applicability.

Correlations with other physical properties have also been used for the calculation of refractive index. Brekke et al.[12] applied the partial least squares (PLS) method to relate physical properties, including the refractive index, of 12 component mixtures of *n*-alkanes, iso-alkanes, cyclo-alkanes, and aromatics, to the intensities and the chemical shifts of ^{13}C NMR spectra. The data set included a total of 24 samples, and the standard error of the prediction was 2.5%.

Despite the wide availability of data on refractive index, no general QSPR relationship has been proposed relating the refractive index of organic compounds with the chemical structure. Therefore, it was decided to use our recently developed CODESSA[13] program in an attempt to develop a general QSPR relationship for the prediction of the refractive index from the information encoded in the chemical structure and quantum-chemical wave function of a compound, without any requirement for experimental data. CODESSA has previously been successfully used to calculate gas chromatographic response factors and retention times,[14] boiling points[15] and melting points[16] of organic compounds,

‡ University of Florida.
§ University of Tartu.

QSPR of Refractive Index of Organic Compounds

J. Chem. Inf. Comput. Sci., Vol. 38, No. 5, 1998 **841**

Table 1. Experimental and Calculated Refractive Index Values for a Diverse Set of Organic Compounds

compd	exptl n	calcd n	Δn	compd	exptl n	calcd n	Δn
1-chloro-2-butene	1.4351	1.4325	−0.0026	*n*-butyl iodide	1.4998	1.4966	−0.0032
1-chloronaphthalene	1.6332	1.6111	−0.0221	*n*-butyl *n*-butyrate	1.4064	1.4034	−0.0030
2-chlorobutane	1.3960	1.3994	0.0034	*n*-butylamine	1.4010	1.3716	−0.0294
2-chlorobutene	1.4205	1.4358	0.0153	*n*-butylbenzene	1.4904	1.5224	0.0320
2-chloroethyl ether	1.4560	1.4483	−0.0077	*n*-decyl alcohol	1.4359	1.4355	−0.0004
2-chloroethyl vinyl ether	1.4380	1.4430	0.005	*n*-dibutylamine	1.4177	1.4181	0.0004
2-ethoxyethyl acetate	1.4055	1.4047	−0.0008	*n*-dipropylamine	1.4046	1.4044	−0.0002
2-ethoxyethyl ether	1.4120	1.3926	−0.0194	*n*-hexane	1.3750	1.3740	−0.0010
2-pentanol	1.4060	1.3834	−0.0226	*n*-propyl chloride	1.3886	1.3910	0.0024
3-chloro-1-butene	1.4150	1.4220	0.007	*o*-bromotoluene	1.5550	1.5567	0.0017
acetal	1.3819	1.3793	−0.0026	*o*-chloroaniline	1.5895	1.5771	−0.0124
acetaldehyde	1.3316	1.3575	0.0259	*o*-dichlorobenzene	1.5515	1.5402	−0.0113
acetyl chloride	1.3898	1.3837	−0.0061	octane	1.3976	1.3958	−0.0018
allyl alcohol	1.4135	1.4263	0.0128	octanol	1.4300	1.4268	−0.0032
allyl bromide	1.4655	1.4621	−0.0034	octyl acetate	1.4204	1.4070	−0.0134
aniline	1.5863	1.5537	−0.0326	*p*-bromotoluene	1.5490	1.5606	0.0116
benzaldehyde	1.5456	1.4874	−0.0582	*p*-fluoroaniline	1.5195	1.5715	0.0520
benzene	1.5011	1.4754	−0.0257	pentane	1.3577	1.3633	0.0056
benzonitrile	1.5289	1.5143	−0.0146	phenyl acetate	1.5030	1.5095	0.0065
benzyl acetate	1.5232	1.5138	−0.0094	propyl acetate	1.3844	1.3861	0.0017
benzyl alcohol	1.5404	1.5276	−0.0128	propyl alcohol	1.3862	1.3788	−0.0074
benzyl ethyl ether	1.4955	1.5137	0.0182	propyl bromide	1.4341	1.4445	0.0104
benzylamine	1.5401	1.4932	−0.0469	propyl butyrate	1.4005	1.3966	−0.0039
bromobenzene	1.5602	1.5477	−0.0125	propylamine	1.3890	1.3592	−0.0298
butyraldehyde	1.3790	1.3849	0.0059	propylbenzene	1.4919	1.5127	0.0208
butyronitrile	1.3839	1.3895	0.0056	*sec*-butyl chloride	1.3960	1.3994	0.0034
butyryl chloride	1.4120	1.4135	0.0015	*tert*-butyl alcohol	1.3847	1.3722	−0.0125
carbon tetrachloride	1.4607	1.4720	0.0113	*tert*-butylamine	1.3700	1.3621	−0.0079
chlorobenzene	1.5248	1.5082	−0.0166	propanoic acid	1.3869	1.4004	0.0135
cycloheptanone	1.4608	1.4381	−0.0227	methyl propionate	1.3775	1.3875	0.0100
cyclohexane	1.4264	1.3953	−0.0311	ethyl propionate	1.3839	1.3877	0.0038
cyclohexanone	1.4507	1.4276	−0.0231	ethanol	1.3611	1.3591	−0.0020
cyclohexene	1.4465	1.4501	0.0036	chloroethane	1.3676	1.3766	0.0090
cyclohexyl chloride	1.4626	1.4419	−0.0207	ethyl cyanide	1.3655	1.3716	0.0061
cyclopentanone	1.4366	1.4126	−0.0240	ethyl acetate	1.3723	1.3848	0.0125
cyclopropyl methyl ether	1.3802	1.3937	0.0135	ethylbenzene	1.4959	1.5007	0.0048
di-*tert*-butyl ether	1.3949	1.4135	0.0186	1-ethyl-2-methylbenzene	1.5046	1.5169	0.0123
diethoxymethane	1.3730	1.3764	0.0034	diethyl ether	1.3526	1.3688	0.0162
diethoxymethyl acetate	1.4000	1.4113	0.0113	butyl propionate	1.4014	1.3987	−0.0027
diethylamine	1.3864	1.3837	−0.0027	ethyl hexyl ether	1.4008	1.4028	0.0020
dimethoxyethane	1.3813	1.3787	−0.0026	methyl isobutyrate	1.3840	1.3875	0.0035
ethyl benzoate	1.5060	1.5110	0.0050	ethyl isobutyrate	1.3869	1.3937	0.0068
ethyl bromide	1.4242	1.4355	0.0113	methanol	1.3288	1.3536	0.0248
ethyl butyrate	1.4000	1.3934	−0.0066	ethanol	1.3611	1.3678	0.0067
ethylaniline	1.5559	1.5119	−0.0440	ethyl octyl ether	1.4127	1.4113	−0.0014
isobutyl iodide	1.4960	1.4913	−0.0047	ethyl pentyl ether	1.3927	1.3941	0.0014
isopropyl alcohol	1.3772	1.3689	−0.0083	ethyl isobutyl ether	1.3739	1.3870	0.0131
isopropyl ether	1.3689	1.3914	0.0225	isobutyl isobutyrate	1.3999	1.4261	0.0262
isopropyl iodide	1.5026	1.4873	−0.0153	2-methylpentane	1.3715	1.3705	−0.0010
isopropylacetone	1.3960	1.4084	0.0124	ethyl chloroacetate	1.4215	1.4165	−0.0050
m-bromotoluene	1.5510	1.5598	0.0088	*n*-propyl isobutyrate	1.3955	1.4005	0.0050
m-dichlorobenzene	1.5459	1.5371	−0.0088	1,2,3,5-tetramethylbenzene	1.5130	1.5331	0.0201
methyl acetate	1.3614	1.3657	0.0043	*p*-xylene	1.4958	1.5062	0.0104
methyl formate	1.3440	1.356	0.0120	3-hexanone	1.4004	1.4050	0.0046
methyl propyl ketone	1.3895	1.3976	0.0081	ethyl benzoate	1.5007	1.5110	0.0103
methyl vinyl ketone	1.4086	1.4005	−0.0081	9-ethylcarbazole	1.6394	1.6677	0.0283
n-amyl bromide	1.4444	1.4715	0.0271	1-ethylnaphthalene	1.6062	1.6047	−0.0015
n-amyl chloride	1.4128	1.4164	0.0036	1,4-dichlorobenzene	1.5285	1.5406	0.0121
n-amyl ether	1.4119	1.4078	−0.0042	2-propanol	1.3776	1.3689	−0.0087
n-butyl acetate	1.3951	1.3964	0.0013	1-butoxy-2-propanol	1.4168	1.4103	−0.0065
n-butyl alcohol	1.3993	1.3942	−0.0051	2-propanethiol	1.4255	1.4380	0.0125
n-butyl bromide	1.4398	1.4587	0.0189	1,1-dichloroethane	1.4164	1.4092	−0.0072
n-butyl chloride	1.4022	1.4050	0.0028				

critical micelle concentrations of surfactants,[17] glass transition temperatures of polymers,[18] and other properties.[19]

METHODOLOGY

The refractive indices of 125 diverse organic compounds used in the QSPR analysis with the CODESSA program were collected from the *Merck Index CD-ROM*[20] and *Handbook of Chemistry and Physics*[21] (as summarized in Table 1). The experimental refractive index values used were measured at 293 K and a wavelength of 589 nm. Since the objective of the current work was to develop a general QSPR relationship widely applicable to different classes of organic compounds, the molecules in the training data set were chosen to form as diverse a data set as possible. This set contains hydroxyl,

Table 2. Best Five Parameter Correlation for Refractive Index (R^2 = 0.945, F = 411.29, s^2 = 0.0003)

X	DX	t-test	descriptor
1.050	4.685×10^{-2}	22.513	intercept
-1.362×10^{-2}	1.629×10^{-3}	-8.358	HOMO−LUMO energy gap
2.487×10^{-3}	2.551×10^{-4}	9.748	min E−N attraction for a C atom
-7.119×10^{-4}	6.636×10^{-5}	-10.728	PPSA-2 [Zefirov's PC]
1.476×10^{-3}	1.370×10^{-4}	10.773	HDSA [semi-MO PC]
2.544×10^{-4}	1.473×10^{-5}	17.268	gravitation index (all bonds)

Figure 1. Plot of best five parameter correlation for refractive index.

amino, ether, ester, carbonyl, cyano, and carboxylic functionalities, halogenated compounds, and both saturated and unsaturated hydrocarbons.

The three-dimensional structures of the molecules were drawn with the PCMODEL[22] program and preoptimized with the MMX molecular mechanics method. The final geometries and quantum-chemical data were obtained with the semiempirical AM1 method[23] using the MOPAC 6.0 program.[24] All geometries and electronic parameters were calculated in a vacuum without taking into account possible solvent effects. The descriptors for each compound in the data set were calculated with the CODESSA program and stored in the database for further regression analysis using the information from the output files of the MOPAC program. The descriptors used in our QSPR study can be divided into four classes: constitutional, topological, electrostatical, and quantum-chemical.[25] For the development of the QSPR model, the heuristic method and the best multilinear correlation method were employed in the framework of CODESSA.[15] The cross-validated regression coefficients were also evaluated to measure the stability of the QSPR model by stepwise removal of any single data point.

By definition, the refractive index in a vacuum is unity. Assuming that all the descriptors involved in the QSPR model have zero values in a vacuum, the refractive index in a vacuum should be determined only by the intercept. Thus, any physically meaningful QSPR model should have an intercept equal to 1. The normal correlation methods yielded numerous statistically significant correlations but all with unacceptable intercept values (see Appendix as an example). Therefore, the best multilinear correlation program module was modified to fix the intercept value as 1. A subset of descriptors selected by the modified correlation program to the correct QSPR model were recalculated with CODESSA's original multilinear correlation method (coefficients for parameters and intercept remained comparable with both original and modified correlation method) for the later prediction of refractive index values.

RESULTS AND DISCUSSION

The CODESSA treatment resulted in a five parameter correlation (details are in Table 2) with a squared correlation coefficient of 0.945 and standard error of 0.0155. The predicted refractive index values along with the deviation from experimental values are listed in Table 1. The linear plot of experimental property versus calculated property is given in Figure 1. The calculated cross-validated correlation coefficient (R_{cv}^2 = 0.937) confirms the stability of the final QSPR model. An alternative method for cross-validation was also used. The data set of experimental refractive indices was divided into three subsets based on their magnitude. When two of the subsets were combined and
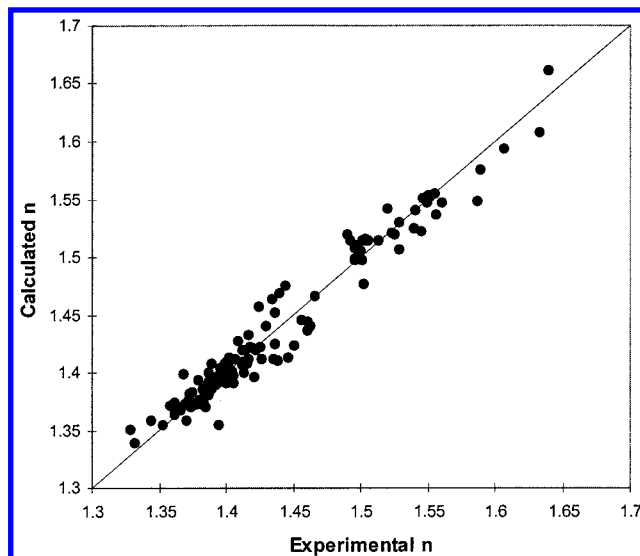
the QSPR model was recalculated, the predicted refractive indices for the third subset gave a squared correlation coefficient of 0.954. We applied similar procedures to calculate the squared correlation coefficients (0.949 and 0.925) for the other two subsets. The average correlation coefficient for the three subsets was 0.943. The results of the two methods are similar to each other; the second method is slightly less sensitive to outliers (the largest outlier is benzaldehyde with a 3.7% prediction error, the average prediction error is 0.8%). We therefore concluded that the final QSPR model can be considered stable.

The first descriptor in the correlation model is the HOMO−LUMO energy gap, which is defined as the energy difference between highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). The involvement of this particular descriptor is to be expected, as previous theoretical work by Hervé and Vandamme et al. showed an empirical relationship between the refractive index and the energy gap in semiconductors.[26] Both the refractive index and the HOMO−LUMO energy gap are related to the polarizability of a molecule. A smaller difference between HOMO−LUMO energies usually means that the molecule is relatively easier to polarize.[25]

The second descriptor is the lowest (absolute value) electron−nucleus attraction energy for a carbon atom. This descriptor represents energy partitioning on atoms in the molecule. For example, carbon atoms with a low electron−nucleus attraction energy are usually part of a polar functional group (carboxylic, carbonyl, etc.). This descriptor is related to the polarity of the molecule.

The third descriptor is the total charge weighted partial positively charged surface area (PPSA2), which is calculated by eq 4, where q_A is a partial charge of an atom calculated by Zefirov's empirical method and S_A is the surface area of the atom. This descriptor is related to the charge distribution in the molecule.

$$PPSA2 = \sum_A q_A \sum_A S_A \qquad (4)$$

The fourth descriptor in the model is the surface area of hydrogen donor atoms (HDSA) and describes specifically

QSPR of Refractive Index of Organic Compounds

J. Chem. Inf. Comput. Sci., Vol. 38, No. 5, 1998 **843**

Table 3. Experimental and Calculated Refractive Index Values for External Validation Set

compd	exptl n	calcd n	Δn
2-methylpyridine	1.501	1.4863	0.0147
2,3-dibromopropene	1.5157	1.5361	0.0204
2,6-di-*tert*-butylpyridine	1.5733	1.5608	0.0125
2-amino-1-butanol	1.453	1.4142	0.0388
2-buten-1-ol	1.4285	1.4584	0.0299
4-methylquinoline	1.619	1.5875	0.0315
acetone	1.3591	1.3689	0.0098
allyl chloride	1.4154	1.4123	0.0031
allyl cyanide	1.406	1.4211	0.0151
allyl ethyl ether	1.3881	1.4156	0.0275
benzothiazole	1.6379	1.6345	0.0034
diethylene glycol	1.4475	1.4576	0.0101
ethyl formate	1.3567	1.3734	0.0167
fluorobenzene	1.4677	1.4764	0.0087
formic acid	1.3714	1.3859	0.0145
isobutylbenzene	1.4928	1.5252	0.0324
2,2′-dichlorodiethyl sulfide	1.5313	1.5276	0.0037
phenyl methyl ketone	1.5339	1.498	0.0359
p-isopropylbenzaldehyde	1.5301	1.524	0.0061
p-methylphenol	1.5395	1.5267	0.0128
propylene	1.3567	1.3902	0.0335
sec-butylamine	1.394	1.3682	0.0258
styrene	1.5463	1.5164	0.0299

the hydrogen bonding interactions in the medium. The HDSA is calculated by eq 5, where S_D is surface area of the hydrogen donor atom.

$$HDSA = \sum_{D} S_D \qquad (5)$$

The fifth descriptor is the gravitation index (G_I), which is calculated over all bonded atoms according to eq 6, where m_i and m_j are the atomic masses of the bonded atoms and r_{ij} is the corresponding bond length. This descriptor is related to the effective mass distribution and describes molecular size dependent effects of the molecule. The gravitation index has been used effectively in recent studies to describe intermolecular dispersion forces in bulk liquid media.[15,27]

$$G_I = \sum_{\substack{\text{bonded atoms}}}^{ij} \left(\frac{m_i m_j}{r_{ij}^2} \right) \qquad (6)$$

To test the applicability of the QSPR model to unknown systems, another data set with 25 randomly chosen diverse structures was compiled. Semiempirical and CODESSA calculations were performed to obtain the descriptors necessary for predicting refractive index. The refractive index values predicted by the CODESSA are listed in Table 3, and a linear plot of experimental versus predicted refractive index values is presented in Figure 2. The performance of the model on the external validation set is satisfactory, with an average prediction error of 1.29%.

CONCLUSION

We have developed a general five parameter QSPR correlation with a correlation coefficient of 0.945 for the refractive index of a diverse set of organic compounds. Unlike previous approaches using the group contribution technique, this method does not rely on experimental parameters. Instead, predictions of refractive index are possible solely on the basis of the molecular structure. The
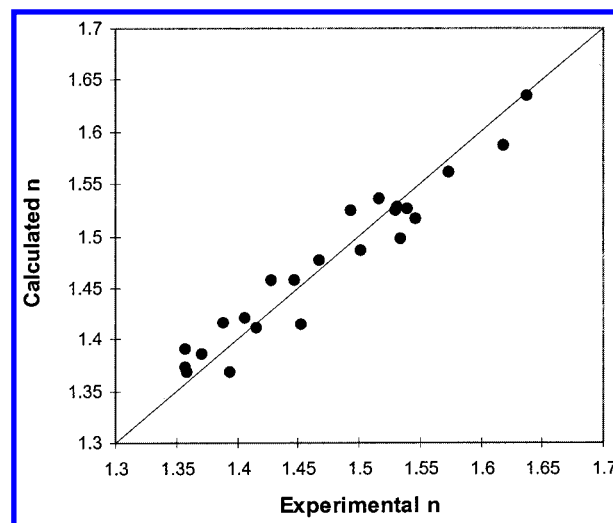
Figure 2. Plot of experimental versus calculated refractive index values for the external validation set.

Table 4. Best Four Parameter Correlation (with Inadmissible Intercept of 1.74) for Refractive Index ($R^2 = 0.955$, $F = 641.83$, $s^2 = 0.0002$)

X	DX	t-test	descriptor
1.7393	3.192×10^{-2}	54.497	intercept
-9.229×10^{-3}	4.252×10^{-4}	-21.704	number of single bonds
6.641×10^{-2}	3.469×10^{-3}	19.142	Kier and Hall index (order 1)
3.302×10^{-2}	2.590×10^{-3}	12.748	HOMO-1 energy
3.809×10^{-2}	3.621×10^{-3}	10.519	HA dependent HDCA-2 [semi-MO PC]

predicted values have an average error of 0.8% when compared with the experimental values; therefore this QSPR relationship can be used for the prediction of refractive indices with a high degree of confidence.

ACKNOWLEDGMENT

APPENDIX

As a cautionary example of a chance correlation, one result from correlation analysis without fixed intercept is also presented (see details in Table 4). This result appears to be statistically better, with a correlation coefficient of 0.955 and one less descriptor, than the result for the physically correct model with the fixed intercept. The descriptors involved are the number of single bonds, the first order Kier and Hall index, the energy of the second highest occupied molecular orbital (HOMO-1), and the charged surface area of hydrogen donors. The first two descriptors are probably related with the bulk properties of the media. Also, note that the first descriptor has a negative slope, which is not reasonable considering that the refractive index in a homologous row should increase when the size of the molecule increases. The third descriptor is related to the HOMO−LUMO energy gap, and the fourth descriptor accounts for hydrogen bonding related effects.

REFERENCES AND NOTES

(1) Finar, I. L. *Organic Chemistry*, 2nd ed.; Longman: London, 1975; Vol. 2.

(2) Partington, J. R. *An Advanced Treatise on Physical Chemistry*; Longmans, Green: London, 1953; Vol. 4.

(3) Bicerano, J. *Prediction of Polymer Properties*, 2nd ed.; Marcel Dekker Inc.: New York, 1996.

(4) Lorentz, H. A. *Wied. Ann. Phys.* **1880**, *9*, 641−665.

(5) Lorentz, L. V. *Wied. Ann. Phys.* **1880**, *11*, 70−103.

(6) Gladstone, J. H.; Dale, T. P. *Philos. Trans. R. Soc. (London)* **1858**, *A148*, 887−894.

(7) Vogel, A. Physical Properties and Chemical Constitution. Part XXIII. Miscellaneous Compounds. Investigation of So-Called Coordinate or Dative Link in Esters of Oxy-Acids and in Nitro-Paraffins by Molecular Refractivity Determinations. Atomic, Structural, and Group Parachors and Refractivities. *J. Chem. Soc.* **1948**, 1833−1855.

(8) Vogel, A. I.; Cresswell, W. T.; Jeffery, G. H.; Leicester, J. Calculation of the Refractive Indices of Liquid Organic Compounds: Bond Molecular Refraction Coefficients. *Chem. Ind.* **1951**, *5*, 376−376.

(9) Huggins, M. Densities and Optical Properties of Organic Compounds in the Liquid State. VI. The Refractive Indices of Paraffin Hydrocarbons and Some of Their Derivatives. *Bull. Chem. Soc. Jpn.* **1956**, *29*, 336−339.

(10) Goedhart, D. J. *Communication Gel Permeation Chromatography International Seminar*, Monaco, Oct 12−15, 1969.

(11) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis*; Research Studies Press, Ltd.: Letchworth, Hertfordshire, England, 1986; p 26.

(12) Brekke, T.; Kvalheim, O. M.; Sletten, E. Prediction of Physical Properties of Hydrocarbon Mixtures by Partial-Least-Squares Calibration of Carbon-13 Nuclear Magnetic Resonance Data. *Anal. Chim. Acta* **1989**, *223*, 123−134.

(13) Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *CODESSA, Reference Manual,* University of Florida: Gainesville, FL, 1994.

(14) Katritzky, A. R.; Ignatchenko, E. S.; Barcock, R. A.; Lobanov, V. S.; Karelson, M. Prediction of Gas Chromatographic Retention Times and Response Factors Using a General Quantitative Structure−Property Relationship Treatment. *Anal. Chem.* **1994**, *66*, 1799−1807.

(15) Katritzky, A. R.; Mu, L.; Lobanov, V. S.; Karelson, M. Correlation of Boiling Points with Molecular Structure. 1. A Training Set of 298 Diverse Organics and a Test Set of 9 Simple Inorganics. *J. Phys. Chem.* **1996**, *100*, 10400−10407.

(16) Katritzky, A. R.; Maran, U.; Karelson, M.; Lobanov, V. S. Prediction of Melting Points for the Substituted Benzenes: A QSPR Approach. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 913−919.

(17) Huibers, P. D. T.; Lobanov, V. S.; Katritzky, A. R.; Shah, D. O.; Karelson, M. Prediction of Critical Micelle Concentration Using a Quantitative Structure−Property Relationship Approach. *J. Colloid Interface Sci.* **1997**, *187*, 113−120.

(18) Katritzky, A. R.; Sild, S.; Lobanov, V.; Karelson, M. QSPR Correlation of Glass Transition Temperatures of High Molecular Weight Polymers. *J. Chem. Inf. Comput. Sci.*, in press.

(19) Katritzky, A. R.; Lobanov, V. S.; Karelson, M. QSPR: The Correlation and Quantitative Prediction of Chemical and Physical Properties from Structure. *Chem. Soc. Rev.* **1995**, 279−287.

(20) *The Merck Index on CD-ROM, User Guide*, 12th ed.; Chapman & Hall: London, 1996.

(21) *CRC Handbook of Chemistry and Physics*, Weast, R. C., Ed.; CRC Press: Boca Raton, FL, 1985.

(22) *PCMODEL User Manual*; Serena Software: Bloomington, IN, 1992.

(23) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.

(24) Stewart, J. J. P. *MOPAC 6.0;* QCPE No 455, 1989.

(25) Karelson, M.; Lobanov, V. S.; Katritzky, A. R. Quantum-Chemical Descriptors in QSAR/QSPR Studies. *Chem. Rev.* **1996**, *96*, 1027−1043.

(26) Hervé, P.; Vandamme, L. K. J. General Relation Between Refractive Index and Energy Gap in Semiconductors. *Infrared Phys. Technol.* **1994**, *35*, 609−615.

(27) Katritzky, A. R.; Mu, L.; Karelson, M. A QSPR Study of the Solubility of Gases and Vapors in Water. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1162−1168.