*Asian Journal of Chemistry* Vol. 22, No. 9 (2010), 6747-6760

# Quantitative Structural Relationship and Estimated Biological and Physicochemical Properties of Paraffin-Base Petroleum ($C_1$-$C_{40}$)

KHOJASTEH KHERADMAND
*Young Researchers Club of Arak Branch, Chemistry Department, Faculty of Science, Islamic Azad University, P.O. Box 38135-567, Arak, Iran*
*E-mail: auni.irb45@gmail.com; khojasteh.kheradmand@gmail.com*

The hydrocarbon compounds in petroleum have different structures. Some of them are linear and normal. This homologues series sometimes has attracted the notice for chemical and biochemical studies. The structural parameters are the useful indices for examination of structure-property relationship. Graph theory is an interesting ground for the exploration of proof techniques in discrete mathematics and its results have applications in many areas of sciences. Topological indices are the numerical value associated with chemical constitution purporting for correlation of chemical structure with various physical properties, chemical reactivity or biological activity. This study represented the quantitative structural relationship (QSR) between the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$, logarithm of calculated octanol-water partitioning coefficients and total biodegradation [log $K_{ow}$ and $TB_d$ (mol/h), respectively], median lethal concentration 50 ($LC_{50}$), water solubility ($S_w$, mg $L^{-1}$/25 ºC) and some of the other calculated chemical and biochemical data in the normal hydrocarbon $C_1$-$C_{40}$ (**1-40**). The results would be extended for some of the other members of these compounds. The interesting results of concerning among log $K_{ow}$, $LC_{50}$, $S_w$, $TB_d$ (g/h) and some of the other data with the descriptors for this compounds are presented and discussed.

**Key Words: Octanol-water partitioning coefficient, Biodegradation, Molecular topology, $LC_{50}$, Water solubility, Topological index, Normal hydrocarbon.**

## INTRODUCTION

Liquid geologically-extracted hydrocarbons are referred to as petroleum (literally "rock oil") or mineral oil, while gaseous geologic hydrocarbons are referred to as natural gas. All are significant sources of fuel and raw materials as a feedstock for the production of organic chemicals and are commonly found in the Earth's subsurface using the tools of petroleum geology. The extraction of liquid hydrocarbon fuel from a number of sedimentary basins has been integral to modern energy development. Hydrocarbons are mined from tar sands, oil shale and potentially extracted from sedimentary methane hydrates. These reserves require distillation and upgrading to produce synthetic crude and petroleum. Oil reserves in sedimentary rocks are

the source of hydrocarbons for the energy, transport and petrochemical industries. Hydrocarbons are of prime economic importance because they encompass the constituents of the major fossil fuels (coal, petroleum, natural gas, *etc*.) and its derivatives plastics, paraffin, waxes, solvents and oils. Hydrocarbons are one of the Earth's most important energy resources. The predominant use of hydrocarbons is as a combustible fuel source. In their solid form, hydrocarbons take the form of asphalt[1-4]. Mixtures of volatile hydrocarbons are now used in preference to the chlorofluorocarbons as a propellant for aerosol sprays, due to chlorofluorocarbons impact on the ozone layer. Methane ($C_1$) and ethane ($C_2$) are gaseous at ambient temperatures and cannot be readily liquefied by pressure alone. Propane ($C_3$) is however easily liquified and exists in 'propane bottles' mostly as a liquid. Butane ($C_4$) is so easily liquified that it provides a safe, volatile fuel for small pocket lighters. Pentane ($C_5$) is a clear liquid at room temperature, commonly used in chemistry and industry as a powerful nearly odorless solvent of waxes and high molecular weight organic compounds, including greases. Hexane ($C_6$) is also a widely used non-polar, non-aromatic solvent, as well as a significant fraction of common gasoline[1-4]. The $C_6$ through $C_{10}$ alkanes, alkenes and isomeric cycloalkanes are the top components of gasoline, naphtha, jet fuel and specialized industrial solvent mixtures. With the progressive addition of carbon units, the simple non-ring structured hydrocarbons have higher viscosities, lubricating indices, boiling points, solidification temperatures and deeper colour. At the opposite extreme from methane ($C_1$) lie the heavy tars that remain as the lowest fraction in a crude oil refining retort. They are collected and widely utilized as roofing compounds, pavement composition, wood preservatives (the creosote series) and as extremely high viscosity sheer-resisting liquids[1-4].

Investigation of the different useful applications of graph theory shows obviously that this area is an exploration of techniques in discrete mathematics and its results can be applied in various fields of sciences. Graph theory has been found to be an effective tool in QSAR and QSPR[5-10]. A graph is a topological concept rather than a geometrical concept of fixed geometry and hence Euclidean metric lengths, angles and three-dimensional spatial configurations have no meaning. Numerous studies have been made relating to these mentioned fields by using what are called topological indices (TI)[10]. All the applications proved that can utilize the topological indices as very useful for molecular structure studies like description and suggestion of chemical and structural properties, biological and toxicological characters of compounds. One of the stages of topological indices (TI) started when Randic introduced the molecular branching index[11]. Randic[11] proposed a topological index that has become one of the most widely used in both QSAR and QSPR studies. However, the most important contribution of this stage is probably the great number of applications of topological indices in several fields of chemistry. The topological indices are based on the original idea of Randic of molecular branching but extended to account for contributions coming from path clusters, clusters and chains of different lengths[12-19].

The new jump of research in topological indices probably have started in the years of 90's. At the middle of the 1990 decade the number of studies and applications of topological indices in chemistry was increased[20,27]. Among the successful topological indices in these applications, it is worth to mention the molecular connectivity indices[20,21], including the Randic index[11,20,30,34,35], the indices of Kier[24], elecro topological state indices[25], Balaban index[26] and Wiener index[23]. Trinajstic and coworkers reported that 39 topological indices are presently available in the literatures[28]. Estrada has made important studies in terms of generalized topological indices with several topological indices in the graph invariant[29]. In 1993 and 1997 were reported a related complex of application of the Wiener and Harary indices in fullerene science[36,37]. The use of the effective mathematical methods for making good correlations between several data properties of chemicals and the indices were reported[36-39].

It needs to use the effective and useful mathematical methods for making good concern between several data of chemical properties, medicinal chemistry and biological activity of chemicals. The octanol-water partition coefficient ($K_{ow}$) is a measure of the equilibrium concentration of a compound between octanol and water that indicates the potential for partitioning in to soil organic matter (*i.e.*, a high $K_{ow}$ indicates a compound which will preferentially partition into soil organic matter rather than water). This coefficient is inversely related to the solubility of a compound in water. The log $K_{ow}$ is used in models to estimate plant and soil invertebrate bioaccumulation factors. log $K_{ow}$ is commonly used in QSAR studies and drug design, since this property is related to drug absorption, bioavailability, metabolism and toxicity. This parameter is also used in many environmental studies to help determine the environmental fate of chemicals[31-33].

Biodegradation ($TB_d$) is a useful and important factors in chemical and biochemical studies[32]. In this study, it is considered the relationship of the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$ (**1-40**) as a important class of the saturated hydrocarbons as molecular descriptions of structure-property relationship studies, logarithm of calculated octanol-water partitioning coefficients and total biodegradation, log $K_{ow}$ and $TB_d$ (mol/h and g/h), respectively, median lethal concentration 50 ($LC_{50}$), water solubility ($S_w$, mg $L^{-1}$/25 ºC), HLC (in atm $m^3$/mol), $EC_{50}$ green algae (in ppm), logarithm values of bioconcentration factor as log [BCF], bio half-life, logarithm values of bioaccumulation factor as log [BAF], adsorption coefficient ($K_{oc}$), fugacity in air and soil, reaction air (kg/h) and advection per cent sediment in $C_1$-$C_{40}$ compounds **1-40** (Table-1)[31-33,40,41].

## GRAPHING AND MATHEMATICAL METHOD

The number of carbon atoms in the normal hydrocarbon $C_1$-$C_{40}$ (**1-40**) "n" factor is the favoured value in the structures. For modeling operations was used and investigated linear (MLR) and nonlinear (ANN) models in this study. To calculate

TABLE-1
SELECTED PHYSICOCHEMICAL PROPERTIES OF $C_1$-$C_{40}$

| No. C | log $K_{ow}$ | $TB_d$ (g/h) $10^{-2}$ | $TB_d$ (mol/h) $10^{-4}$ | Exp. log $K_{oa}$ | Est. log $K_{oa}$ | Exp. HLC (atm $m^3$/mol) | log $K_{aw}$ | -log WSol. (mol/lit) | $LC_{50}$ fish (mg/lit) | $EC_{50}$ green algae (ppm) | log BCF | Bio half life | log BAF | Soil $K_{oc}$ | Fugacity air $10^{-12}$ | Fugacity soil $10^{-8}$ | Reaction air $10^3$ (kg/h) | Advection per cent sediment |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 1.09 | 0.177 | 1.10 | -0.38 | -0.340 | 0.66 | 1.43 | 0.789 | 147.544 | 19.372 | 0.39 | 0.1004 | 0.28 | 3.98 | 374 | 85.1 | 455 | 0.000246 |
| **2** | 1.81 | 0.182 | 0.60 | 0.42 | 0.499 | 0.50 | 1.311 | 1.506 | 91.243 | 16.474 | 0.86 | 0.2706 | 0.84 | 13.20 | 186 | 45.4 | 166 | 0.000293 |
| **3** | 2.36 | 0.191 | 0.43 | 0.97 | 0.899 | 0.71 | 1.461 | 2.077 | 49.472 | 11.888 | 1.22 | 0.3887 | 1.32 | 21.70 | 104 | 30.9 | 582 | 0.000337 |
| **4** | 2.89 | 0.220 | 0.38 | 1.53 | 1.301 | 0.95 | 1.589 | 2.632 | 24.109 | 7.711 | 1.57 | 0.5954 | 1.79 | 39.60 | 58.9 | 23.4 | 1.22 | 0.000307 |
| **5** | 3.39 | 0.316 | 0.44 | 1.96 | 1.682 | 1.25 | 1.708 | 3.161 | 11.065 | 4.710 | 1.9 | 0.8255 | 2.17 | 72.20 | 38.8 | 18.8 | 1.42 | 0.00041 |
| **6** | 3.90 | 0.555 | 0.46 | 2.40 | 2.033 | 1.80 | 1.867 | 3.699 | 4.886 | 2.768 | 2.24 | 1.153 | 2.49 | 132 | 26.6 | 15.8 | 1.55 | 0.00058 |
| **7** | 4.66 | 1.200 | 1.20 | 2.95 | 2.747 | 2.00 | 1.913 | 4.450 | 2.101 | 1.584 | 2.74 | 1.921 | 2.84 | 240 | 19.6 | 13.6 | 1.65 | 0.000852 |
| **8** | 5.18 | 1.550 | 1.40 | 3.35 | 3.062 | 3.21 | 2.118 | 4.996 | 0.885 | 0.888 | 3.08 | 2.922 | 3.06 | 437 | 15.1 | 11.9 | 1.73 | 0.00128 |
| **9** | 5.65 | 1.710 | 1.30 | 5.65 | 3.507 | 3.40 | 2.143 | 5.500 | 0.368 | 0.491 | 2.02 | 3.965 | 3.88 | 796 | 12.0 | 10.6 | 1.79 | 0.00198 |
| **10** | 5.01 | 1.450 | 1.00 | _ | 2.687 | 5.15 | 2.323 | 5.055 | 0.151 | 0.268 | 1.6 | 2.454 | 2.97 | 1450 | 9.8 | 9.53 | 1.83 | 0.00321 |
| **11** | 5.74 | 1.780 | 1.10 | _ | 3.843 | 1.93 | 1.897 | 5.784 | 0.061 | 0.145 | 2.08 | 4.013 | 3.3 | 2640 | 8.02 | 8.58 | 1.86 | 0.00528 |
| **12** | 6.10 | 1.770 | 1.00 | _ | 3.576 | 8.18 | 2.524 | 6.190 | 0.025 | 0.078 | 2.32 | 5.026 | 2.86 | 4820 | 6.94 | 7.93 | 1.85 | 0.00887 |
| **13** | 6.73 | 2.300 | 1.20 | _ | 4.659 | 2.88 | 2.071 | 6.827 | 0.010 | 0.041 | 2.73 | 7.617 | 4.6 | 8780 | 5.67 | 7.16 | 1.86 | 0.0138 |
| **14** | 7.20 | 2.260 | 1.10 | _ | 4.625 | 9.20 | 2.575 | 7.334 | 0.004 | 0.022 | 3.43 | 10.37 | 5.71 | 16000 | 4.42 | 6.74 | 1.81 | 0.0191 |
| **15** | 7.71 | 2.840 | 1.30 | _ | 4.998 | 12.60 | 2.712 | 7.870 | 0.00155 | 0.012 | 3.18 | 14.46 | 4.86 | 29200 | 3.47 | 6.26 | 1.34 | 0.0197 |
| **16** | 8.20 | 7.470 | 3.30 | _ | 6.914 | 0.47 | 1.286 | 8.391 | 0.000613 | 0.006 | 2.94 | 19.92 | 5.28 | 53200 | 2.14 | 2.86 | 1.7 | 0.0156 |
| **17** | 8.69 | 4.210 | 1.70 | _ | 5.493 | 38.50 | 3.197 | 8.913 | 0.000241 | 0.003 | 2.7 | 27.44 | 4.8 | 96900 | 3.04 | 5.52 | 1.71 | 0.0104 |
| **18** | 9.18 | 5.460 | 2.10 | _ | 5.859 | 51.20 | 3.321 | 9.435 | 0.0000942 | 0.00165 | 2.46 | 37.8 | 4.64 | 177000 | 2.71 | 5.25 | 1.71 | 0.0152 |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **19** | 9.67 | 6.51 | 2.4 | _ | 6.227 | 67.9 | 3.443 | 9.956 | 0.0000367 | 0.000855 | 2.22 | 52.08 | 4.43 | 322000 | 2.42 | 4.94 | 1.71 | 0.00915 |
| **20** | 10.16 | 7.1 | 2.5 | _ | 6.593 | 90.2 | 3.567 | 10.478 | 0.0000143 | 0.000443 | 1.98 | 71.75 | 4.19 | 586000 | 2.17 | 4.65 | 1.71 | 0.00543 |
| **21** | 10.65 | 7.5 | 2.5 | _ | 6.959 | 120 | 3.691 | 10.999 | 0.00000555 | 0.000229 | 1.74 | 98.85 | 3.91 | 1070000 | 1.96 | 4.38 | 1.71 | 0.00321 |
| **22** | 11.15 | 7.66 | 2.5 | _ | 7.337 | 59 | 3.813 | 11.521 | 0.00000215 | 0.000118 | 1.5 | 136.2 | 3.61 | 1950000 | 1.76 | 4.11 | 1.71 | 0.00185 |
| **23** | 11.64 | 7.73 | 2.4 | _ | 7.704 | 211 | 3.936 | 12.042 | 0.00000083 | 0.0000606 | 1.26 | 187.6 | 3.29 | 3550000 | 1.60 | 3.84 | 1.7 | 0.0117 |
| **24** | 12.13 | 7.76 | 2.3 | _ | 8.071 | 280 | 4.059 | 12.564 | 0.00000032 | 0.0000311 | 1.61 | 258.5 | 2.94 | 6400000 | 1.44 | 3.57 | 1.68 | 0.00691 |
| **25** | 12.62 | 7.78 | 2.2 | _ | 8.438 | 372 | 4.182 | 13.085 | 1.23E-07 | 0.000016 | 1.37 | 356.1 | 2.57 | 11800000 | 1.30 | 3.3 | 1.65 | 0.00408 |
| **26** | 13.11 | 7.78 | 2.1 | _ | 8.805 | 494 | 4.305 | 13.607 | 4.74E-09 | 0.00000816 | 1.13 | 490.6 | 2.18 | 21500000 | 1.16 | 3.01 | 1.62 | 0.0024 |
| **27** | 13.6 | 7.79 | 2 | _ | 9.172 | 655 | 4.428 | 14.128 | 1.82E-09 | 0.00000417 | 0.89 | 675.9 | 1.77 | 39100000 | 0.969 | 2.77 | 1.57 | 0.00142 |
| **28** | 14.09 | 7.79 | 2 | _ | 9.539 | 870 | 4.551 | 14.65 | 6.97E-09 | 0.00000213 | 0.65 | 931.2 | 1.35 | 71300000 | 0.909 | 2.4 | 1.52 | 0.000837 |
| **29** | 14.58 | 7.79 | 1.9 | _ | 9.908 | 1150 | 4.672 | 15.171 | 2.67E-09 | 0.00000108 | 0.5 | 1283 | 0.94 | 130000000 | 0.798 | 2.08 | 1.46 | 0.000494 |
| **30** | 15.07 | 7.79 | 1.8 | _ | 10.274 | 1530 | 4.796 | 15.693 | 1.02E-09 | 5.52E-07 | 0.05 | 1767 | 0.57 | 237000000 | 0.692 | 1.76 | 1.39 | 0.000291 |
| **31** | 15.57 | 7.79 | 1.8 | _ | 10.649 | 2040 | 4.921 | 16.214 | 3.9E-09 | 0.00000028 | 0.05 | 2435 | 0.27 | 432000000 | 0.596 | 1.47 | 1.31 | 0.000168 |
| **32** | 16.06 | 7.79 | 1.7 | _ | 11.017 | 2700 | 5.043 | 16.736 | 1.49E-09 | 1.42E-07 | 0.5 | 3355 | 0.09 | 787000000 | 0.505 | 1.19 | 1.24 | 0.0000991 |
| **33** | 16.55 | 7.79 | 1.7 | _ | 11.751 | 4760 | 5.289 | 17.257 | 5.67E-09 | 7.23E-08 | 0.5 | 4622 | 0 | 1430000000 | 0.37 | 0.964 | 1.15 | 0.0000587 |
| **34** | 17.04 | 7.79 | 1.6 | _ | 12.118 | 6320 | 5.412 | 17.779 | 1.6E-09 | 3.66E-08 | 0.5 | 6367 | -0.03 | 2610000000 | 0.42 | 1.35 | 1.3 | 0.0000462 |
| **35** | 17.53 | 7.79 | 1.6 | _ | 12.485 | 8390 | 5.535 | 18.3 | 8.22E-09 | 1.86E-08 | 0.5 | 8772 | -0.04 | 4760000000 | 0.355 | 1.09 | 1.22 | 0.0000273 |
| **36** | 18.02 | 7.79 | 1.5 | _ | 12.853 | 1110 | 5.657 | 18.822 | 3.13E-09 | 9.39E-08 | 0.5 | 12090 | -0.05 | 8680000000 | 0.157 | 0.917 | 1.1 | 0.0000162 |
| **37** | 18.51 | 7.79 | 1.5 | _ | 12.853 | 1110 | 5.657 | 19.343 | 1.9E-09 | 4.75E-08 | 0.5 | 6550 | -0.05 | 1E + 10 | 0.0187 | 1.04 | 1.08 | 0.00000609 |
| **38** | 19.00 | 7.79 | 1.5 | _ | 13.218 | 1480 | 5.782 | 19.856 | 4.51E-09 | 2.4E-08 | 0.5 | 22940 | -0.05 | 1E + 10 | 0.0128 | 1.23 | 1.15 | 0.00000197 |
| **39** | 19.45 | 7.79 | 1.4 | _ | 13.586 | 1960 | 5.904 | 20.386 | 1.71E-09 | 1.21E-08 | 0.5 | 31600 | -0.05 | 1E + 10 | 0.00119 | 1.42 | 1.22 | 6.39E-07 |
| **40** | 19.99 | 7.79 | 1.4 | _ | 13.962 | 2610 | 6.028 | 20.908 | 6.49E-09 | 6.12E-08 | 0.5 | 43540 | -0.05 | 1E + 10 | 0.0245 | 1.59 | 1.31 | 2.02E-07 |

the scales that were not reported previously were utilized the eqns. 1-10 of this study. Some of the other indices were examined and the best results and equations for extending the physicochemical data were chosen. All graphing operations were performed using the Microsoft Office Excel-2003 program. The data of octanol-water partitioning coefficients and total biodegradation (log $K_{ow}$ and $TB_d$ (g/h), respectively), median lethal concentration 50 ($LC_{50}$) and water solubility ($S_w$, mg $L^{-1}$ /25 ºC) were calculated by EPI-suit v3.12 package[41].

## RESULTS AND DISCUSSION

It is reported and accepted that the toxicity property of organic compounds can be predicted on the basis of the log $K_{ow}$[38]. The QSAR results hold true for quite a lot of organic compounds, the most commonly used for test organism, follows this standard pattern[39]. Biodegradation is usually quantified by incubating a chemical compound in presence of a degrader and measuring some factors like oxygen or production of $CO_2$. The biodegradation QSAR studies demonstrate that microbial biosensors are a viable alternative means of reporting on potential biotransformation. However, a few chemicals are tested and large data sets for different chemicals need for QSAR modeling[30-39]. This study shows the structural relationship between the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$ (**1-40**) as a important class of the saturated hydrocarbons as molecular descriptions of structure-property relationship studies, logarithm of calculated octanol-water partitioning coefficients and total biodegradation, log $K_{ow}$ and $TB_d$ (mol/h and g/h), respectively, median lethal concentration 50 ($LC_{50}$), water solubility ($S_w$, mg $L^{-1}$/ 25 ºC), HLC (in atm $m^3$/mol), $EC_{50}$ green algae (in ppm), log [BCF], bio half-life, log [BAF], $K_{oc}$, fugacity in air and soil, reaction air (kg/h) and advection per cent sediment in $C_1$-$C_{40}$ compounds **1-40**.

An $LC_{50}$ value is the concentration of a material in air that will kill 50 % of the test subjects (animals, typically mice or rats) when administered as a single exposure (typically 1 or 4 h). Also called the median lethal concentration and lethal concentration 50, this value gives an idea of the relative acute toxicity of an in halable material. Other variants that are occasionally used are the $LC_{25}$ and $LC_{75}$ which refer to the lethal concentration that kills 25 and 75 % of test subjects, respectively. Both $LC_{50}$ and $LD_{50}$ values state the animal used in the test. This is important because animal toxicity studies do not necessarily extrapolate (extend) to humans. For example, dioxins are highly toxic to guinea pigs and ducklings at extremely low levels, but have never been conclusively linked to a single human death even at very high levels of acute (short term) exposure. However, it is best to be on the safe side when evaluating animal toxicity studies and assume that most chemicals that are toxic to animals are toxic to humans. Typical units for $LC_{50}$ values are parts/ million (ppm) of material in air, micrograms ($10^{-6}$ = 0.000001 g)/L of air and mg ($10^{-3}$ = 0.001 g)/cubic meter of air[36]. There are two types of the fishes that were often considered in the more of veterinary studies: *Carassius auratus* (Gold fish) and *Perca fluviatilis* (the sea fish).

The biological half-life of water in a human is about 7-10 days. It can be altered by behaviour. Drinking large amounts of alcohol will reduce the biological half-life of water in the body. This has been used to decontaminate humans who are internally contaminated with tritiated water $T_2O$ (tritium). Drinking the same amount of water would have a similar effect, but many would find it difficult to drink a large volume of water. The basis of this decontamination method is to increase the rate at which the water in the body is replaced with new water.

An effective dose in pharmacology is the amount of drug that produces a therapeutic response in 50 % of the people taking it, sometimes also called ED-50. In radiation protection it is an estimation of the stochastic effect that a non-uniform radiation dose has on a human. In pharmacology, effective dose is the median dose that produces the desired effect of a drug. The effective dose is often determined based on analyzing the dose-response relationship specific to the drug. The dosage that produces a desired effect in half the test population is referred to as the ED-50, for "effective dose, 50 %".

Bioconcentration factor (BCF) used to describe the accumulation of chemicals in organisms, primarily aquatic, that live in contaminated environments. According to EPA guidelines, "the bioconcentration factor is defined as the ratio of chemical concentration in the organism to that in surrounding water. Bioconcentration occurs through uptake and retention of a substance from water only, through gill membranes or other external body surfaces. In the context of setting exposure criteria it is generally understood that the terms "BCF" and "steady-state BCF"' are synonymous. A steady-state condition occurs when the organism is exposed for a sufficient length of time that the ratio does not change substantially." Bioconcentration factors (BCFs) are used to relate pollutant residues in aquatic organisms to the pollutant concentration in ambient waters[42-50]. The bioconcentration factor (BCF) is related to biomagnification efects. Many chemical compounds, especially those with a hydrophobic component partition easily into the lipids and lipid membranes of organisms and bioaccumulate. If the compounds are not metabolized as fast as they are consumed, there can be significant magnification of potential toxicological effects up the food chain. The concern about bioaccumulation and biomagnification comes mainly from experience with chlorinated compounds, especially pesticides and PCBs and their deleterious effects on vulnerable species, especially birds, frogs and fish. Only minimal experimental and monitoring information has been gathered on the bioaccumulation properties of many other currently used chemical compounds. In fact, the biomagnification of many widely available chemicals has not been observed or predicted in aquatic systems. Bioconcentration factor (BCF) or bioaccumulation factor (BAF) values are based on US Environmental Protection Agency publications pursuant to Section 304(a) of the Federal Water Pollution Control Act as amended, literature values or site-specific bioconcentration data. Current EPA guidelines for the derivation of human health water quality criteria use BCFs as well[42-50].

The organic carbon adsorption coefficient (KOC), is crucial for estimating a chemical compound's mobility in soil and the prevalence of leaching from soil. The adsorption of a compound increases with an increase in organic content, clay content and surface area of the soil. The presence of a chemical compound can also be detected in ground water and inference can be made about its residence time in the soil and the degradation period before reaching the water table. The presence of continuous pores or channels in soil will increase the mobility of a chemical compound in the soil[51-55].

Fugacity is a measure of a chemical potential in the form of 'adjusted pressure.' It reflects the tendency of a substance to prefer one phase (liquid, solid or gas) over another and can be literally defined as "the tendency to flee or escape". At a fixed temperature and pressure, a homogeneous substance will have a different fugacity for each phase. The phase with the lowest fugacity will be the most favourable and will have the lowest Gibbs free energy. Fugacity has the same units as pressure (*e.g.*, atm, psia, bars, *etc.*) As well as predicting the preferred phase of a single substance, fugacity is also useful for multi-component equilibrium involving any combination of solid, liquid and gas equilibrium. It is useful as an engineering tool for predicting the final phase and reaction state of multi-component mixtures at various temperatures and pressures without doing the actual lab test. Fugacity is not a physical property of a substance; rather it is a calculated property which is intrinsically related to chemical potential. When a system approaches the ideal gaseous state (very low pressure), chemical potential approaches negative infinity, which for the purposes of mathematical modeling is undesirable. Under the same conditions, fugacity approaches the ideal pressure of the substance and the fugacity coefficient (defined below) approaches 1. Thus, fugacity is much easier to manipulate mathematically[1,42-55].

In Figs. 1-8 were shown two dimensional diagrams of the relationship between the values of log $K_{ow}$, $TB_d$, log (BCF), log (BCF), bio-half life, soil fugacity and estimated log($K_{oa}$) with number of carbon atoms in the normal hydrocarbon $C_1$-$C_{40}$ **1-40** (*n*-factor). In some figures, the logarithmic values of these amounts were demonstrated. The values of the relative structural coefficients of the logarithm of calculated octanol-water partitioning coefficients and total biodegradation, log $K_{ow}$ and $TB_d$ (mol/h and g/h), respectively, median lethal concentration 50 ($LC_{50}$), water solubility ($S_w$, mg $L^{-1}$/25 ºC), HLC (in atm $m^3$/mol), $EC_{50}$ green algae (in ppm), log [BCF], bio half-life, log [BAF], $K_{oc}$, fugacity in air and soil, reaction air (kg/h) and advection per cent sediment in $C_1$-$C_{40}$ compounds **1-40** data were shown in Table-1. Table-2 shows the eqns. (1-13) that indicate the relationships between the selected chemical and biochemical data logarithm of calculated octanol-water partitioning coefficients and total biodegradation, log $K_{ow}$ and $TB_d$ (mol/h and g/h), respectively, median lethal concentration 50 ($LC_{50}$), water solubility ($S_w$, mg $L^{-1}$/25 ºC), HLC (in atm $m^3$/mol), $EC_{50}$ green algae (in ppm), log [BCF], bio half-life, log [BAF], $K_{oc}$, fugacity in air and soil, reaction air (kg/h) and advection per cent sediment in $C_1$-$C_{40}$ compounds **1-40**.

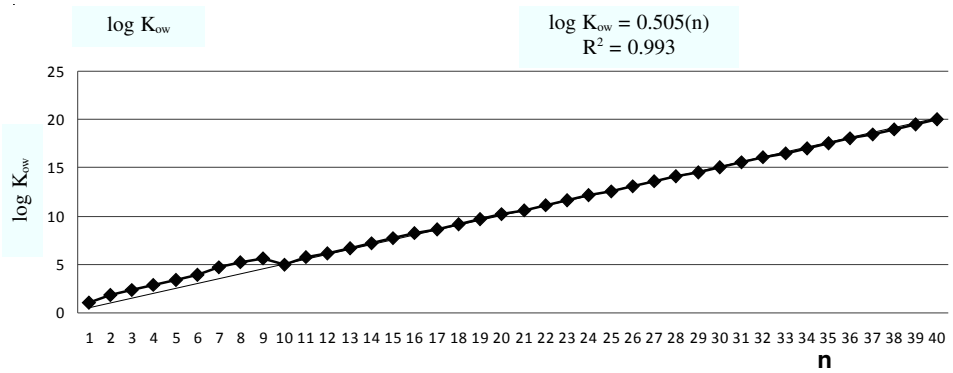


Fig. 1. Relationship between the values of log ($K_{ow}$) *versus* the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$ (**1-40**)

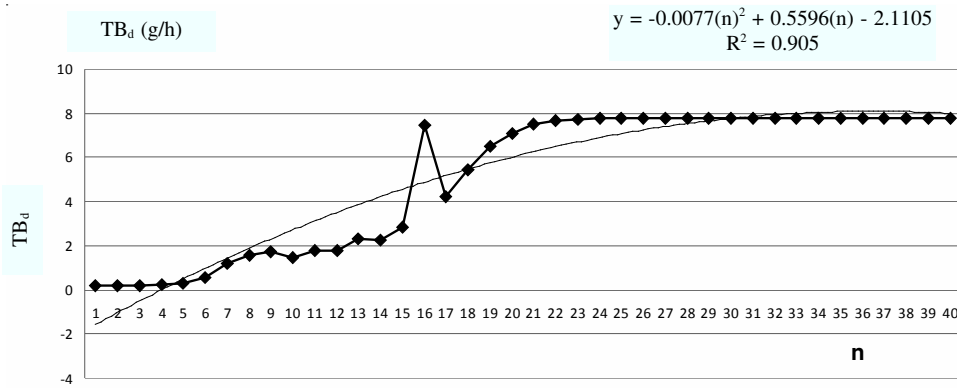


Fig. 2. Plot of the total biodegradation ($TB_d$) in g/h *versus* the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$ (**1-40**)

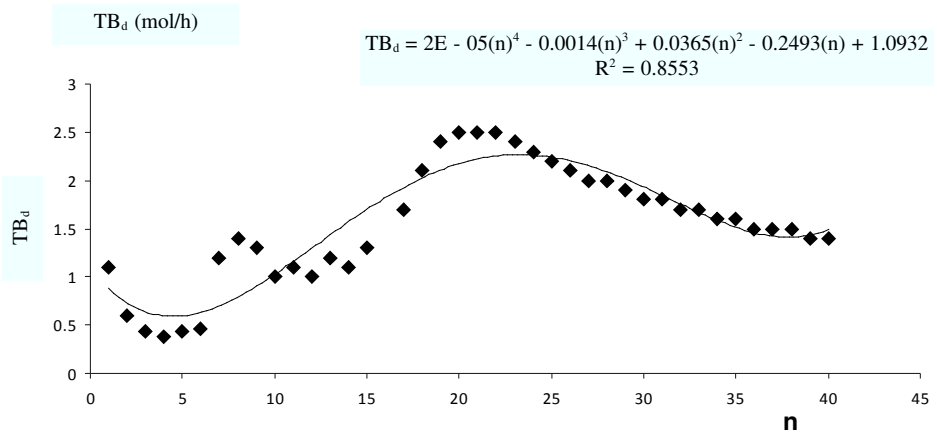


Fig. 3. Plot of the total biodegradation ($TB_d$) in mol/h *versus* the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$ (**1-40**)

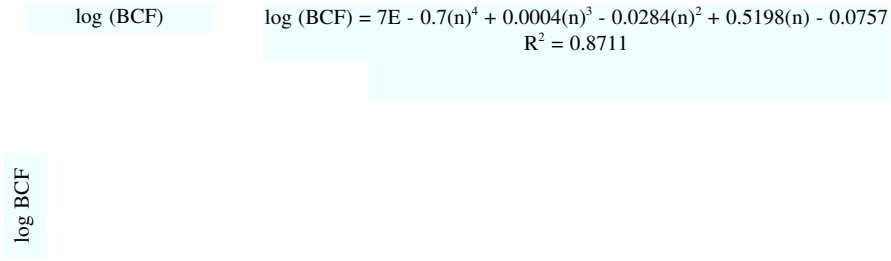


Fig. 4. Relationship between the values of log (BCF) *versus* the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$ (**1-40**)

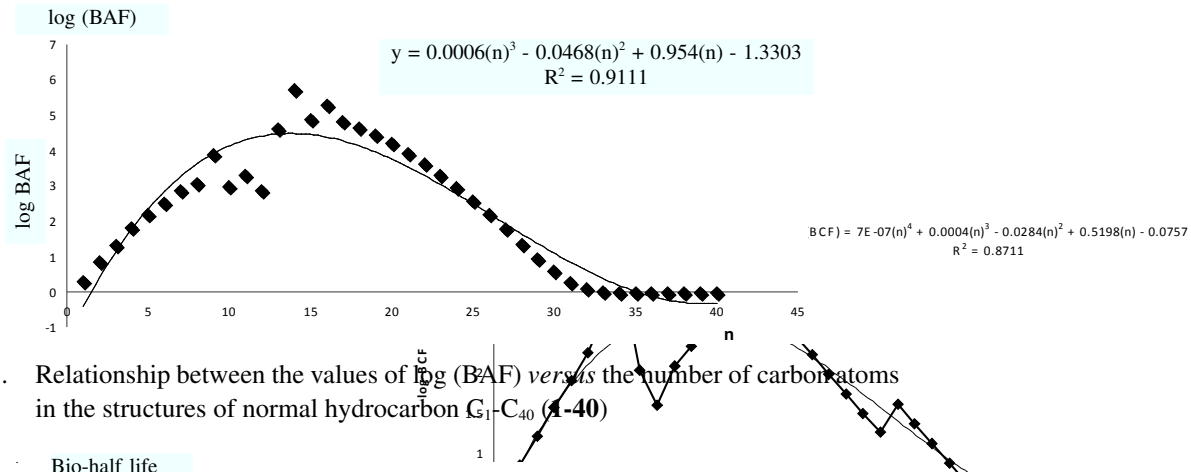


Fig. 5. Relationship between the values of log (BAF) *versus* the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$ (**1-40**)

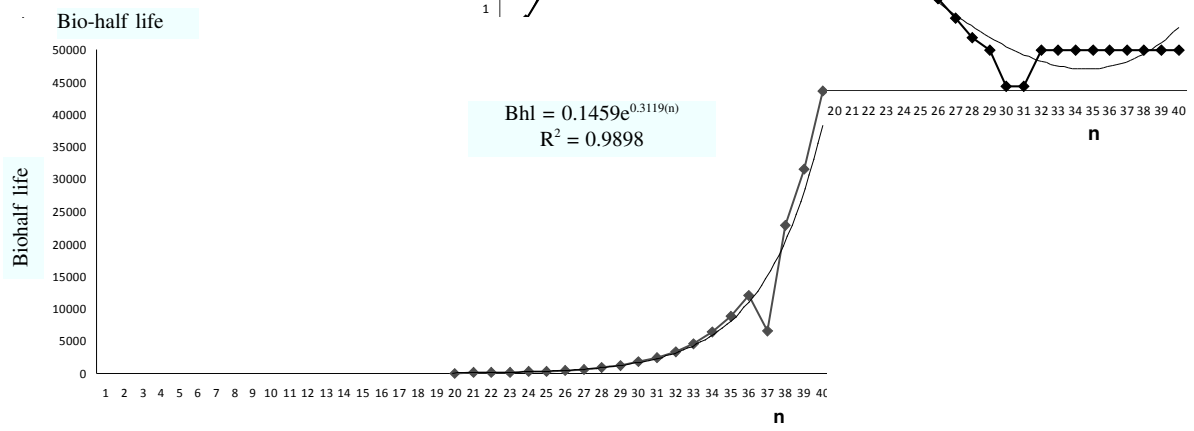


Fig. 6. Relationship between the values of bio-half life (B-hl) *versus* the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$ (**1-40**)

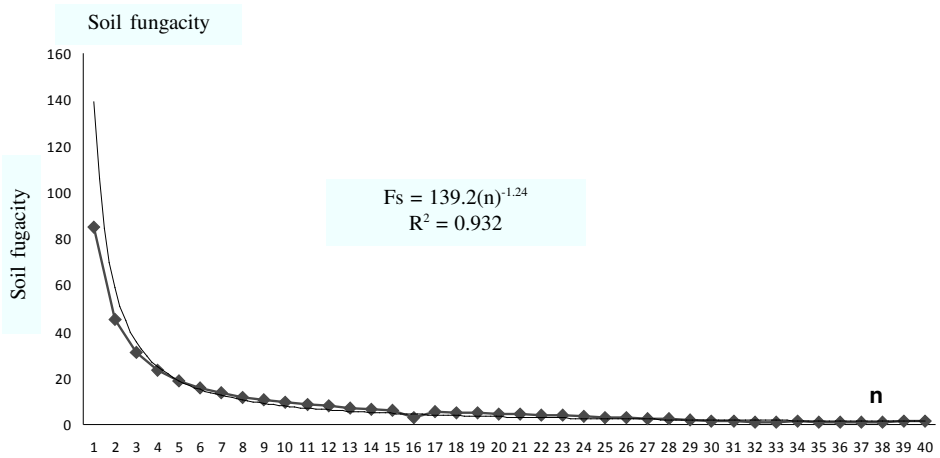


Fig. 7. Relationship between the values of soil fugacity *versus* the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$ (**1-40**)

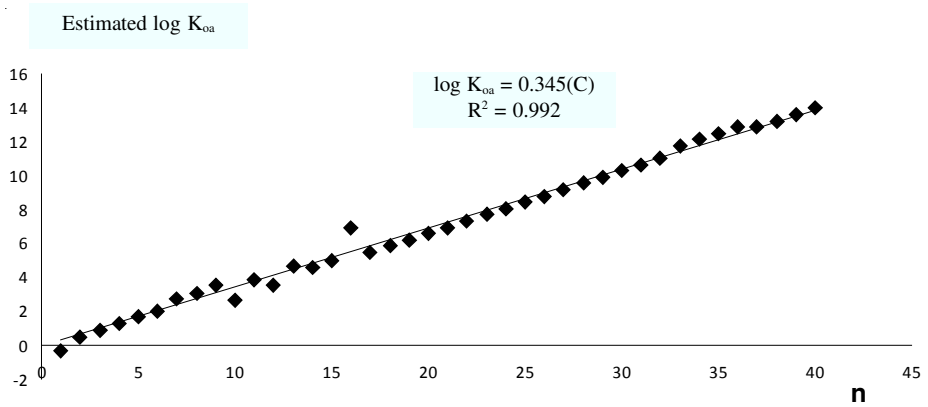


Fig. 8. Curve of the relationship between the values of estimated log ($K_{oa}$) *versus* the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$ (**1-40**)

Eqn. 1 has shown the linear relationships between log ($K_{ow}$) and number of carbon atoms in $C_1$-$C_{40}$ compounds **1-40**. The R-squared value ($R^2$) for this graph shows 0.9930. Eqn. 2 has shown a second order relationship between $TB_d$ (in g/h) and number of carbon atoms in $C_1$-$C_{40}$. The $R^2$ value for this graph shows 0.9050. The eqn. 3 is a four-order polynomial equation between $TB_d$ (in mol/h) and the number of carbon atoms in $C_1$-$C_{40}$. The $R^2$ value for this graph shows 0.8553. In eqn. 4 was demonstrated the correlations between log $S_w$ with the number of carbon atoms in $C_1$-$C_{40}$ compounds **1-40**. The value of $R^2$ value for this equation and its graph is 0.9970. The relationships between log (BCF) and the number of carbon atoms in the linear hydrocarbons $C_1$-$C_{40}$. The $R^2$ value for this graph shows 0.8711. The exponential relationship of bio-half life (B-hl) and "n" as the number of carbon atoms of **1-40** were shown in eqn. 5. In this equation the $R^2$ value is 0.9898. Eqn. 6

TABLE-2
EQNS. (1-13) THAT INDICATE THE RELATIONSHIPS BETWEEN THE SELECTED CHEMICAL AND BIOCHEMICAL DATA LOGARITHM OF CALCULATED OCTANOL-WATER PARTITIONING COEFFICIENTS AND TOTAL BIODEGRADATION, log $K_{ow}$ AND $TB_d$ (mol/h AND g/h), RESPECTIVELY, MEDIAN LETHAL CONCENTRATION 50 ($LC_{50}$), WATER SOLUBILITY ($S_w$, mg $L^{-1}$/25 ºC), HLC (IN atm $m^3$/mol), $EC_{50}$ GREEN ALGAE (IN ppm), log [BCF], BIO HALF-LIFE, log [BAF], $K_{oc}$, FUGACITY IN AIR AND SOIL, REACTION AIR (kg/h) AND ADVECTION PER CENT SEDIMENT IN $C_1$-$C_{40}$ COMPOUNDS **1-40**

| No. of Equations | Equations | $R^2$ |
|---|---|---|
| 1 | $\log K_{ow} = 0.505(n)$ | 0.9930 |
| 2 | $TB_d$ (g/h) = $-0.0077(n)^2 + 0.5596(n) - 2.1105$ | 0.9050 |
| 3 | $TB_d$ (mol/h) = $2 \times 10^{-5}(n)^4 - 0.0014(n)^3 + 0.0365(n)^2 - 0.2493(n) + 1.0932$ | 0.8553 |
| 4 | $\log S_w = -0.524(n)$ | 0.9970 |
| 5 | $\log (BCF) = 7 \times 10^{-7}(n)^4 + 0.0004(n)^3 - 0.0284(n)^2 + 0.5198(n) - 0.0757$ | 0.8711 |
| 6 | B-hl = 0.1459 exp [0.3119(n)] | 0.9898 |
| 7 | $\log [BAF] = 0.0006(n)^3 - 0.0468(n)^2 + 0.954(n) - 1.3303$ | 0.9111 |
| 8 | Soil fugacity = $139.2(n)^{-1.24}$ | 0.9320 |
| 9 | $\log K_{oa} = 0.345(n)$ | 0.9920 |
| 10 | $\log K_{aw} = 0.128(n) + 0.927$ | 0.9580 |
| 11 | $LC_{50}$ = 30.417 exp [-0.7402(n)] | 0.8108 |
| 12 | $EC_{50}$ = 76.82 exp [-0.60(n)] | 0.9800 |
| 13 | $K_{oc}$(Soil) = 4.4401 exp [0.5827(n)] | 0.9962 |

was demonstrated the relationship of log [BAF] and the number of carbon atoms in the linear hydrocarbons $C_1$-$C_{40}$ ($R^2$ = 0.9111). In eqn. 7 was shown the correlations between soil fugacity and n-values of $C_1$-$C_{40}$ compounds **1-40**. The $R^2$ value for this graph shows 0.9320. The eqns. 9 and 10 were shown the linear correlations of log $K_{oa}$ and log $K_{aw}$ with the number of carbon atoms in $C_1$-$C_{40}$ compounds **1-40**. The values of $R^2$ value for eqns. 9 and 10 and their graph are 0.9920 and 0.9580, respectively. In eqns. 11-13 were demonstrated the exponential correlations between $LC_{50}$, $EC_{50}$ and $K_{oc}$ (soil) with the number of carbon atoms in $C_1$-$C_{40}$ compounds **1-40**. The values of $R^2$ value for eqns. 11-13 and their graph are 0.8108, 0.9800 and 0.9962, respectively. By the use of descriptor "n" for linear hydrocarbons $C_1$-$C_{40}$ compounds **1-40** in the eqns. 1-13 can achieve to a good approximation for the selected chemical and biochemical data logarithm of calculated octanol-water partitioning coefficients and total biodegradation, log $K_{ow}$ and $TB_d$ (mol/h and g/h), respectively, median lethal concentration 50 ($LC_{50}$), water solubility ($S_w$, mg $L^{-1}$/ 25 ºC), HLC (in atm $m^3$/mol), $EC_{50}$ green algae (in ppm), log [BCF], bio half-life, log [BAF], $K_{oc}$, fugacity in air and soil, reaction air (kg/h) and advection per cent sediment in $C_1$-$C_{40}$ compounds **1-40**.

## Conclusion

The chemical, biochemical and environmental factors as the important points have high importance for predicting the useful models in the chemical and bio-

chemical, medical and drug structure designs studies of the compounds. The toxicity property of organic compounds can be predicted on the basis of the log $K_{ow}$. The biodegradation QSAR studies demonstrate that microbial biosensors are a viable alternative means of reporting on potential biotransformation. Here, was considered the relationship of the structural relationship between the number of carbon atoms in the structures of normal hydrocarbon $C_1$-$C_{40}$ (**1-40**) as a important class of the saturated hydrocarbons as molecular descriptions of structure-property relationship studies, logarithm of calculated octanol-water partitioning coefficients and total biodegradation, log $K_{ow}$ and $TB_d$ (mol/h and g/h), respectively, median lethal concentration 50 ($LC_{50}$), water solubility ($S_w$, mg $L^{-1}$/25 ºC), HLC (in atm m3′mol), $EC_{50}$ green algae (in ppm), log [BCF], bio half-life, log [BAF], $K_{oc}$, fugacity in air and soil, reaction air (kg/h) and advection per cent sediment in $C_1$-$C_{40}$ compounds **1-40**. The structural parameters (n) show good differences between the values of the selected calculated data which they were computed by EPI-suit v3.12 package, as important factors in chemical and biochemical studies in these compounds.

## REFERENCES


1. (a) McMurry, J. Organic Chemistry, edn. 5, Brooks/Cole: Thomson Learning (2000); (b) Hydrocarbon-Wikipedia, the free encyclopedia.htm.
2. J. Clayden, N. Greeves, S. Warren and P. Wothers, Organic Chemistry, Oxford University Press, Oxford (2000).
3. M. Silberberg, Chemistry: The Molecular Nature of Matter and Change, New York: McGraw-Hill Companies (2004).
4. W. Meierhenrich, Amino Acids and the Asymmetry of Life, Springer-Verlag (2008).
5. P.J. Hansen and P. Jurs, *J. Chem. Educ*., **65**, 574 (1988) and the literature cited therein.
6. H. Hosoya, *Bull. Chem. Soc. (Japan)*, **44**, 2332 (1971).
7. M. Randic, *Acta Chim. Slov*., **45**, 239 (1998).
8. G. Rücker and C. Rücker, *J. Chem. Inf. Cmput. Sci*., **39**, 788 (1999).
9. H. Wiener, *J. Am. Chem. Soc*., **69**, 17 (1947).
10. Y.P. Du, Y.Z. Liang, B.Y. Li and C.J. Xu, *J. Chem. Inf. Cmput. Sci*., **42**, 1128 (2002).
11. M. Randic, *J. Am. Chem. Soc*., **97**, 6609 (1975).
12. A. Sabljic and N. Trinajstic, *Acta Pharm. Ugosl*., **31**, 189 (1981).
13. P.G. Sybold, M. May and U.A. Bagal, *J. Chem. Educ*., **64**, 575 (1987).
14. L.B. Kier and L.H. Hall, Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York (1976).
15. M. Randic, *J. Math. Chem*., **7**, 155 (1991).
16. M. Randic, D. Mills and S.C. Basak, *Int. J. Quantum Chem*., **80**, 1199 (2000).
17. M. Randic, D. Plavšic and N. Lerš, *J. Chem. Inf. Cmput. Sci*., **41**, 657 (2001).
18. M. Randic and S. C. Basak, *J. Chem. Inf. Cmput. Sci*., **41**, 614 (2001).
19. M. Randic and M. Pompe, *J. Chem. Inf. Cmput. Sci*., **41**, 575 (2001).
20. L.B. Kier and L.H. Hall, Molecular Connectivity in Chemistry and Drug Research, Academic Press, New Yourk (1976).
21. L.B. Kier and L. H. Hall, Molecular Connectivity in Structure-Activity, Research Studies Press, Letchwork (1986).
22. I. Gutman and M. Randic, *Chem. Phys. Lett*., **47**, 15 (1977).
23. H. Wiener, *J. Am. Chem. Soc*., **69**, 17 (1947).
24. L.B. Kiere, *Quant. Struc.-Act. Rel*., **4**, 109 (1985).

25. L.B. Kiere and L.H. Hall, Molecular Structure Description: The Electrotopological State, Academic Press, New York (1999).
26. A. Balaban, *Chem. Phys. Lett*., **89**, 399 (1982).
27. Q.-N. Hu and Y.-Z. Liang, *Internet Electron. J. Mol. Des*., **3**, 335 (2004).
28. M. Barysz, D. Plavšic and N. Trinajstic, *Match*, **19**, 89 (1986).
29. E. Estrada, *Chem. Phys. Lett*., **336**, 284 (2000).
30. J.G. Leahy and R.R. Colwell, *Microbiol. Mol. Biol. Rev*., **54**, 305 (1990)
31. C. Hansch, A. Leo and D. Hoekman, Exploring QSAR: Hydrophobic, Electronic, Steric Constants, ACS, Washington, DC, USA (1995).
32. J.G. Bundy, A.W.J. Morriss, D.G. Durham, C.D. Campbell and G.I. Paton, *Chemosphere*, **42**, 885 (2001) and the literature cited there in.
33. A. Li and S.H. Yalkowsky, *Ind. Eng. Chem. Res*., **37**, 4470 (1998).
34. A.A. Taherpour and F. Shafiei, *J. Mol. Struct. Theochem*., **726**, 183 (2005).
35. S.D. Bolboaca and L. Jantschi, *Int. J. Mol. Sci*., **8**, 335 (2007).
36. Z. Slanina, M.-C. Chao, S.-L. Lee and I. Gutman, *J. Serb. Chem. Soc*., **62**, 211 (1997).
37. D. Plavšic, S. Nikolic, N. Trinajstic and Z. Mihalic, *J. Math. Chem*., **12**, 235 (1993).
38. A.A. Taherpour, *Fullerenes Nanotubes Carbon Nanostruc*., **15**, 405 (2007).
39. A.A. Taherpour, *Fullerenes Nanotubes Carbon Nanostruc*., **16**, 142 (2008).
40. For answering to some questions about the $LC_{50}$ was used: http://www.ilpi.com/msds/ref/lc50.html.
41. For study about the EPI-suit v3.12, See US Environmental Protection Agency site: http://www.epa.gov/epahome/docs.
42. J.W. Davis and C.L. Carpenter, *Rev. Environ. Contam. Toxicol*., **149**, 87 (1997).
43. C. Svendsen and J.M. Weeks, *Ecotoxicol. Environ. Safety*, **36**, 80 (1997).
44. C. Svendsen and J.M. Weeks, *Ecotoxicol. Environ. Safety*, **36**, 72 (1997).
45. J. Falandysz and A. Chwir, *Sci. Total Environ.*, **203**, 221 (1997).
46. T.P. Traas, R. Luttik and R.H. Jongbloed, *Ecotoxicol. Environ. Safety*, **34**, 264 (1996).
47. X. Wang, S. Harada, M. Watanabe, H. Koshikawa, K. Sato and T. Kimura, *Chemosphere*, **33**, 865 (1996).
48. A. Pant, S.C. Srivastava and S.P. Singh, *Ecotoxicol. Environ. Safety*, **32**, 87 (1995).
49. T.P. Makela and A.O. Oikari, *Chemosphere*, **31**, 3651 (1995).
50. D.G. Brown, R.P. Lanno, M.R. van den Heuvel and D.G. Dixon, *Ecotoxicol. Environ. Safety*, **30**, 302 (1995).
51. F. Vaussenat, J.Y. Bosc, M. LeBlanc and B. Canaud, *ASAIO J.*, **43**, 910-5.
52. K.M. Jop, P.D. Guiney, K.P. Christensen and E.M. Silberhorn, *Ecotoxicol. Environ. Safety*, **37**, 229 (1997).
53. D.M. Heuman, R.S. Bajaj and Q. Lin, *J. Lipid Res.*, **37**, 562 (1996).
54. D.W. Underhill, *Health Phys.*, **71**, 160 (1996).
55. S.C. Scarpitta, *Health Phys.*, **70**, 673 (1996).

(*Received*: 2 November 2009; *Accepted*: 4 June 2010) AJC-8759