

The QSPR Study of Water - Octanol Partition Coefficients for a Series of 24 Benzene Derivatives

P.G. ANOICA⁽¹⁾, EMILIA AMZOIU⁽¹⁾, L. GIUBELAN⁽²⁾

⁽¹⁾ Faculty of Pharmacy, University of Medicine and Pharmacy, Craiova, ⁽²⁾ Faculty of Medicine, University of Medicine and Pharmacy, Craiova,

ABSTRACT A series of 24 benzene derivatives are commonly used as solvents or reactants in the chemical process of synthesis. For this reason, a circumstantial study of skin permeability (or biologic membranes) to such classes of substances is enforced. This paper presents a correlation between the *log P* partition coefficients for these compounds and the descriptors characterizing the molecular structures. These descriptors were calculated using quantum *ab initio* molecular methods as Polarizable Continuum Model (PCM) for water and 1-octanol solvents which feign biological phase (the interface between water medium and cellular membrane). The partition coefficients analysis depends on the energies of solvation, molecular shape descriptors (area / volume of molecular cavity and Connolly area).

KEY WORDS benzene derivatives, partition coefficient, transmembranar model, QSPR study

Introduction

The measurement of integument skin absorption from exogenous substance is of big interest in interconnection areas of chemical applicability, as pharmaceuticals or cosmetics, for the toxicological evaluation and risk of this utility [1].

It is nearly impossible to estimate the penetrability of skin for all these new types of substance, especially the new substance or prescription, as it uses merely “in vivo” experiments. Therefore, the models “ex vivo” and “in vitro” were used for the evaluation of risk and injuriousness associated with exposure of human skin to exogenous substances [2,3].

The methodology of QSAR – type, rather QSPR [4] method Error! Bookmark not defined. is very useful for us to establish a predictive model for the biological activity of a homolog series of chemical substances used commonly solvents or mediators in a different chemical industrial process.

Although the biological activity for skin penetrability refers precisely (in an essential way) to the process of diffusion, we consider that such a diffusion succeeds at last in a partition between two immiscible phases of solvents, among which one represents the human skin.

The solvent 1-octanol feigns biologic membrane [5,6,7] in the best way due to its special structure, of hydrophilic –OH specificity and

hydrophobic (lipophilic) character owing to a long catena of carbon atoms (Figure 1).

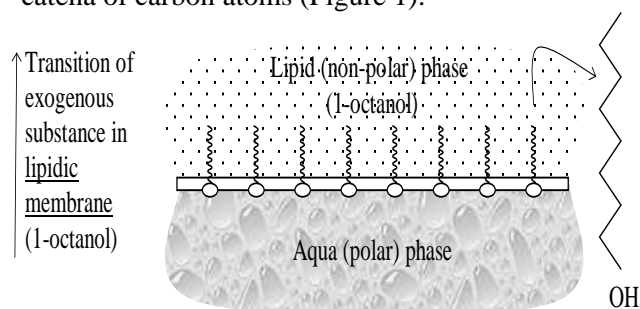


Figure 1: The biologic membrane simulation in 1-octanol model

At the interface between water and lipidic phase (the latered simulate by 1-octanol) an organized appears structure due to the polar character of water and also to the amphipathic character of the 1-octanol molecule [8].

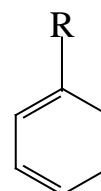


Figure 2: Substituted derivatives of benzene shown in Table 1

Such arrangement of molecules at the interface water – octanol will better simulate the partition processes in the cellular membrane.

The chemical substance studied in this article was represented by 24 substituted derivatives of benzene (figure 2). From these, the partition coefficients ($\log P$) between water and octanol

were established experimentally and were known in literature [9]. In Table 1 this 24 value was shown.

Table 1: The partition coefficients for 24 substituted derivatives of benzene

Nr.	R	Molecule	$\log P$	Nr.	R	Molecule	$\log P$
1	H-	C ₆ H ₆	2.13	13	H ₂ C=CH-	C ₆ H ₅ -CHCH ₂	2.95
2	F-	C ₆ H ₅ -F	2.27	14	HC≡C-	C ₆ H ₅ -CCH	2.53
3	Cl-	C ₆ H ₅ -Cl	2.84	15	-C ₃ H ₇	C ₆ H ₅ -C ₃ H ₇	3.68
4	Br-	C ₆ H ₅ -Br	2.99	16	H ₂ C=CH-CH ₂ -	C ₆ H ₅ -CH ₂ -CHCH ₂	3.23
5	HS-	C ₆ H ₅ -SH	2.52	17	i-C ₃ H ₇ -	i-C ₃ H ₇ -C ₆ H ₅	3.66
6	CH ₃ -	C ₆ H ₅ -CH ₃	2.69	18	n-C ₄ H ₉ -	n-C ₄ H ₉ -C ₆ H ₅	4.26
7	CH ₂ Br-	C ₆ H ₅ -CH ₂ Br	2.92	19	t-butyl	t-butyl-C ₆ H ₅	4.11
8	CH ₂ Cl-	C ₆ H ₅ -CH ₂ Cl	2.30	20	ciclopropil	ciclopropil-C ₆ H ₅	3.27
9	CH ₃ S-	C ₆ H ₅ -SCH ₃	2.74	21	ciclopentil	ciclopentil-C ₆ H ₅	4.27
10	CF ₃ -	C ₆ H ₅ -CF ₃	3.01	22	ciclohexil	ciclohexil-C ₆ H ₅	4.64
11	CF ₃ S-	C ₆ H ₅ -SCF ₃	3.57	23	C ₆ H ₅ -	C ₆ H ₅ -C ₆ H ₅	4.09
12	C ₂ H ₅ -	C ₆ H ₅ -C ₂ H ₅	3.15	24	C ₆ H ₅ CH ₂ -	C ₆ H ₅ -CH ₂ -C ₆ H ₅	4.14

A comparative study between structure and property (QSPR) assumes in fact a statistical correlation between bio-chemical properties ($\log P$ in this case) and physical or chemical properties (descriptors) which represent in fact the chemical structure of the analyzed substance. These descriptors constitute the interface correlation between bio-chemical property and chemical structure of the substance [10].

There are large numbers of these descriptors in special literature publications of last decade [11]. The reason for large numbers of descriptors is the incapacity to explain – in many cases - of the cause – effect interaction between structures and the activity or mechanism of this interaction (for ex: interaction of medicament (ligand) and receptor – in pharmacology).

In this QSPR study we analyze the structural descriptors involved in partition coefficients for benzene derivatives. Such a study allows the better knowledge of interaction between solvents and dermatome-membranes. This QSAR way helps the labor study and was carried out with success between pollutants or nocive substances and derma interaction, with hope of artificial membrane optimization [12].

For molecular descriptor calculation involved in partition process we effectuated the chemical design of the structure that was studied with MM+ Molecular Mechanics and RHF and PM3 semi-empirical method (MOPAC 7.0). The atomic coordinate obtained in this way was used at input data in quantum-molecular *ab initio* calculus RHF, PCM, STO 6G (GAMESS) about Restricted Hartree Fock approximation. In this method the atomic orbital – tip Slater – compute with 6 Gaussian function (STO – 6G) [13,14,15,16] was used.

Descriptors for molecular shape and electrostatic interaction

The solvents water and 1-octanol were in PCM model (Polarizable Continuum Model) incorporate in GAMMES-soft. Thus, the solvent is described as a macroscopic continuum medium with a specified dielectric constant. The immersion of the solute molecule inside the solvent created a specific molecular cavity. In this model the interaction between solute and solvent molecules computed such a modification of electro-magnetic field in reaction with other dielectric media.

In PCM approach were calculated the surface and volume for molecular cavity with the intersection and reunion of Van der Waals sphere of atoms in molecule – Figure 3.

For the water solvent this dielectric constant $\epsilon_r = 78.4$ (relative to vide or air) was used with maximal raze of rotating sphere $r_w = 1.385 \text{ \AA}$ and for the 1-octanol solvent $\epsilon_r = 10.3$ and $r_o = 3.25 \text{ \AA}$.

In correlation with these molecular descriptors of molecular shape in solute we evaluate the Connolly descriptors of molecular shape [17] such as CSAA – Connolly Solvent Accessible Area and CSEV - Connolly Solvent Excluded Volume expose in Figure 4.

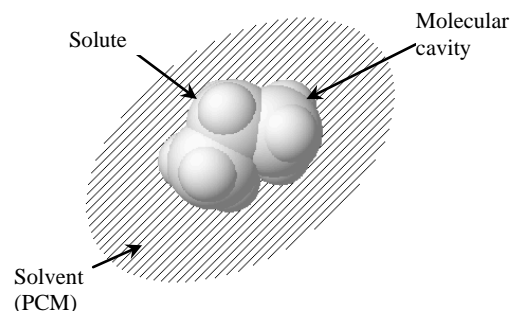


Figure 3: Molecular cavity in solute – solvent interaction model

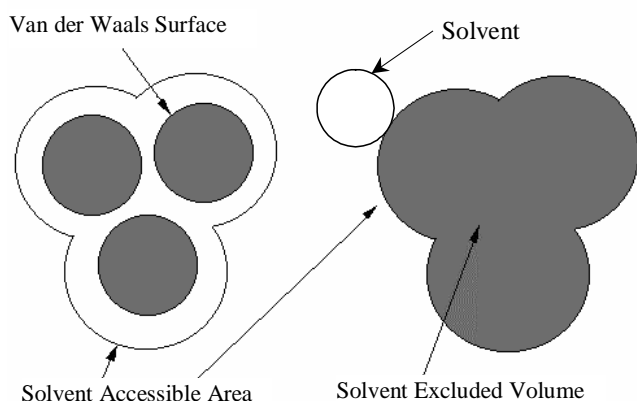


Figure 4: Connolly Solvent Accessible Area (CSAA) and Connolly Solvent Excluded Volume (CSEV)

Results and Discussion

In Table 2 the value of molecular descriptors of shape for these 24 of molecule involved was described.

We also calculated, apart from of shape descriptors, the energy of electrostatic interaction between solute and solvent with *ab initio* quantum-molecular procedure. These values depend on the partition of solute among two immiscible phases.

Table 2: Molecular shape descriptors value for the 24 benzene derivatives

Nr.	log P	CSAA	CSAA [#]	CSEV	CSEV [#]	CA	CV
1	2.13	232.453	481.432	70.805	74.213	109.703	83.510
2	2.27	239.083	491.119	74.263	77.986	113.953	86.385
3	2.84	255.698	515.450	84.871	89.023	127.229	98.645
4	2.99	263.831	527.309	90.550	94.879	132.924	104.640
5	2.52	259.492	520.844	87.328	91.609	132.387	102.729
6	2.69	262.983	525.375	87.374	92.608	131.075	100.111
7	2.92	293.926	569.043	108.830	116.671	155.075	121.793
8	2.30	286.710	559.165	102.966	110.094	149.397	116.090
9	2.74	291.838	568.098	105.990	111.748	152.501	119.610
10	3.01	280.295	549.957	98.054	104.410	143.827	109.182
11	3.57	310.590	595.142	117.229	124.178	166.943	129.191
12	3.15	291.963	565.854	105.544	113.464	152.769	116.574
13	2.95	276.319	546.316	94.661	99.237	141.371	110.606
14	2.53	273.975	545.588	88.029	93.076	136.931	105.864
15	3.68	322.621	610.523	122.700	132.790	174.593	133.717
16	3.23	312.620	596.975	114.351	123.486	166.249	128.337
17	3.66	315.517	597.930	122.558	133.357	172.148	133.147
18	4.26	353.247	655.312	139.821	151.548	196.316	150.412
19	4.11	329.335	616.486	138.489	148.857	186.396	148.893
20	3.27	306.745	586.314	112.864	122.774	164.598	127.017
21	4.27	346.534	641.780	145.293	156.329	194.418	156.196
22	4.64	369.906	673.817	163.646	178.173	214.055	172.604
23	4.09	339.269	637.668	131.694	138.044	184.394	153.636
24	4.14	372.197	682.178	153.250	164.204	213.642	172.219

CSAA – Connolly Solvent Access Area (\AA^2);

CSEV – Connolly Solvent Excluded Volume (\AA^3);

CA – Surface Area (\AA^2); CV – Cavity Volume (\AA^3)

The # signifies the value of interaction energy between solvent and solute (ours) molecule, while the other values was for water.

Table 3 summarizes the calculated value of interaction for solvent – solute energies.

The molecular descriptor values in Table 2-3 were calculated using the *ab initio* quantum-molecular procedure (RHF, STO-6G GAMESS) for water and 1-octanol solvents. For the CSAA (Connolly Accessible Surface Area) and CSEV (Connolly Solvent Excluded Volume) molecular descriptors, we used Chem 3D programs (ChemOffice 2005 Ultra).

All molecular descriptors summarized in Table 2 and 3 were imported in CODESSA file and the correlation with partition coefficients (log P) was effectuated with *multiple* linear regression technique.

The examination of the contribution of shape descriptors and solute – solvent interaction descriptors summarized in Table 4 is very suggestive

As it is shown in Table 4, all 4 combinations of two descriptors of shape have the same correlation coefficient, $R^2 = 0.88$. This aspect denotes the importance of molecular shape in the process of partition between the two non-miscible phases, due to the of interaction access of solvent - solute molecules. In this direction the solvent – solute interaction must depend on the solvent medium radius of the two molecules involved. To record this aspect, the most sensible descriptor is the Connolly Solvent Excluded Volume (CSEV). This shape descriptor indirectly describes the access of the solute molecule to the solvent.

Also, we present in table 5 the results of the correlation in multiple linear regressions to mark the contribution of these shape descriptors.

In fact the electrostatic interaction (EI) alone does not have a significant contribution in the partition process. In this case the correlation

coefficient is insignificant ($R^2 = 0.3832$). Therefore, the single descriptor TFES contribution is irrelevant ($R^2 = 0.043543$).

Actually, this descriptor summarizes many interactions ($TFES = IES + TI = EI + PCE + RFE + IES$). Individually grouped, this specifies a very

good correlation with a great coefficient (0.9161)^[12].

Similarly, TI consists of many specific interaction ($TI = EI + PCE + RFE$). If that were individually grouped the correlational coefficient would be better ($R^2 = 0.906$).

Table 3: Solute – solvent interaction energy (kcal/mol)

Nr	TFES	TFES [#]	RFE	RFE [#]	PCE	PCE [#]	EI	EI [#]	TI	TI [#]	IES
1	-144392.68	-144396.59	1.96	1.96	14.79	10.75	-0.94	-0.79	15.81	11.92	-144408.49
2	-206127.92	-206131.97	1.74	1.74	15.32	11.12	-1.37	-1.18	15.69	11.68	-206143.61
3	-431417.48	-431422.03	1.78	1.78	16.70	11.96	-1.64	-1.41	16.84	12.32	-431434.31
4	-1749668.2	-1749671.9	1.70	1.70	17.26	12.29	-10.32	-8.64	8.64	5.34	-1749676.8
5	-393011.82	-393016.34	2.10	2.08	17.51	12.59	-2.98	-2.48	16.63	12.19	-393028.45
6	-168836.15	-168840.84	2.35	2.32	17.82	13.01	-1.02	-0.85	19.15	14.49	-168855.30
7	-1774076.7	-1774081.7	2.09	2.05	20.40	14.63	-5.70	-4.84	16.79	11.83	-1774093.4
8	-455859.70	-455864.9	2.16	2.13	19.83	14.29	-2.78	-2.34	19.21	14.08	-455878.91
9	-417456.73	-417462.01	2.51	2.48	20.37	14.74	-2.94	-2.47	19.94	14.75	-417476.67
10	-354043.11	-354048.14	1.71	1.69	19.43	14.13	-2.36	-2.02	18.78	13.80	-354061.89
11	-602655.71	-602661.66	1.93	1.90	22.20	16.01	-2.07	-1.75	22.06	16.16	-602677.78
12	-193277.02	-193282.5	2.75	2.69	20.90	15.31	-1.04	-0.85	22.61	17.16	-193299.64
13	-192509.12	-192514.17	2.39	2.35	19.03	13.82	-1.27	-1.05	20.14	15.13	-192529.26
14	-191748.54	-191753.40	2.12	2.13	18.18	13.10	-1.77	-1.50	18.54	13.72	-191767.08
15	-217718.18	-217724.49	3.12	3.00	24.00	17.63	-1.08	-0.88	26.03	19.75	-217744.22
16	-216947.27	-216953.23	2.84	2.78	22.58	16.47	-1.24	-1.02	24.18	18.24	-216971.44
17	-217717.99	-217724.15	3.13	3.06	23.66	17.38	-1.11	-0.90	25.68	19.55	-217743.67
18	-242159.30	-242166.34	3.52	3.37	27.08	19.93	-1.23	-0.93	29.37	22.36	-242188.67
19	-242154.38	-242161.07	3.32	3.23	25.67	18.87	-1.23	-1.00	27.76	21.10	-242182.14
20	-216952.35	-216958.24	2.89	2.87	22.43	16.40	-1.17	-0.98	24.15	18.29	-216976.00
21	-265873.59	-265880.56	3.61	3.57	26.80	19.72	-1.07	-0.89	29.34	22.40	-265902.93
22	-290322.14	-290329.83	3.88	3.78	29.63	21.85	-1.14	-0.92	32.38	24.71	-290354.52
23	-288069.37	-288075.92	3.05	3.06	24.67	17.85	-1.83	-1.52	25.89	19.38	-288095.26
24	-312510.50	-312518.12	3.40	3.33	28.86	21.00	-1.82	-1.46	30.44	22.87	-312540.94

TFES – Total Free Energy in Solvent ($TFES = IES + TI$);
 RFE – Repulsion Free Energy;
 PCE – Pierotti Cavity Energy;
 TI – Total Energy ($TI = RFE + PCE + EI$);
 IES – Internal Energy in Solvent

Table 4: Multiple linear regression: $\log P = a_0 + a_1 \cdot Desc.1 + a_2 \cdot Desc.2$

R ²	Desc.							
	CSAA	CSAA [#]	CSEV	CSEV [#]	CA	CA [#]	CV	CV [#]
0.882	x	x						
0.896			x	x				
0.878					x	x		
0.876							x	x

“Desc.” is one of the descriptors: CSAA - Connolly Solvent Access Area (Å²), CSEV - Connolly Solvent Excluded Volume (Å³), CA – Surface Area (Å²) and The sign # symbolise values computed for 1-octanol solvent (non mark – for water solvent); R² is correlational coefficient.

Moreover, if TFES descriptor (which represents the total energy of salvation in the interaction between solute and solvent) was associated with shape descriptors, an optimization in previous correlations would result with an increase of regressional coefficients S and R² = 0.88÷0.90.

1. TFES + CA:

$$\log P = -0.361 + 0.0236 \cdot CA - 0.00600 \cdot TFES + 0.00600 \cdot TFES^{\#}$$

$$S = 0.2646; R^2 = 0.883$$

2. TFES + CV:

$$\log P = -0.025 + 0.0276 \cdot CV - 0.00619 \cdot TFES + 0.00619 \cdot TFES^{\#}$$

$$S = 0.2663 \quad R^2 = 0.881$$

3. TFES + CSEV:

$$\log P = 0.146 + 0.0866 \cdot CSEV - 0.0528 \cdot CSEV^{\#} - 0.00951 \cdot TFES$$

$$S = 0.2433; R^2 = 0.906$$

4. TFES + CSAA:

$$\log P = 4.18 - 0.00351 \cdot TFES + 0.00351 \cdot TFES^{\#} + 0.0730 \cdot CSAA - 0.0393 \cdot CSAA^{\#}$$

$$S = 0.2698; R^2 = 0.884$$

This result is in concordance with Tanaka results [17] for shape descriptors CSAA and salvation energies in their case.

Table 5: Multiple linear regression: $\log P = a_0 + \sum (a_i \text{ Desc.}_i)$

R ²	Desc.										
	TI	TI [#]	EI	EI [#]	PCE	PCE [#]	RFE	RFE [#]	TFES	TFES [#]	IES
0.383			x	x							
0.868	x	x									
0.043									x	x	
0.862	x	x									x
0.906			x	x	x	x	x	x			
0.916			x	x	x	x	x	x			x

Now "Desc." is one of the descriptors: EI – Electrostatic Interaction; PCE – Pierroti Cavity Energy; RFE - Repulsion Free Energy; TFES – Total Free Energy in Solvent; IES – Internal Energy in Solvent, TI = EI + PCE + RFE; TFES = IES + TI = EI + PCE + RFE + IES

The # signifies the value of interaction energy between solvent and solute (ours) molecule, while the other values were for water.

Conclusion

In this paper a possible correlation between partition coefficient ($\log P$) and 24 structural descriptors for benzene derivatives was studied.

1. The shape molecular descriptors for benzene derivatives (CSAA, CSEV, CA, CV) have the same correlation coefficient $R^2 = 0.88$. These results have shown the role played by the shape of the molecules in the partition process between the two non-miscible phases. This fact is correct because the interaction solute-solvent depends on the access of the molecular solvent to the molecular solute.

2. The single electrostatic interaction (EI) does not contribute significantly to the partition process. Now the correlational coefficient has only the value $R^2 = 0.3832$. Also, the contribution of TFES descriptor is irrelevant ($R^2 = 0.043543$). The electrostatic interaction in detail $TFES = TI + IES = EI + PCE + RFE + IES$ shows a correlation coefficient very close to unity (0.9161). Similarly, the TI descriptor consists of many specific interactions ($TI = EI + PCE + RFE$) very well correlated in separate mode ($R^2 = 0.906$).

3. The TFES descriptor (it represents a total energy of salvation in solute- solvent interaction) in association with shape descriptors calculate a very relevant correlational coefficient $0.88 \div 0.90$.

References

1. E. Merică, "Tehnologia produselor cosmetice", Editura Kolus, Iași, p. 225-226
2. Understanding variation in partition coefficient", Kd, values, USA Environmental Protection Agency, Office of Air and Radiation, I, **1999**, 212
3. G.S.D. Ayton, S. Bardenhagen, P. McMurtry, D. Sulsky and G.A. Voth, *IBM J. Res. & Dev.*, **2001**, 45, 3/4, 417-420
4. C. Amovilli and B. Mennucci, *J. Phys. Chem. B*, **1997**, 101, 1051-1054
5. B. Mennucci and J. Tomasi, *J. Chem. Phys.*, **1997**, 106, 5151-5155
6. C.W. Wimley, T.P. Creamer and S.H. White, *Biochemistry*, **1996**, 35, 5109-5111
7. C.W. Wimley, S.H. White, *Nat. Struct. Biol.*, **1996**, 10/3, 842-845
8. H. Chuman, A. Mori and H. Tanaka, *Analytical Sciences, Japan Society for Analytical Chemistry*, **2002**, 18/9, 1015-1018
9. L. Tarko, *Rev. Chim. (Bucharest)*, **2007**, 58/2, 191
10. M. Karelson, "Molecular Descriptors in QSAR/QSPR studies", Wiley Intersc., New York, **2000**
11. C.T. Zawodzinski, M. Eikerling, L. Pratt, A. Redondo, T. Rockward, M. Hickner and J. McGrath, "Membranes for Operation above 100, Materials Science and Technology Division", Los Alamos National Laboratory, Report, **2002**
12. M.W. Schmidt, K.K. Baldrige, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S.J. Su, T.L. Windus, M. Dupuis and J.A. Montgomery, "The General Atomic and Molecular Electronic Structure System.", *J. Comput. Chem.*, **1993**, 14, 1347.
13. R.G. Parr and W. Yang, "Density functional Theory of Atoms and Molecules", New York, Oxford University Press, **1989**
14. R.G. Pearson, Absolute electronegativity and hardness: applications to organic chemistry, *J. Org. Chem.*, **1989**, 54, 1423.
15. R.G. Parr and Z. Zhou, "Absolute Hardness: Unifying concept for identifying shells and subshells in nuclei, atoms, molecules, and metallic clusters", *Chem. Res.*, **1993**, 26, 256
16. Z. Zhou and R.G. Parr, "Activation hardness: new index for describing the orientation of electrophilic aromatic substitution", *J. Am. Chem. Soc.*, **1990**, 112, 5720
17. M.L. Connolly, "Solvent - accessible surfaces of proteins and nucleic acids", *Science*, **1983**, 221, 709-713

Corresponding Adress: P.G. Anoaica PhD, Faculty of Pharmacy, University of Medicine and Pharmacy, Craiova gabriel@umfcv.ro