QSPR Correlation of Physico-chemical Descriptors with the Molecular Surface Area and Rf of Ten Polyphenolic Compounds, Separated from Vegetal Extracts by TLC

RENATA MARIA VARUT[1]*, LUCIANA TEODORA ROTARU[2], MARIUS CIPRIAN VARUT[1]
[1]University of medicine and pharmacy, Pharmacy Department I, 2-4 Petru Rares, 200349, Craiova, Romania
[2]University County Hospital Craiova, Emergency & First Aid Department, University of Medicine and Pharmacy of Craiova, Emergency Department - SMURD, 1 Tabaci Str., 200642, Craiova, Romania

Polyphenols are important constituents of plants, being demonstrated that long term consumption of them have potential health benefits as antioxidant. In the experimental part of the study we separated ten polyphenolic compounds by thin layer chromatography (TLC) from hydroalcoholic hypoglycemic plant extracts (Ericaceae family). In the theoretical study part we consider important to characterize the structural properties of the substances using molecular mechanics studies. The selected physico-chemical descriptors were: HOMO and LUMO orbitals frontier, water solubility, partition coefficient, dipole moment, heat of formation, stabilization energy of the molecule and number of proton donor groups. We realized a comparative QSPR regression correlation study of molecular surface area and Rf with the selected desriptors. From studies we observed the relationship between the two parameters Rf and MSA, which correlates very well with common descriptors.

Keywords: antioxidants, TLC, QSPR

Polyphenols are secondary metabolites of plants, being involved in defense against ultraviolet radiation or aggression by pathogens. Long term consumption of diets rich in plant polyphenols have positive antioxidant effects on the human body, reducing de development risk of cancers, cardiovascular diseases, diabetes, osteoporosis and neurodegenerative diseases. The most important classes are classified in phenolic acids, flavonoids, stilbenes and lignans. All plant phenolic compounds arise from a common intermediate, phenylalanine, or a close precursor, shikimic acid, having a similar basic skeleton [1].

Researchers are often encouraged to use theoretical data in their studies in order to support their experimental results. Therefore, many publications in different research areas contain substantial quantum chemical calculations. Formulas used for molecular modeling programs allow molecular geometry optimization and calculation of physico-chemical parameters, that requires time and material resources to be experimentally determined. The information thus obtained are now used in various industries, including the pharmaceutical industry to improve the therapeutic activity of the drugs, or to contribute to the design and synthesis of new compounds with improved pharmaceutical properties [2,3].

In the present study were calculated different descriptors that can contribute to thin layer chromatographic separation of ten polyphenolic compounds, using molecular mechanics studies.

We also performed simple and multiple regressional correlations of molecular surface area (MSA) and Rf with physico-chemical and structural properties of the analyzed compounds.

Experimental part
Thin Layer Chromatography (TLC) analysis
The polyphenolic compounds were identified using aluminum backed silica gel GF_{254} Merck TLC plates (20×20cm, activated for 60 min. at 105°C) as the stationary phase and ethyl acetate-formic acid -water (80:8:12, v/v/v) as mobile phase [4]. Analyzed samples were 20% ethanol solutions. To obtain the hydrolyzed solutions, 3 mL of each extract were refluxed with 3 mL of 10% HCl solution for 30 min in an electric bath. After cooling they were extracted twice with 7 mL of diethyl ether. Apolar layers were combined, filtered through anhydrous sodium sulphate and brought to the residue which was solubilized in 3 mL of absolute methanol. 0.1 mg/mL methanol solutions of rutin, hyperoside, isoquercitrin, kaempferol, luteolin, chlorogenic acid (Roth, Germany), quercetol, ferulic acid (Sigma-Aldrich, Germany) and caffeic acid (Merck, Germany) were used as reference solutions. 10 µl of the test samples/reference solutions were applied to the starting line, the tapes having 1 cm width with 1.5 cm between them. The plates were developed over a path of 8 cm. Revelation was done by sputtering 10% ethanol solution of diphenylboriloxyethylendiamine (DFBOA, Sigma-Aldrich, Germany). The plates were examined in UV light (λ 365 nm) and photographed before and after the revelation, using an Epson Photo PC 3000 Z Mod G 790A and Camag Reprostar 3 System [5].

Descriptors tested for having an important role in migration differences of the selected compounds on chromatographic plate are: HOMO and LUMO orbitals frontier, water solubility, partition coefficient, dipole moment, heat of formation, stabilization energy of the molecule and number of proton donor groups. The water solubility, partition coefficient and the number of hydrogen donor groups were taken from literature [6], the stabilization energy of the molecules was calculated using the program HyperChem 8, and the other descriptors value was obtained using software package MOPAC 2016.

For the QSPR study we browse the following steps:
- molecular modeling of chemical structures for the ten substances using HyperChem program 8 [7] (semi-empirical optimization PM 3/ SCF) [8].
- quantum molecular calculations of the molecular geometries using MOPAC 2016. The output data contains

physico-chemical information about the selected substances.

- correlation between molecular descriptors and MSA/Rf, using Regression Excel function from Microsoft Office package.

Results and discussions

Thin layer chromatography is a planar chromatography, wherein the compounds of a mixture are separated because of their solubility in a particular solvent mixture and thanks to their interactions (London forces, interactions dipol-dipole, intermolecular hydrogen bonds) with the stationary phase (the chromatographic planar surface) [9].

In normal phase TLC chromatography, the stationary phase (silica gel) is more polar than the mobile phase. To quantify the results, the distance traveled by the substance is divided to the total distance traveled by the mobile phase. (is not allowed to reach the end of the stationary phase.) This ratio is called the retention factor or Rf Value (Distance from Baseline travelled by Solute/Distance from Baseline travelled by Solvent (Solvent Front)) [10]. The main retention mechanism between the molecule and the two phases involves wan der Waals interactions and hydrogen bonds.

MSA is a molecular descriptors characterizing the geometry of the molecule, which plays an essential role in interaction with other molecules [11]. The COSMO procedure for calculating it generates a conducting polygonal surface around the system (ion or molecule), at the van der Waals' distance, accesible to the polar groups of solvent and stationary phase [12].
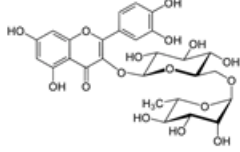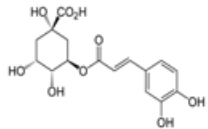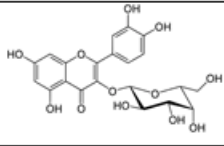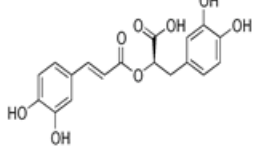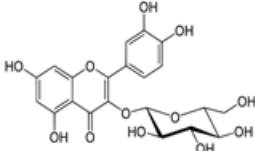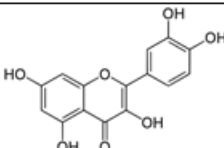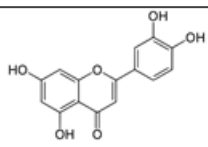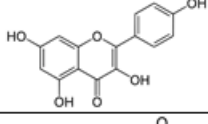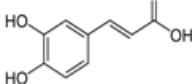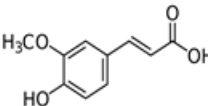
No.	Polyphenol	Rf	MSA (A^2)	
1	Rutoside	0.2	474.95	
2	Chlorogenic acid	0.38	342.5	
3	Hyperoside	0.43	382.58	
4	Rosmarinic acid	0.45	366.58	
5	Isoquercitrin	0.5	392.11	
6	Quercetol	0.64	286.84	
7	Luteol	0.66	279.58	
8	Kaempferol	0.69	278.62	
9	Caffeic acid	0.92	204.74	
10	Ferulic acid	0.95	222.32	

Table 1
THE Rf VALUES OF THE STUDIED POLYPHENOLS

No.	ΔE (eV)	S (g/l)	Log P	μ (D)	ΔH_f (kcal/mol)	E_{st} (kcal/mol)	H+ donor groups
1.	7.817	0.125	1.76	1.677	-636.70	-7455.99	10
2.	8.077	404.7	-0.356	5.802	-369.70	-4558.46	6
3.	7.827	27.82	1.75	3.744	-446.07	-5769.28	8
4.	7.708	1.425	1.7	4.158	-289.00	-4669.58	5
5.	7.793	27.82	1.75	6.098	-450.17	-5502.06	8
6.	7.961	2.472	2.075	5.524	-232.50	-3726.01	5
7.	7.991	0.387	2.401	4.201	-196.65	-3627.57	4
8.	8.121	1.191	2.05	4.449	-189.83	-3623.33	4
9.	8.029	54.07	1.424	3.762	-139.98	-2334.14	3
10.	7.885	5.97	1.641	4.079	-134.26	-2600.12	2

The MSA of studied substances play an important role in migration distance, a bigger value of this descriptor correlated with a high polarity of the molecule means a greater interaction with the stationary phase, which is reflected in a small distance of migration (low Rf). From table 1 we can see that the value of Rf varies inversely with the molecular surface area.

The difference in energy between HOMO and LUMO levels (ΔE = ELUMO - EHOMO) is an important chemical molecular descriptor which explains the stability of the molecule, a low value indicating that the molecule is highly reactive [13, 14].

The partition coefficient (Log P) express the distribution of the molecule between two immiscible liquids (octanol/ water) and has not a good correlation with Rf, since the distribution of polyphenols is between two polar phases [15].

Another important parameter of the molecular quantum chemical calculations is the electric dipole moment (μ), reflecting partial separation of electric charge in the molecule. The molecular descriptor is also a predictor of the chemical reactivity of the molecules, being a measure of the polarity of the molecular system.

The enthalpy of formation (ΔH_f) is the amount of heat released or absorbed during the formation of the chemical substances, a small value reflecting a stable substance [16].

Table 3
BEST CORRELATIONS BETWEEN Rf/MSA AND DESCRIPTORS; D_i- DESCRIPTOR

No. D_i	Rf		MSA (A^2)		Descriptors							
	R^2	F	R^2	F	ΔE	S	ΔH_f	Log P	μ	Est	H+ donor groups
3	*0.982*	*106.7*	0.983	113.71		X	X			X	
3	0.975	78.16	0.985	130.21		X		X		X	
3	0.948	36.48	*0.986*	*136.51*	X	X				X	
3	0.934	28.56	0.984	121.88	X			X		X	
2	*0.947*	*63.19*	0.978	155.1		X				X	
2	0.933	48.98	0.978	157.15				X		X	
2	0.90	30.36	*0.983*	*200.01*					X	X	
2	0.89	28.44	0.979	169.94	X					X	
1	0.88	57.74	0.976	332.7						X	
1	0.82	35.44	0.901	73.42							X

Table 4
HANSCH CORRELATION, Rf= a_0 + $\Sigma a_i D_i$; MSA= a_0 + $\Sigma a_i D_i$; D_i = DESCRIPTORS INVOLVED IN THE REGRESSION

Rf= 1.498 - 0.00074* S - 0.0016* ΔH_f + 0.00031* Est				MSA= 683.2622--73.4251* ΔE +0.05827* S-0.04978 * Est			
Regression Statistics				**Regression Statistics**			
Multiple R	0.990758			Multiple R	0.992754		
R Square	0.981601			R Square	0.985561		
Std. Error	0.039925			Std. Error	12.31101		
	Coefficients	*Std. Error*	*P-value*		*Coefficients*	*Std. Error*	*P-value*
Intercept	1.498869	0.0835	$1.9*10^{-6}$	Intercept	683.2622	334.4391	0.087081
S	-0.00074	0.00013	0.00115	ΔE	-73.4251	41.24082	0.125307
ΔH_f	-0.0016	0.00048	0.016	S	0.05827	0.037729	0.173438
Est	0.00031	$5.04*10^{-5}$	0.0008	Est	-0.04978	0.003281	$5.18*10^{-6}$

Table 5
Rf VALUES AND THE DEGREE OF CONFIDENCE FOR HANSCH EQUATION

No.	Theoretical Rf	Experimental Rf	Difference
1	0.20614	0.2	0.00614
2	0.37779	0.38	-0.00221
3	0.40352	0.43	-0.02648
4	0.51264	0.45	0.06264
5	0.49292	0.5	-0.00708
6	0.71398	0.64	0.07398
7	0.68868	0.66	0.02868
8	0.67848	0.69	-0.01152
9	0.95924	0.92	0.03924
10	0.90323	0.95	-0.04677
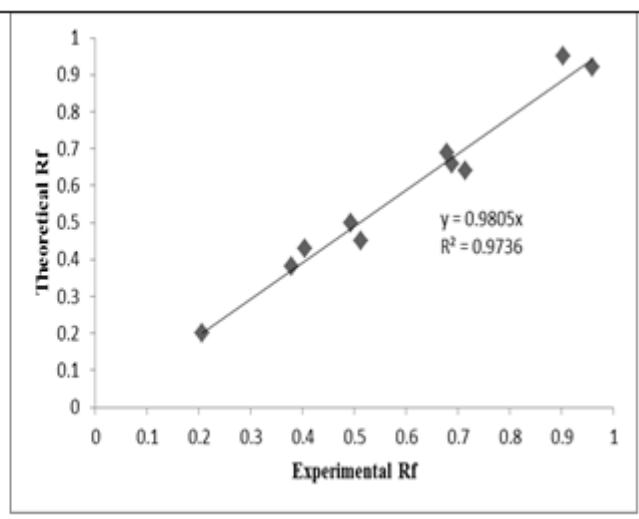
The stabilization energy value (Est) are in relationship with the number of hidrogen donor groups, the substances with the most donor groups (substance 1) being very rea)ctive and interacting the most with the stationary phase. This explains the small Rf value caused by the formation of hydrogen bonds with silica gel.

In the case of the MSA, the best regression correlation indicates the participation of two common descriptors, as in the case of the Rf (S and energy of stabilization) and a different one (ΔE**)**. The best Hansch correlation for both parameters (Rf and MSA) using two descriptors involves a common descriptor, respectively the energy of stability (R^2: 0.947; F:63.19/ R^2: 0.983; F: 200.01).

From our results we can see that the optimal regression with three descriptors for Rf of the studied polyphenols involves: water solubility of the compounds, heat of formation and the energy of stability of the molecule [15]. With these descriptors we obtain the best Hansch equation for Rf, with good predictive power for regression (Rf: R^2 = 0.982/ F= 106.7).

Conclusions

In TLC chromatography with normal phase the stationary phase is polar and the mobile phase is consider to be less polar. Two from three descriptors (S and Est) which contributes to the multiple linear regression of the MSA is also implicated in the best regression of the Rf obtained from TLC. Considering the correlation with two descriptors we can see that the optimal regression for MSA and Rf have in common the descriptor Est.

In conlusion the stabilization energy of the molecule and the solubility plays an major role in TLC separation of phenolic compounds.

References

1. PANDEY, K.B., RIZVI, S.I., Oxid Med Cell Longev., **2**, no. 5, 2009, p 270–278.
2. RUIZ I.L., GOMEZ, M.A., Computational Science and Its Applications – ICCSA, 2008, p 986-996.
3. VARUT, R.M., ROTARU, L.T., Rev. Chim. (Bucharest), **68**, no. 7, 2017, p. 1419
4. WICHTL, M., ANTON, R., Plantes therapeutiques: tradition, pratique officinale, science et therapeutiques, Lavoisier Tec & Doc, Paris, 2003, p 529–530.
5. WALL P.E., Thin-layer Chromatography: A Modern Practical Approach, RSC chromatography monographs, **10**, no. 4, 2006, p 800-850.
6. *** www.chemspider.com/ACD
7. *** www.filestube.com/116d34a4d3cd050403e9,g/Poetable-HyperChem-8-0-7.html
8.*** http://www.4shared.com/file/E6zKU0pg/portable_hyperchem_807.html
9. SPANGENBERG, B., POOLE, C.F., WEINS, C., Quantitative Thin-Layer Chromatography, **2**, 2011, p 1-52.
10.*** https://en.wikipedia.org/wiki/Thin-layer_chromatography#cite_note-hptlc_med-3
11. CONNOLLY, M.L., Science, **221**,1983, p 709-713.
12.*** http://openmopac.net/manual/cosmo.html
13. AMZOIU, E., AMZOIU, M.O., ANOAICA, P.G., Rev. Roum.Chim., **54**, no. 8, 2010, p 671-679.
14. CHIRIGIU, L, BUBULICA, M.V., AVERIS L.M., Rev. Chim. (Bucharest), **63**, no. 8, 2012, p. 764
15. LEO, A., HANSCH,C., ELKINS, D., Chem Rev., **71**, no. 6, 1971, p 525–616.
16. OXTOBY, D.W., GILLIS, H.P., CAMPION, A., Principles of Modern Chemistry, chapter 12, 7TH edition, 2011, p 546-549.