APTEFF, 46, 1-269 (2015) UDC: 577.17:543.544
DOI: 10.2298/APT1546219K BIBLID: 1450-7188 (2015) 46, 219-227
Original scientific paper

CHEMOMETRIC ESTIMATION OF THE RETENTION BEHAVIOR OF SELECTED ESTRADIOL DERIVATIVES

Milica Ž. Karadžić[1], Davor M. Lončar[2], Lidija R. Jevrić[1], Sanja O. Podunavac-Kuzmanović[1], Strahinja Z. Kovačević[1] and Stela D. Jokić[3]*

[1] University of Novi Sad, Faculty of Technology, Bulevar cara Lazara 1, 21000 Novi Sad, Serbia
[2] SUPERLAB, Milutina Milankovića 25, 11070 Novi Beograd, Serbia
[3] J.J. Strossmayer University of Osijek, Faculty of Food Technology, F. Kuhača 18, 31000 Osijek, Croatia

Quantitative structure-retention relationship (QSRR) analysis has been performed in order to correlate the retention of selected estradiol derivatives with their calculated molecular lipophilicity. The lipophilicity descriptors were derived computationally and most important were selected. Linear regression (LR) Wad used for model establishing. Statistical quality of the generated models was determined by standard statistical and cross-validation statistical parameters. Statistically significant and physically meaningful models were obtained. The prediction results are very well correlated with the experimentally observed data. Given predictive ability of the established models indicates that they could be used for predicting the chromatographic behavior of the similar molecules in normal-phase high-performance thin-layer chromatography.

KEY WORDS: estradiol derivatives, chromatography, QSRR, lipophilicity, linear regression

INTRODUCTION

One of the steroid hormones that naturally occurs in the human body is estradiol. It is also an estrogen sex hormone, the primary female sex hormone. Estradiol can be found in most vertebrates as well as many insects, fish and other animal species (1). In many studies, the identification and quantitation of estradiol is a subject of interest. The most sensitive method and widely used technique for estradiol determination is liquid chromatography (LC) (2-5). In environmental analyses, reversed-phase high pressure liquid chromatography (RP HPLC) with different detectors is used (6). Although being less sensitive, the method of thin layer chromatography (TLC) can be used for estradiol determination (7,8).

For understanding the chromatographic processes, it is very useful to establish the mathematical models. Quantitative structure-retention relationship (QSRR) is a technique for determining relationships between chromatographic properties of investigated mole-

* Corresponding author: Milica Ž. Karadžić, University of Novi Sad, Faculty of Technology, Bulevar cara Lazara 1, 21000 Novi Sad, Serbia, e-mail: mkaradzic@hotmail.com

cules and molecular descriptors. In QSRR studies, two groups of input data are needed, dependent and independent variables (9). The use of computational methods increases, particularly for prediction of the influence of various factors on biological activity and for establishing correlations between them. A very important step in establishing the QSRR models is to select the most suitable descriptors for predicting retention. In this paper, molecular lipophilicity descriptors were calculated using suitable software for molecular design.

One of the most important physicochemical properties of a molecule that determines its bioactivity is its lipophilicity. Estimation of the lipophilic character of new, potentially biologically active compounds is considered as one of the first parameters to be determined. Lipophilicity is a molecular property that expresses the relative affinity of a solute for aqueous and organic phases (10). Rapidly increasing use of lipophilicity in modeling of the biological processes indicates the great need for valid procedures for quantification of this physicochemical property.

The aim of this study was to characterize the physicochemical properties of estradiol derivatives and to find the possible relationship between retention characteristics and lipophilicity parameters of investigate compounds in order to understand the separation mechanism in the given chromatographic system. Additionally, the task was to evaluate if established models can be used for the prediction of the retention behavior in given chromatographic systems and to mark the best predictive model for ten estradiol derivatives. Ten estradiol derivatives were examined using normal-phase high-performance thin-layer chromatography (NP HPTLC).

EXPERIMENTAL

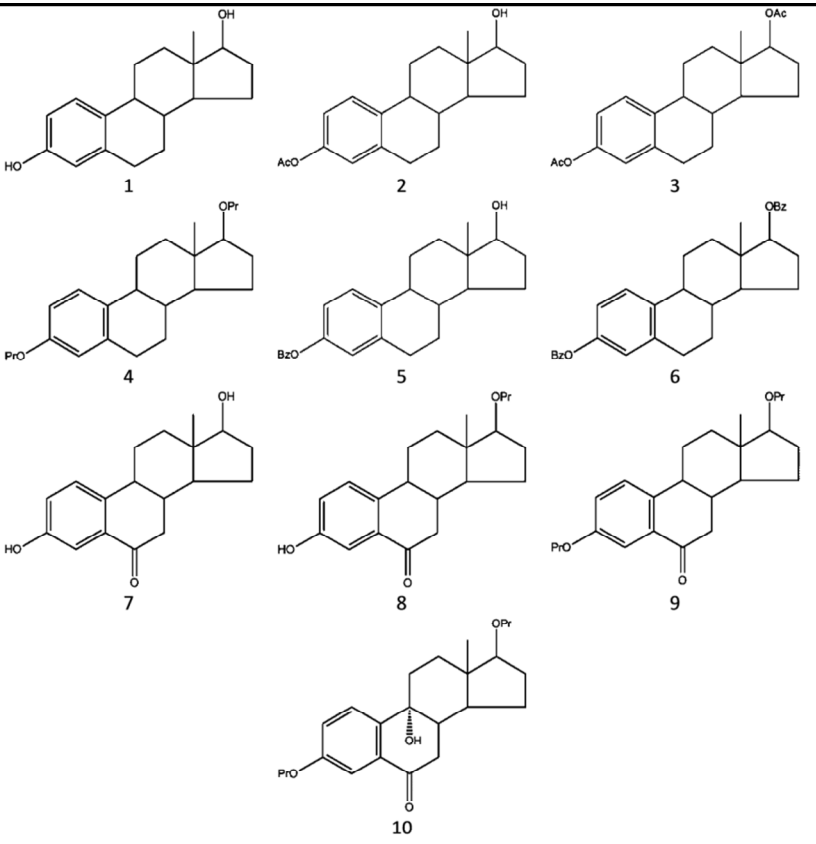
Thin-layer chromatography

High performance TLC was performed on 10 x 10 cm plates that were pre-coated with silica gel 60 with fluorescence indicator (Merck, Darmstadt, Germany). The investigated compounds were dissolved in methanol (2 mg/ml). Glass capillaries were used to apply samples on the plates 1 cm from the edge. The mobile phases were benzene-acetone (9:1 v/v) and benzene-ethyl acetate (9:1; 8:2 v/v). After chromatograms development, the plates were dried at the room temperature (20-22^{o}C) and spots were detected under UV light (λ = 254 nm). The R_F values are averages from at least three chromatograms developed for each solute-mobile phase combination. All solvents used in the analysis were of analytical grade purity.

The names and 2D chemical structures of 10 investigated estradiol derivatives are presented in Table 1.

APTEFF, 46, 1-269 (2015)
DOI: 10.2298/APT1546219K
UDC: 577.17:543.544
BIBLID: 1450-7188 (2015) 46, 219-227
Original scientific paper

Table 1. 2D chemical structures and names of studied compounds



1	3,17-β-dihydroxy-esta-1,3,5(10)-trien
2	2-acetoxy-17-β-hydroxy-estra-1,3,5(10)-trien
3	3,17-β-diacetoxy-estra-1,3,5(10)-trien-1,3,5(10)-trien
4	3,17-β-dipropionoxy-estra-1,3,5(10)-trien
5	3-benzoyloxy-17-β-hydroxy-estra-1,3,5(10)-trien
6	3,17-β-dibenzoyloxy-estra-1,3,5(10)-trien
7	3,17-β-dihydroxy-estra-1,3,5(10)-trien-6-on
8	3-hydroxy-17-β-propionoxy-estra-1,3,5(10)-trien-6-on
9	3,17-β-dipropionoxy-estra-1,3,5(10)-trien-6-on
10	3,17-β-dipropionoxy-9-α-hydroxy-estra-1,3,5(10)-trien-6-on
AcO - acetoxy group; BzO - benzoyl group; Pr - propionyl group	

APTEFF, 46, X-XXX (2015) UDC: 577.17:543.544
DOI: 10.2298/APT1546219K BIBLID: 1450-7188 (2015) 46, 219-227
Original scientific paper

Molecular modeling and molecular descriptors

For 2D molecular structures, ChemBioDraw Ultra v12.0 software was used and 3D structures were obtained by ChemBio3D Ultra v12.0 program (11). The 3D models were subjected to energy minimization using molecular mechanism force field method (MM2). The minimization was performed until the root mean square (RMS) value reached a value smaller than 0.1 kcal/Åmol. The values of the lipophilicity descriptors for each molecule in the data set are computed using the software MarvinSketch v6.1, ChemBio3D Ultra v12.0, ALOGPS 2.1, Molinspiration and PaDEL Descriptor (12-15). ChemBio3D Ultra v12.0 calculates lipophilicity descriptors from 3D structure and all other programs calculate lipophilicity descriptors from 2D structure. Software ALOGPS 2.1. is considered as an accurate tool for predicting the lipophilicity and aqueous solubility of molecules (16).

QSRR analysis

LR was used for QSRR analysis. The LR analysis is used for modeling the relationship between the dependent variable and just one independent variable. It attempts to model the relationship between two variables by fitting a linear equation to observed data. Standard statistical parameters: Fisher's criteria (F), correlation coefficient (r) and standard deviation (s) and *cross*-validation parameters: *cross*-validated coefficient of determination (r^2_{cv}), adjusted coefficient of determination (r^2_{adj}), predictive residual sum of squares (PRESS), total sum of squares (TSS), PRESS/TSS ration and standard deviation based on predicted residual sum of squares (S_{PRESS}) were used for model validation. High values of *cross*-validation parameters (r^2_{cv}, $r^2_{adj} > 0.5$) indicate that the derived QSRR models have a high predictive ability (17). The QSRR modeling was conducted using NCSS&GESS software (18).

RESULTS AND DISCUSSION

Chromatographic data

Chromatographic retention data (R_M) of the investigated compounds are shown in Table 2. In the mobile phases, benzene (Bz) was used as a diluent and acetone (Ac) and ethyl acetate (EtAc) as the modifiers. As it can be concluded, at the same concentration of the modifier the retention of compounds is higher in the eluent with ethyl acetate. This suggests a higher solubility of the compounds in acetone. The higher concentration of the modifier in the mobile phase Bz-EtAc causes a decrease of retention. Also, the replacement of the hydroxy group with an acetyl group (molecules 1 and 2, 2 and 3, 1 and 3) leads to the decrease of the retention values. The same pattern can be noticed when the hydroxy group is replaced with a propionyl group (molecules 1 and 4, 3 and 4, 7 and 8, 7 and 9, 8 and 9). If the hydroxy group is replaced with a benzoyl group (molecules 1 and 5) a decrease of the retention values is also observed. The compounds in which the hydroxyl group is replaced with an acetyl, propionyl or benzoyl group are less retained, and they are less polar than the compounds that have hydroxyl group. This chromatographic behavior is described by the

phenomenon in the normal-phase chromatography: non-polar components pass more quickly through the layer/column than polar components.

Table 2. Chromatographic retention data of studied compounds

	R_M		
	Bz-An 9:1	**Bz-EtAc 9:1**	**Bz-EtAc 8:2**
1	0.259	0.589	0.087
2	-0.158	0.043	-0.358
3	-0.525	-0.399	-1.005
4	-0.589	-0.410	-1.817
5	-0.337	-0.176	-0.513
6	-0.395	-1.065	-1.399
7	0.753	1.235	0.550
8	0.347	0.689	0.158
9	-0.167	0.035	-0.327
10	0.288	0.653	0.122

QSRR analysis

The calculated lipophilicity descriptors are presented in Table 3. It can be concluded that the values of the lipophilicity descriptors are similar, with exception of AlogP and AlogP2, which are very small for the majority of the compounds.

Table 3. Calculated lipophilicity descriptors for ten investigated compounds

Comp	**logPm**[1]	**LogP**[2]	**AlogPs**[3]	**AClogP**[3]	**miLogP**[3]	**AlogP**[3]	**MlogP**[3]	**XlogP2**[3]	**XlogP3**[3]	**miLogP**[4]	**ALogP**[5]	**ALogp2**[5]	**CrippenLogP**[5]	**MLogP**[5]	**XLogP**[5]
1	3.75	3.91	3.57	3.84	3.43	3.81	3.63	4.23	4.01	3.43	0.44	0.19	3.61	3.22	3.47
2	3.66	3.88	3.91	4.07	3.46	3.85	3.98	4.42	2.81	3.46	-0.25	0.06	3.58	3.45	1.95
3	4.10	4.11	4.88	4.55	4.16	4.23	4.34	5.16	3.38	4.17	0.47	0.23	3.83	3.33	3.66
4	5.50	5.42	5.49	5.48	5.19	5.56	4.77	5.67	4.32	5.19	0.85	0.73	4.40	3.44	4.40
5	5.71	5.78	5.25	5.54	5.78	5.51	5.08	6.14	4.47	5.78	1.01	1.01	5.18	3.66	4.91
6	8.21	7.91	7.07	7.51	8.19	7.56	6.42	8.60	6.70	8.19	1.67	2.80	5.12	3.88	4.73
7	2.60	2.55	2.52	3.33	2.53	2.82	2.98	2.86	2.81	2.53	3.25	10.56	6.99	4.54	6.53
8	3.74	3.44	3.39	4.29	3.59	3.87	3.58	3.85	3.86	3.59	-0.11	0.01	3.19	3.11	1.88
9	4.35	4.07	4.03	4.98	4.29	4.57	4.16	4.30	4.43	4.29	0.35	0.12	4.16	3.34	2.88
10	3.27	3.00	3.41	4.23	3.72	3.56	3.39	2.92	3.17	3.72	0.46	0.21	4.77	3.55	3.32

[1]MarvinSketch v6.1; [2]ChemBio3D Ultra v12.0; [3]ALOGPS 2.1; [4]Molinspiration; [5]PaDEL Descriptor

The LR analysis was carried out in order to correlate chromatographic retention (R_M) and calculated lipophilicity values and to determine the best QSRR models that can predict retention behavior. The best LR models that cover all three mobile phases (Table 2) are described by the following equations:

Bz-An 9:1 $R_M = 1.7631 - 0.4389 \cdot \text{AlogPs}$ (1)
Bz-EtAc 9:1 $R_M = 2.4695 - 0.5478 \cdot \text{AlogPs}$ (2)
Bz-EtAc 8:2 $R_M = 2.3460 - 0.6644 \cdot \text{AlogPs}$ (3)

It can be seen that the retention behavior is best described by the AlogPs lipophilicity descriptor. The statistical quality of established equations was determined by the standard statistical and cross-validation parameters (Table 4).

Table 4. Statistical and *cross*-validation parameters for established LR models described by equations 1-3

Eq.	r	F	s	r^2_{cv}	r^2_{adj}	PRESS	TSS	PRESS/TSS	S_{PRESS}
1	0.9057	67.24	0.1479	0.8436	0.8922	0.2541	1.6246	0.1564	0.1594
2	0.8927	58.24	0.1984	0.8201	0.8774	0.4618	2.5676	0.1799	0.2149
3	0.8244	32.87	0.3203	0.6073	0.7994	1.6061	4.0898	0.3927	0.4008

In this study, the high values of r^2_{cv}, r^2_{adj} and F and the low values of PRESS and S_{PRESS} indicate a high predictive ability of the QSRR models described by equations 1-3.

For the confirmation of the absence of a systematic error, the residuals of the predicted R_M values were plotted against the experimentally observed R_M values (Figure 1).

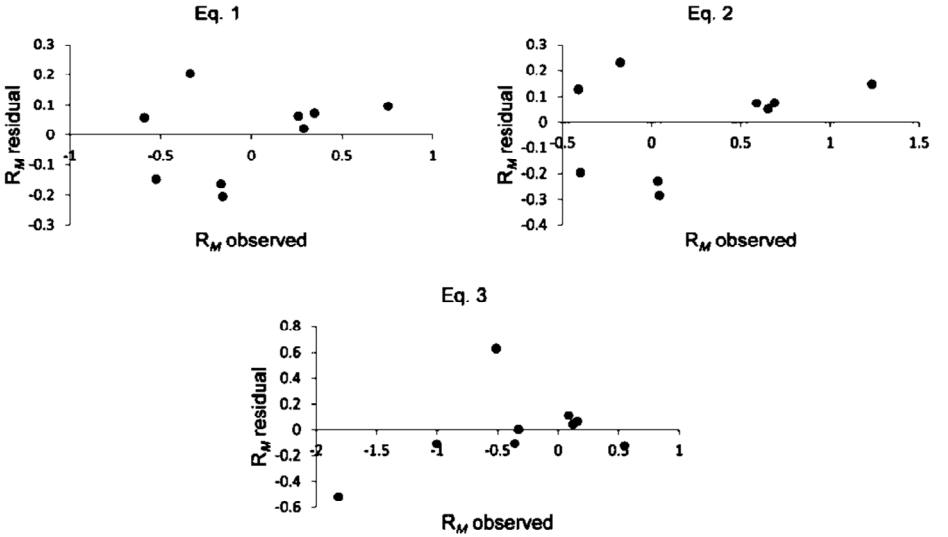


Figure 1. Experimentally observed *versus* residual R_M values

All presented results indicate that the LR models can be successfully used to predict the R_M values of the investigated estradiol derivatives. Based on the result of the given *cross*-validation parameters (Table 4) and graphs (Fig.1) it can be concluded that the best models are described by equations 1 and 2, i.e. for the eluent that contains 10 % acetone and ethyl acetate as modifiers. By the same criteria, the model described by equation 3 is also satisfactory.

In some future NP HPTLC experiments, for suitable solvent systems, the established LR models with the AlogPs lipophilicity descriptor can be used for the prediction of the retention behavior of estradiol derivatives and similar compounds. In addition, it can be concluded that by the lipophilicity parameter, the retention in NP HPTLC can be predicted as good as in reversed-phase thin-layer chromatography.

CONCLUSION

In this paper, the QSRR analysis using LR models was performed in order to identify the most important factors and quantify their influence, to select descriptors that best describe the retention behavior, and to derive mathematical models that could be able to predict the NP HPTLC chromatographic behavior of the investigated estradiol derivatives. The most appropriate molecular descriptor is AlogPs and the best statistical results were obtained with 10% acetone and 10% ethyl acetate as modifiers. Predictive power of the established models allows us to estimate the retention behavior for estrogen derivatives and similar compounds, and understand their behavior in an NP HPTLC system. Considering the small set of molecules encompassed by this study, the predictive ability of the generated models is limited.

Acknowledgment

This paper was performed within the framework of the research projects CMST COST Action No. 1105, and projects No. 172012 and 172014 and 31055 supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

REFERENCES

1. Mechoulam, R.; Brueggemeier, R.W.; Denlinger, D.L. Estrogens in insects. *Cell. Mol. Life Sci.* **1984**, *40* (9), 942-944.
2. López de la Alda, M. J.; Barceló, D. New developments in liquid chromatography mass spectometry for the determination of micropollutants. *J. Chromatogr. A* **2000**, *892* (1-2), 391-406.
3. Lagana, A.; Fago, G.; Marino, A.; Santarelli, D. Liquid chromatography tandem mass spectrometry applied to the analysis of natural and synthetic steroids in environmental waters. *Anal. Lett*. **2001**, *34* (6), 913-926.
4. Ingrand, V.; Herry, G.; Beausse, J.; de Roubin, M. R. Analysis of steroid hormones in effluents of wastewater treatment plants by liquid chromatoraphy–tandem mass spectrometry. *J. Chromatogr. A* **2003**, *1020* (1), 99-104.

5. Kimura, A.; Taguchi, M.; Arai, H.; Hiratsuka, H.; Namba, H.; Kojima, T. Radiation-induced decomposition of trace amounts of β-estradiol in water. *Radiat. Phys. Chem*. **2004**, *69* (4), 295-301.
6. Lopez de Alda, M. J.; Barcelo, D. Use of solid-phase extraction in various of its modalities for sample preparation in the determination of estrogens and progestogens in sediment and water. *J. Chromatogr. A* **2001**, *938* (1-2), 145-153.
7. Medina, M. B.; Schwartz, D. P.; Thin-layer chromatographic detection of zeranol and estradiol in fortified plasma and tissue extracts with Fast Corinth V. *J. Chromatogr.* **1992**, *581* (1), 119-128.
8. Novakovič, J.; Pacáková, V.; Ševčík, J.; Cserháti, T. Quantitative structure-chromatographic retention relationship study of six underivatized equine estrogens. *J. Chromatogr. B* **1996**, *681* (1), 115-123.
9. Karadžić, M.Ž.; Jevrić, L.R.; Podunavac-Kuzmanović, S.O.; Lončar, E.S.; Kovačević, S.Z. Structure-retention relationship study of 2,4-dioxotetrahydro-1,3-thiazole derivatives. *J. Liq. Chromatogr. R. T.* **2015**, *38* (12), 1247-1253.
10. Jevrić, L. R.; Velimirović, S. D.; Koprivica, G. B.; Mišljenović, N. M.; Kuljanin T. A.; Tepić, A. N. 2012. Prediction of s-triazine components lipophilicity of total herbicides. *Romanian Biotechnological Letters* 17:6882-6892.
11. ChemBioOffice 2012. Perkin Elmer Informatics. http://www.cambridgesoft.com/.
12. Chem Axon, Ltd. http://www.chemaxon.com/.
13. VCCLAB, 2005. Virtual Computational Chemistry Laboratory. http://www.vcclab.org.
14. Molinspiration Cheminformatics, http://www.molinspiration.com/.
15. Yap, C.W. PaDEL-Descriptor. An open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem*. **2011**, ***32*** (7), 1466-1474.
16. Tetko, I.; Tanchuk, V. Application of associative neural networks for prediction of lipophilicity in ALOGPS 2.1 program. *J. Chem. Inf. Comp. Sci*. **2002**, *42* (5), 1136-1145.
17. Karadžić, M.Ž.; Jevrić, L.R.; Kovačević, S.Z.; Podunavac-Kuzmanović, S.O. Retention data from normal-phase thin-layer chromatography in characterization of some 1,6-anhydrohexose and D-aldopentose derivatives by QSRR method. *J. Liq. Chromatogr. R. T.* **2015**, *38* (10), 1044-1051.
18. Hintze, J. NCSS and GESS, NCSS, LLC, Kaysville, Utah. http://www.ncss.com/.

ХЕМОМЕТРИЈСКА ПРОЦЕНА РЕТЕНЦИОНОГ ПОНАШАЊА ОДАБРАНИХ ДЕРИВАТА ЕСТРАДИОЛА

Милица Ж. Карацић[1], Давор М. Лончар[2], Лидија Р. Јеврић[1], Сања О. Подунавац-Кузмановић[1], Страхиња З. Ковачевић[1], Стела Д. Јокић[3]

[1] Универзитет у Новом Саду, Технолошки факултет, Булевар цара Лазара 1, 21000 Нови Сад, Србија
[2] СУПЕРЛАБ, Милутина Миланковића 25, 11070 Нови Београд, Србија
[3] J.J. Штросмајер Универзитет у Осијеку, Прехрамбено-технолошки факултет, Ф. Кухача 18, 31000 Осијек, Хрватска

Анализа корелација између хемијске структуре и ретенције (QSRR) је примењена у циљу дефинисања корелације између ретенције одабраних деривата естрадиола и њихове рачунате молекулске липофилности. Дескриптори липофилности су

APTEFF, 46, 1-269 (2015) UDC: 577.17:543.544
DOI: 10.2298/APT1546219K BIBLID: 1450-7188 (2015) 46, 219-227
Original scientific paper

добијени компјутерски и изабрани су најбитнији. Линеарна регресија (LR) је коришћена за успостављање модела. Статистички квалитет генерисаних модела је одређен стандарним статистичким и крос валидационим статистичким параметрима. Статистички значајни и физички смислени модели су задржани. Предикт резултати су веома добро корелирани са експерименталним подацима. Предиктивна способност утврђених модела указује да они могу бити коришћени за предикцију хроматографског понашања сличних молекула у нормално фазној танкослојној хроматографији високог учинка (NP HPTLC).

Кључне речи: деривати естрадиола, хроматографија, анализа корелација између хемијске структуре и ретенције, липофилност, линеарна регресија

Received: 1 September 2015.
Accepted: 2 October 2015.