

A CAMD-QSAR METHODOLOGY FOR PROSPECTING THE POTENTIAL ACTIVE ANTIHISTAMINIC SUBSTANCES: CLASS OF MANDELIC ACID ESTERS

PAUL-GABRIEL ANOICA^{1*}, MIHAELA-BIANCA ANOICA²,
EMILIA AMZOIU¹, LUCIAN GIUBELAN³, MARIANA TĂTUCU¹

¹*Department of Physical Chemistry, Faculty of Pharmacy*

²*Molinette Hospital, Neurology II, Department of Neurosciences, Headache Centre, University of Turin, Via Cherasco 15, I-10126 Torino, Italy*

³*Department of Infectious Diseases, Faculty of Medicine Craiova*

^{1,3}*University of Medicine and Pharmacy of Craiova, 2-4 Petru Rareș Street, 200349 Craiova, Romania*

**corresponding author: gabriel@umfcv.ro*

Abstract

A first step in a pharmaceutical micro-production laboratory is to create new potential active substances, starting from a specified class of pro-drugs. In this regard, the article presents a way of handling two biological properties for a class of substances. It was chosen the mandelic acid esters class – with antihistaminic activity. In this case, we observed the logarithm of partition coefficient for biological membrane ($\log P$) and the $\log (1/C)$ antihistaminic activity, where C represents the minimum molar concentration needed to obtain a clear antihistaminic response. Knowing these properties for 11 substances, predictions can be made for other substances from this pharmaceutical class. Note that, if we increase the biological pursued properties the accuracy and credibility of default predictions will increase. We are interested in the extreme values of activity (minimum and especially maximum) for the biological properties handled.

Rezumat

Un prim demers în activitatea unui laborator farmaceutic de microproducție este de a crea noi substanțe potențial active, plecând de la o clasă specificată de prodroguri. În acest scop, prezentăm modul de manevrabilitate a două proprietăți biologice pentru o clasă de substanțe. A fost aleasă clasa esterilor acidului mandelic, care prezintă activitate antihistaminică. În acest caz s-au urmărit logaritmul coeficientului de partiție pentru membrana biologică ($\log P$) și în special activitatea antihistaminică $\log (1/C)$, unde C reprezintă concentrația molară minimă necesară obținerii unui răspuns antihistaminic evident. Cunoscând aceste proprietăți la 11 dintre substanțe, se pot face predicții referitoare la alte substanțe din clasă. De remarcat că, dacă se crește numărul proprietăților biologice urmărite, se mărește implicit acuratețea și credibilitatea predicției. Ne interesează valorile extreme ale activității (minimă și mai ales maximă) pentru proprietățile biologice analizate.

Keywords: quantitative structure-activity relationship (QSAR), computer aided molecular design (CAMD); antihistaminic activity; mandelic acid esters.

Introduction

The objective of this study was to present a process of chemical modeling for new chemical structures design. For this, we monitored the biological activity by the development of predictive algorithms that use different types of experimental data.

The material is based upon a *Computer Assisted Molecular Design* (CAMD) methodology. It parameterizes and designs the molecular structures of pharmaceutical interest based on the relationship between the biological activity (measured experimentally) and the molecular descriptors (which represent the chemical structure of the analyzed substances).

These descriptors, of electrostatic fingerprint that should cover both the form and quantum-molecular properties [1, 2], constitute the interface through which chemical or biological activity is correlated with the structure of the chemical substances. This is the *Quantitative Structure–Activity Relationship* (QSAR) method [3].

Our hope, for a set of molecules with determined biological activity was to achieve a statistical invariant and predictive model.

This method should be able to answer some questions, as follows:

1. Can predictions be made at the pharmaceutical level based on preliminary experimental data?
2. How can the numerical information (molecular descriptors) be used in the evaluation of the biological activity?
3. How efficient and far can we go in the statistical correlation?

Materials and methods

The concrete method is based on a set of 11 mandelic acid esters (Figure 1). This class has antihistaminic activity and is parameterized using $\log (1/C)$ – Table I (no. 1 to no. 11) [4].

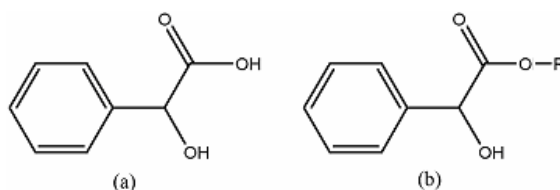


Figure 1

(a) mandelic acid; (b) esters of mandelic acid: R – alkyl radical (according to the Table I).

We intended to extend this class and to calculate the antihistaminic activities of another two related substances (no. 0 and no. 12 in Table I).

Table I
Antihistaminic activity of mandelic acid esters,
parameterized using $\log (1/C)$

Ester (R)	$\log P$	$\log (1/C)$
0: <i>Mandelic Acid</i>	—	?
1: Methyl	0.41	-0.52
2: Ethyl	0.91	-0.22
3: Propyl	1.41	0.20
4: Butyl	1.91	0.59
5: Pentyl	2.41	1.08
6: Hexyl	2.91	1.52
7: Heptyl	3.41	1.70
8: Octyl	3.91	2.18
9: Nonyl	4.41	2.26
10: Decyl	4.91	1.45
11: Undecyl	5.41	1.28
12: <i>Dodecyl</i>	—	?

C – minimum molar concentration needed to obtain a clear antihistaminic response;

P – partition coefficient for biological membrane

Calculus of trans-membrane partition coefficient

As a first step in our work, we wanted to double this basic biological property by another bio-molecular property. Thus, we wanted to control and restrict most efficiently our results.

It was chosen the partition coefficient $\log P$, because the molecular set has a lipophilic trend. This coefficient was measured for 11 esters by Funcke AB *et al.* (Table I) [4]. They correlated $\log P$ with the antihistaminic activity obtaining good results in correlation ($N = 11$; $R^2 = 0.988$; $S = 0.121$; $F = 195.13$) in a “cut-off” model [4] but with a relatively low predictive power.

Therefore, we intended to determine $\log P$ for substances no. 0 and no. 12 from Table I.

In addition, at the interface between the aqueous and lipidic phase (simulated by 1-octanol) structures are organized because of the polar character of water molecules and the amphoteric character of 1-octanol molecules. Such bonded molecules between the aqueous phase and 1-octanol can simulate well enough the cell membrane [5].

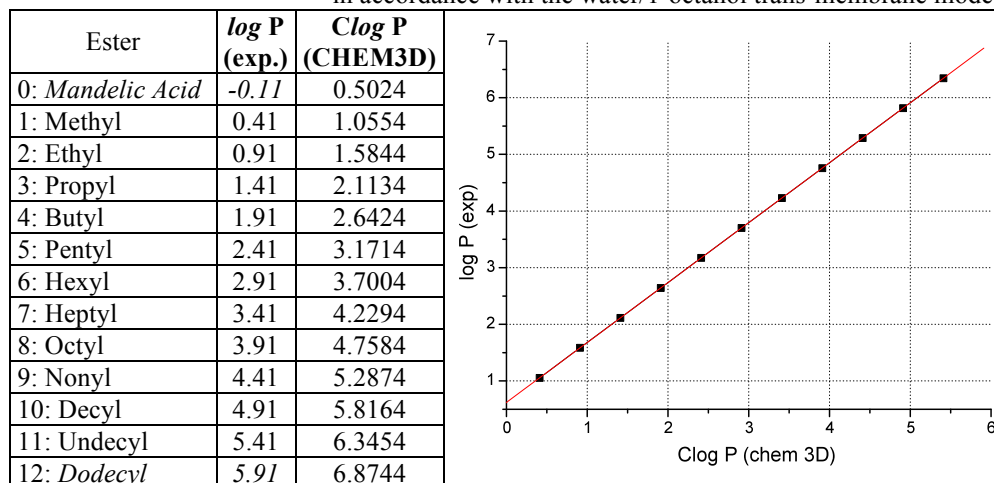
We have made a linear correlation (Table II) between the experimental value of $\log P$ and the values calculated in accordance with the water/1-octanol trans-membrane model ($C\log P$ by CHEM3D – ChemOffice Ultra computational program).

The result of our linear correlation is summarized in Table III.

In order to have a confirmation of this prediction we re-correlated $\log P$ values in a Hansch-QSAR equation. The Hansch-equation provides a mono-variate linear correlation, in our case, between the experimental $\log P$ values related to the structural descriptors of 11 esters.

Table II

Mathematical estimation of $\log P$ based on the value of $C\log P$ coefficient calculated with CHEM3D (ChemOffice Ultra) computational program – in accordance with the water/1-octanol trans-membrane model



$\log P$ (exp.) – experimental; $C\log P$ (CHEM3D) – calculated.

Table III

Linear regression: $\log P$ [experimental] = A + B * $C\log P$ [CHEM3D]

Parameter	Value	Error	
A (intercept)	0.62162	$6.93314 \cdot 10^{-16}$	
B	1.058	$2.09346 \cdot 10^{-16}$	
R^2	SD	N	P
0.9999	$1.098 \cdot 10^{-15}$	11	<0.0001

R^2 – square of the regression coefficient of correlation or coefficient of determination (0 designates minimum of correlation and 1 maximum of correlation between this two variables); SD – standard deviation for N (11) values; P – significance level.

To estimate the molecular descriptors that may contribute to the process followed it was carried out the structural modeling of the studied chemical substances.

We first used the Molecular Mechanics approach (MM+) and then molecular geometries were obtained using quantum-molecular approximation MOPAC 7.0 (RHF, PM3). Atomic coordinates thus obtained were used as input data for CODESSA software. This program calculates about 400 different descriptors with a large variety (topological, quantum-chemical, constitutional, geometrical or electrostatic descriptors) [6].

We calculated Connolly molecular CHEM 3D form descriptors [7]

and then we imported them into the CODESSA program.

Therefore, we can get a very good QSAR correlation for $\log P$ in one descriptor, according to Connolly Excluded Solvent Volume (CSEV) (Table IV).

Table IV

Linear regression: $\log P$ [experimental] = $a + b \cdot \text{CSEV}$

Parameter (description)	Value	Error	t-test
a (intercept)	$-3.3919 \cdot 10^{+00}$	$4.3867 \cdot 10^{-02}$	-77.3216
b (CSEV coefficient)	$2.8938 \cdot 10^{-02}$	$1.9538 \cdot 10^{-04}$	148.1088
R^2	F	N	s^2
0.9996	21936.2207	11	0.0013

R^2 – coefficient of determination; F – Fisher index for multiple groups of N (11) values; s^2 – standard deviation of the regression; CSEV – Connolly Excluded Solvent Volume.

Based on Table IV, the partition coefficient values can be recalculated and compared with the values obtained in the water/1-octanol model (Table III).

There were obtained good correlations for $\log P$, in these two models (structure in trans-membrane model and QSAR equation), as it can be seen in table V, for the compounds no. 0 to no. 12.

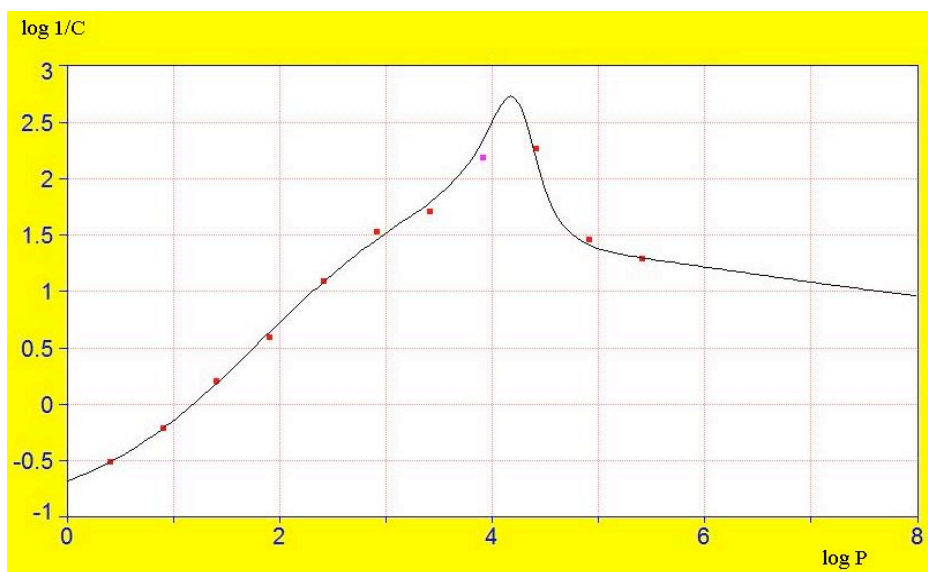
Table V

Comparison of $\log P$ for no. 0 and no. 12 structures in trans-membrane model (column 2) and QSAR equation (column 4)

Ester	$\log P$	CSEV	$\log P$
		$[A^3]$	QSAR
0: Mandelic Acid	-0.11	114.634	-0.0746
1: Methyl	0.41	134.147	0.4901
2: Ethyl	0.91	148.153	0.8954
3: Propyl	1.41	165.279	1.3910
4: Butyl	1.91	182.225	1.8814
5: Pentyl	2.41	200.101	2.3987
6: Hexyl	2.91	216.674	2.8783
7: Heptyl	3.41	234.26	3.3872
8: Octyl	3.91	252.291	3.9090
9: Nonyl	4.41	269.398	4.4040
10: Decyl	4.91	288.228	4.9489
11: Undecyl	5.41	304.722	5.4262
12: Dodecyl	5.91	321.134	5.9011

Calculation of antihistaminic activity

In order to calculate the antihistaminic activity, we computed the dependence between the two biological parameters: $\log (1/C) = f(\log P)$. The chart variation obtained with TableCurve program was very suggestive, even if its polynomial form was more complicated (Figure 2, Table VI).



$$\log (1/C) = [a + c \cdot \log P + e \cdot (\log P)^2 + g \cdot (\log P)^3] / [1 + b \cdot \log P + d \cdot (\log P)^2 + f \cdot (\log P)^3 + h \cdot (\log P)^4]$$

Figure 2

Representation $\log (1/C) = f(\log P)$ – polynomial interpolation.

Table VI

Dependence parameters represented in Figure 2

Coefficient	Value	Standard error	t-value	99% Confidence limits	
a	-0.68765	0.48841	-1.40794	-3.53727	2.16197
b	-0.80592	1.80405	-0.44673	-11.3317	9.71983
c	0.88816	0.65994	1.34581	-2.96228	4.73859
d	0.29968	1.26622	0.23667	-7.08809	7.68744
e	-0.29847	0.51605	-0.57837	-3.30938	2.71244
f	-0.05896	0.33732	-0.17478	-2.02707	1.90915
g	0.03023	0.09482	0.31878	-0.52303	0.58348
h	0.00473	0.0327	0.14476	-0.18604	0.19551

r^2 coef. det.	DF adj. r^2	Fit std. err.	F-value
0.9945473703	0.9727368515	0.1250962841	78.170462714

r^2 coef. det. – coefficient of determination; DF adj. r^2 – the adjusted r^2 ; Fit std. err. – mean square deviation; F-value – Fisher index of the correlation.

This equation presented in Figure 2 allowed us to estimate the antihistaminic activities corresponding to $\log P$'s structures no. 0 and no. 12 (Table VII).

For compliance, it was designed the QSAR equation – this time between the intrinsic activity, $\log (1/C)$, of the 11 basic structures and the absolute entire collection of molecular descriptors of the CODESSA program (obviously with Conolly descriptors). A linear QSAR relationship resulted in the following three parameters (Table VIII).

Table VII

Mathematical estimation for antihistaminic activities structures no. 0 and no. 12, based on the mathematical equation presented in Figure 2

Ester	$\log (1/C)$	$\log P$
0: <i>Mandelic Acid</i>	-0.72 $\Leftarrow$	-0.11
1: Methyl	-0.52	0.41
2: Ethyl	-0.22	0.91
3: Propyl	0.2	1.41
4: Butyl	0.59	1.91
5: Pentyl	1.08	2.41
6: Hexyl	1.52	2.91
7: Heptyl	1.7	3.41
8: Octyl	2.18	3.91
9: Nonyl	2.26	4.41
10: Decyl	1.45	4.91
11: Undecyl	1.28	5.41
12: <i>Dodecyl</i>	1.22 $\Leftarrow$	5.91

Table VIII

QSAR equation: antihistaminic activity $\leftrightarrow$ molecular descriptors

$$\text{Linear regression: } \log (1/C) = x_0 + \sum_{i=1}^3 x_i \cdot d_i$$

Parameter (descriptor)	Value	Error	t-test
x_0 (intercept)	$2.1315 \cdot 10^{+02}$	$4.3867 \cdot 10^{+01}$	3.3425
x_1 (molecular volume / XYZ Box)	$-1.8043 \cdot 10^{+01}$	$6.3770 \cdot 10^{+00}$	-3.7432
x_2 (max. bonding contribution of a MO)	$-9.9750 \cdot 10^{+01}$	$3.1190 \cdot 10^{+01}$	-3.1982
x_3 ((1/2)X BETA polarizability (DIP))	$8.1009 \cdot 10^{-02}$	$2.3463 \cdot 10^{-02}$	3.4527
R^2	F	N	s^2
0.9550	49.55	13	0.0553

XYZ – a parallelepiped box; MO – molecular orbital; DIP – molecular dipole; R^2 – coefficient of determination; F – Fisher index for multiple groups of N (11) values; s^2 – standard deviation of the regression.

It could thus be recalculated the antihistaminic activity for the two unknown structures (no. 0 and no. 12) (Table IX).

Table IX

Antihistaminic activity value, based on QSAR equation (Table VIII)

Structure	Calculated $\log (1/C)$	Experimental $\log (1/C)$	Difference	Absolute difference [%]
0: <i>Mandelic Acid</i>	-0.8340	–	–	
1: Methyl	-0.4691	-0.5200	0.0509	9.788462
2: Ethyl	-0.3652	-0.2200	-0.1452	66
3: Propyl	0.2173	0.2000	0.0173	8.65
4: Butyl	0.8033	0.5900	0.2133	36.15254
5: Pentyl	0.8994	1.0800	-0.1806	16.72222

Table IX (continued)

Structure	Calculated <i>log</i> (1/C)	Experimental <i>log</i> (1/C)	Difference	Absolute difference [%]
6: Hexyl	1.7458	1.5200	0.2258	14.85526
7: Heptyl	1.7901	1.7000	0.0901	5.3
8: Octyl	1.8850	2.1800	-0.2950	13.53211
9: Nonyl	1.9993	2.2600	-0.2607	11.5354
10: Decyl	1.7159	1.4500	0.2659	18.33793
11: Undecyl	1.2983	1.2800	0.0183	1.429688
12: Dodecyl	1.6890	–	–	

The trend of overall variation of *log* (1/C) has also been recorded.

Results and discussion

We estimated the antihistaminic activities, *log* (1/C), by two independent methods of prediction: a polynomial equation (depends on *log* P – Table VII) and QSAR equation – Table IX.

Now, we can compare the predicted values for the antihistaminic activities on the two tested compounds (no. 0 and no. 12) (Table X).

Table X

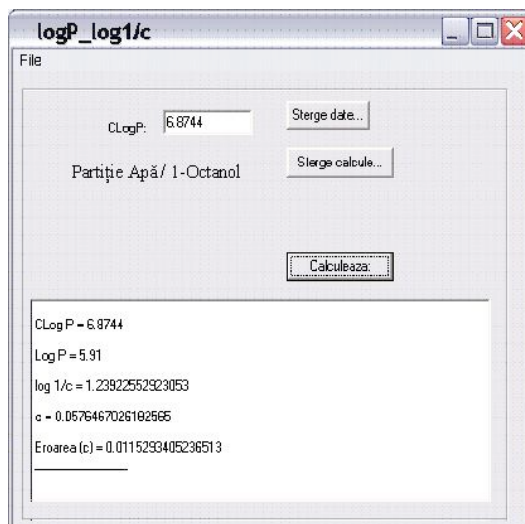
Independent test of predicted antihistaminic activities calculated using the polynomial equation (by *log* P – Table VII) / QSAR (Table IX) equation

Structures	Calculation with the polynomial model (<i>log</i> P)	Calculation with Hansch / QSAR equation
0: <i>Mandelic Acid</i>	-0.72	-0.8340
12: <i>Dodecil</i>	1.22	1.6890

As it can be seen in Table X, the values calculated using the QSAR method validate the trans-membrane partition model.

Conclusions

This methodology was applied in order to achieve a *predictive and invariant statistical model* for other substances in the antihistaminic tested class. We can obtain, in this case, new information on the parameterization process of interaction between the drug and the active sites. The results were conclusive and encouraging for the experimental practice and calculation of other classes of substances. The increasing number of biological parameters control will automatically lead to several ways of validating the predictive outcome and increase the credibility of results. All data obtained were incorporated into a “window of analysis” performed by our team. It can automatically calculate the antihistaminic activity using the partition coefficient (*Clog* P) water/1-octanol model on the considered antihistaminic class (Figure 3).

**Figure 3**

Window of analysis of antihistaminic activity in correlation with the partition coefficient ($Clog P$ water/1-octanol in the trans-membrane model).

Even if the calculation model can be improved using other biological properties (e.g. toxicity), the validation results can certify the synthesis and experimental testing of theoretically evaluated compounds.

References

1. Anoaica P. G., Lepădatu C. I., QSAR study for classes with a broad range of biological activity using electronegativity descriptors for OMO–UMO quantum states. Clotrimazole imidazole derivatives with antifungal activity, *Revue Roumaine de Chimie*, 2007, 52(3):293–297.
2. Anoaica P. G., Amzoiu Emilia, Lepădatu C. I., The hierarchy of the molecular descriptors and the ligand–receptor interaction, *Revue Roumaine de Chimie*, 2007, 52(8–9):789–793.
3. Vlaia V., Olariu T., Vlaia Lavinia, Butur M., Ciubotariu C., Medeleanu M., Ciubotariu D., Quantitative structure–activity relationships (QSAR). V. Analysis of the toxicity of aliphatic esters by means of molecular compressibility descriptors, *Farmacia*, 2009, 57(5):549–561.
4. Funcke A. B., Ernsting M. J., Rekker R. F., Nauta W. T., Studies of spasmolytics. I. Mandelic acid esters, *Arzneimittel-Forschung*, 1953, 3(10):503–506.
5. Wimley W. C., White S. H., Experimentally determined hydrophobicity scale for proteins at membrane interfaces, *Nature Structural Biology*, 1996, 3(10):842–848.
6. Schmidt M. W., Baldrige K. K., Boatz J. A., Elbert S. T., Gordon M. S., Jensen J. H., Koseki S., Matsunaga N., Nguyen K. A., Su S., Windus T. L., Dupuis M., Montgomery J. A. Jr, General atomic and molecular electronic structure system, *Journal of Computational Chemistry*, 1993, 14(11):1347–1363.
7. Connolly M. L., Solvent-accessible surfaces of proteins and nucleic acids, *Science*, 1983, 221(4612):709–713.

Manuscript received: October 1st 2009