MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **75** (2016) 533-558

ISSN 0340 - 6253

A Statistical Method for Calculation of Intramolecular Synergy

Laszlo Tarko

Centre of Organic Chemistry – Romanian Academy, Romania, Bucharest, Sector 6, Spl. Independentei 202B, PO box 35-108, MC 060023

ltarko@cco.ro ; tarko_laszlo@yahoo.com

(Received December 2, 2015)

Abstract

Frequently, there is an intramolecular reciprocal influence between fragments (or substituents) grafted on a common skeleton and the bioactivity of a specific compound depending on this synergy or antagonism. The synergism or antagonism of molecular fragments can be computed using molar refractivity MR or the percentage in weight of fragments and the value of bioactivity of molecules. This paper proposes some new statistical formulae for calculating the sign and the level of intramolecular synergy for DrugDesignFocus index. The proposed formulae are useful for the identification of the substituted positions and fragments that could be valuable in drug design for the synthesis of compounds having a greater than zero value of the DrugDesignFocus index. The MR sums or the sums of percentages of fragments, proposed as new (synergy) descriptors, are frequently used as QSAR predictors. The paper presents the results obtained after the analysis of three groups of molecules including 78 flavones, 50 phenols and 76 N-methyl-phenyl urethanes, respectively.

1. Introduction

Consider a group of molecules having the same studied property P. Each molecule possesses the characteristics F_1, F_2 … F_n and the effects of certain molecular features on P in said molecule are E_1, E_2 … E_n. An analysis of these molecules shows that the presence of F_i *and* F_j may or may not lead to the total effect $E_i + E_j$.

In principle, '*similar* structures have *similar* bio-activities'. This statement is viewed as a 'QSAR axiom', where QSAR means *Quantitative Structure-Activity Relationship.* This 'axiom' is challenged by the fact that there are similar enough molecules that present non-similar values of bioactivity, an example of which are the enantiomers and certain fluoro-ethers anesthetics.

Moreover, it was also observed that if $E_i + E_j \sim 0$, then two *non-similar* molecules can have *similar* bioactivity. Also, assuming that the presence of the molecular fragment X increases the bioactivity and the presence of fragment Y decreases the bioactivity, then a molecule including both X *and* Y and a molecule without X *and* without Y can have similar bioactivity. For example, ethylene glycol, β-pinene and *cis*-3-hexenol present the same value 4700 mg/kg of LD_{50} (rat, oral route) [1]. Consequently, the observance of QSAR axiom does not imply the observance of the 'opposite QSAR axiom', i.e. the statement '*non-similar* structures present *non-similar* values of activities' does not stand.

There is an intramolecular influence of fragments (or substituents grafted on common skeleton) which affects bioactivity.

Synergy is a widely used term in biochemistry, pharmacology, mathematics and other fields [2, 3]. This term is derived from the Greek *syn-ergos,* meaning *working together.* Simply put, if *the whole is different than the sum of its parts,* then there is a *synergy of its parts.* Accordingly, the synergy is considered an additive property of the analyzed phenomenon. By definition, if the effect of the presence of features F_i *and* F_j *is different* (higher or lower) than $E_i + E_j$ sum, there is said to be a *synergistic* (positive or negative) effect of F_i and F_j. A negative synergistic effect is sometimes referred to as *antagonistic* effect.

In the QSAR field there are many qualitative (non-mathematical) comments regarding the synergistic effects of molecular features or components in mixtures [4-7].

Based on the experimental value of bioactivity, the synergy of two molecules can be calculated using advanced methods [8-14] or with simple formula (1), where TU_{sum} represents the sum of toxic units [15]. If the calculated $TU_{sum} < 0.8$ then, there is a synergistic effect, but if the calculated $TU_{sum} > 1.2$, there is said to be an antagonistic effect.

$$TU_{sum} = C_A / EC_{50-A} + C_B / EC_{50-B} \quad (1)$$

where

C_A is the concentration of component A in the mixture at median inhibition
C_B is the concentration of component B in the mixture at median inhibition
EC_{50-A} is the median effective inhibition concentration of chemical A
$_{-B}$ is the median effective inhibition concentration of chemical B

Experimental values of synergy can be used in QSAR computations [16].

The multilinear QSARs describe the analyzed systems better than the sum of the individual descriptors *because* the descriptors' action is synergistic [17, 18]. This point of view regarding the synergistic effect is, in our opinion, debatable, because any multilinear QSAR is a weighted sum of the individual descriptors. In addition, the forward selection procedure [19-21] yields QSARs including n_1 descriptors that *always* have a higher predictive power than any QSAR including n_2 descriptors ($n_1 > n_2$), even if the synergistic effects are non-existent. Actually, an increase of predictive power emphasizes a *low* intercorrelation of descriptors, not the existence of some synergistic effects of descriptors.

Thus, this paper proposes a first-hand formula for the calculation of synergistic effects of molecular fragments in analyzed molecules or substituents grafted on common skeleton, before QSAR computation. Thereafter, the application of QSAR methodology verifies if some proposed synergy descriptors are predictors. As an example, we analyzed three groups of molecules, including derivatives of 78 flavones, 50 phenols and 76 N-methyl-phenyl urethanes, respectively.

2. Analyzed groups of molecules

The main inhibitory neurotransmitter in the vertebrates' central nervous system is γ-amino-butyric acid (GABA). The $GABA_A$ type receptor is a ligand-gated ion channel and presents three active sites: site *A* and *B* for GABA, and site *C* for benzodiazepine [22].

The interaction of flavonoid ligands with benzodiazepine site of $GABA_A$ is measured, as bioactivity, by the value of the logarithm of the binding affinity constant log K, quoted in literature [23, 24]. The analyzed flavones are presented in Figure 1 and Table 1.

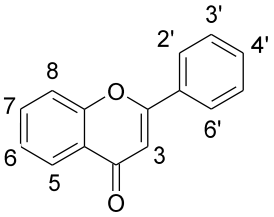

Figure 1 The flavone (common skeleton of analyzed derivatives)

Table 1 The structure of flavone's derivatives and observed values of logK

No.	Substituent(s)	logK	No.	Substituent(s)	logK
1	6-fluoro, 3'-methoxy	0.398	**40**	6,2'-dihydroxy	- 1.469
2	6-bromo, 3'-methoxy	- 0.215	**41**	5,7,2'-trihydroxy, 6-methoxy	- 1.420
3	6-chloro, 4'-methoxy	0.097	**42**	5,7,2'-trihydroxy	- 1.125
4	6-bromo, 4'-methoxy	0.322	**43**	2'-hydroxy	- 0.678
5	6-chloro, 2'-fluoro	- 0.380	**44**	5,7-dihydroxy, 6,8-dimethoxy	- 0.699
6	6-bromo, 2'-fluoro	- 0.424	**45**	7,2'-dihydroxy	- 0.252
7	6,3'-difluoro	- 0.036	**46**	5,7-dihydroxy, 6-methoxy	- 0.051
8	6-chloro, 3'-fluoro	- 0.920	**47**	5,7-dihydroxy, 8-methoxy	0.182
9	6-bromo, 3'-fluoro	- 1.377	**48**	6-hydroxy	0.422
10	4'-fluoro	0.556	**49**	7-hydroxy	0.623
11	6,4'-difluoro	0.398	**50**	5,6,7-trihydroxy	0.747
12	6-chloro, 4'-fluoro	- 0.742	**51**	6-hydroxy, 2'-methoxy	0.976
13	3'-chloro	- 0.212	**52**	2'-methoxy	1.508
14	6,3'-dichloro	- 1.638	**53**	6-methoxy, 2'-amino	0.544
15	3'-bromo	- 0.384	**54**	none	0.000
16	6-fluoro, 3'-bromo	- 0.627	**55**	5,7-dihydroxy	0.477
17	6-chloro, 3'-bromo	- 1.638	**56**	5,3',4'-trihydroxy, 6,7-dimethoxy	2.301
18	6,3'-dibromo	- 1.721	**57**	5,4'-dihydroxy, 6,7-dimethoxy	1.362
19	6-bromo, 4'-nitro	- 0.699	**58**	5,7,4'-trihydroxy, 6-methoxy	0.000
20	6-bromo	-1.155	**59**	5,7-dihydroxy, 2'-chloro	0.903
21	6-chloro	- 0.785	**60**	5,7-dihydroxy, 2'-fluoro	0.903
22	6-nitro	- 0.561	**61**	5,7-dihydroxy, 6,8-dibromo	- 0.155
23	6-methoxy	- 0.066	**62**	5,7,4'-trihydroxy	0.602
24	6-fluoro	0.653	**63**	3,5,7,4'-tetrahydroxy	1.969
25	6-bromo, 3'nitro	- 3.000	**64**	5-hydroxy, 7-methoxy	1.699
26	6-methyl, 3'nitro	- 2.252	**65**	5,7-dihydroxy, 6,8-diiodo	0.000
27	6-chloro, 3'-nitro	- 2.097	**66**	6-fluoro, 3'-hydroxy	0.400

28	6,3'-dinitro	- 1.585	**67**	6-chloro, 3'-hydroxy	- 0.070
29	6-fluoro, 3'-nitro	- 0.745	**68**	6-bromo, 3'-hydroxy	- 0.220
30	3'-nitro	- 0.545	**69**	6-bromo, 2'-nitro	- 0.680
31	6-methyl, 3'-bromo	- 1.886	**70**	6-nitro, 4'-bromo	-1 .600
32	6-nitro, 3'-bromo	- 1.602	**71**	3'-methoxy	0.380
33	6-hydroxy, 3'-bromo	0.000	**72**	6-chloro, 3'-methoxy	- 0.072
34	6-methoxy, 3'-bromo	0.000	**73**	3'-fluoro	0.550
35	6,3'-dimethyl	- 0.682	**74**	6-bromo, 4'-fluoro	- 0.939
36	3'-methyl	1.000	**75**	6-fluoro, 3'-chloro	- 0.701
37	5,2'-dihydroxy, 6,7,8,6'-tetramethoxy	- 0.444	**76**	6-bromo, 3'-chloro	- 1.770
38	5,7,2'-trihydroxy, 6,8-dimethoxy	- 2.215	**77**	6-methyl	- 0.903
39	2'-hydroxy, 7,8-benzo	- 1.569	**78**	6-bromo, 3'-methyl	- 0.812

Protozoa are used as indicators of pollution of water due to their sensitivity to alterations of environment. *Tetrahymena pyriformis* is one of the most widely used protozoa in laboratory research because it can grow in non-expensive media, has a short generation time and is non-pathogenic. Consequently, the literature is flooded by papers addressing the toxic effect of various molecules to *Tetrahymena pyriformis* protozoan. Here, we analyzed the derivatives of phenol in Table 2. The value of toxicity T, here weighted within {0, 2.638} range, is quoted in literature [25, 26].

Table 2 The structure of phenol's derivatives and observed value of toxicity

No.	**Substituent(s)**	**T**	**No.**	**Substituent(s)**	**T**
1	none	0.000	**26**	2,5-dichloro	1.559
2	2,6-difluoro	0.827	**27**	2,3-dichloro	1.702
3	2-fluoro	0.679	**28**	2-methyl, 4-chloro	1.131
4	4-fluoro	0.448	**29**	3-methyl, 4-chloro	1.226
5	3-fluoro	0.904	**30**	2,4-dichloro	1.467
6	4-methyl	0.239	**31**	3-*tert*-butyl	1.161
7	3-methyl	0.369	**32**	4-*tert*-butyl	1.344
8	2-chloro	0.708	**33**	3,5-dichloro	1.993
9	2-bromo	0.935	**34**	2-phenyl	1.525
10	4-chloro	0.976	**35**	2,4-dibromo	1.834
11	3-ethyl	0.660	**36**	2,3,6-trimethyl	0.849
12	2-ethyl	0.607	**37**	3,4,5-trimethyl	1.361
13	4-bromo	1.112	**38**	2,4,6-trimethyl	2.126
14	2,3-dimethyl	0.553	**39**	3,5-dimethyl, 4-chloro	1.634
15	2,4-dimethyl	0.559	**40**	2,6-dichloro, 4-bromo	2.210
16	2,5-dimethyl	0.440	**41**	2,4,5-trichloro	2.531
17	3,4-dimethyl	0.553	**42**	2-methyl, 4-bromo, 6-chloro	1.708
18	3,5-dimethyl	0.544	**43**	2,6-dimethyl, 4-bromo	1.709
19	3-chloro, 4-fluoro	1.273	**44**	2,4,6-tribromo	2.481

20	2-chloro, 5-methyl	1.071	**45**	2-*tert*-butyl, 4-methyl	1.728
21	4-iodo	1.285	**46**	2-*iso*-propyl, 4-chloro, 5-methyl	2.293
22	3-iodo	1.549	**47**	2,4-dimethyl, 6-*tert*-butyl	1.676
23	2-*iso*-propyl	1.234	**48**	2,6-diphenyl	2.544
24	3-*iso*-propyl	1.040	**49**	2,4-dibromo, 6-phenyl	2.638
25	4-*iso*-propyl	0.904	**50**	2,6-di- *tert*-butyl, 4-methyl	2.219

The N-methyl-phenyl urethanes in Figure 2 and Table 3 are used as insecticides. The molar concentration IC_{50} of the toxic induces, *in vitro*, 50% inhibition of acetyl cholinesterase. The value of bioactivity A = log ($1/IC_{50}$) is quoted in literature [27-29].

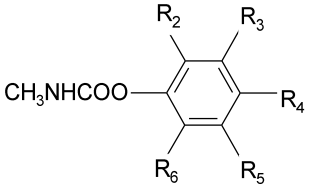

Figure 2 N-methyl-phenylurethane's derivatives

Table 3 Structure and observed bioactivity of urethanes

No.	**R_2**	**R_3**	**R_4**	**R_5**	**R_6**	**A**
1	H	*iso*-Pr	H	*iso*-Pr	H	7.48
2	H	Me	H	H	H	4.85
3	H	*iso*-Pr	H	H	H	6.47
4	H	t-Bu	H	H	H	6.40
5	H	Me	H	Me	H	5.22
6	H	Me	H	*iso*-Pr	H	7.25
7	Me	H	H	*iso*-Pr	H	5.70
8	O- *iso*-Pr	H	H	H	H	6.17
9	F	H	H	H	H	4.80
10	Cl	H	H	H	H	5.30
11	Br	H	H	H	H	5.66
12	I	H	H	H	H	6.10
13	H	H	Cl	H	H	3.62
14	H	NMe_2	H	H	H	5.10
15	H	Me	S-Me	Me	H	5.92
16	H	H	NMe_2	H	H	3.62
17	H	NMe_2	H	*iso*-Pr	H	6.72
18	H	*iso*-Pr	NMe_2	H	H	6.82
19	H	NMe_2	H	NMe_2	H	5.59

20	H	*iso*-Pr	NMe_2	H	Me	6.41
21	*iso*-Pr	H	NMe_2	Me	H	5.89
22	H	S-Me	H	H	H	5.16
23	H	H	S-Me	H	H	4.47
24	H	H	H	H	H	3.70
25	Me	H	H	H	H	3.85
26	H	H	Me	H	H	4.00
27	Et	H	H	H	H	4.89
28	H	Et	H	H	H	5.32
29	H	H	Et	H	H	4.42
30	*iso*-Pr	H	H	H	H	5.22
31	H	H	*iso*-Pr	H	H	4.16
32	*tert*-Bu	H	H	H	H	5.22
33	H	H	*tert*-Bu	H	H	5.82
34	*sec*-B	H	H	H	H	5.96
35	H	*sec*-Bu	H	H	H	6.80
36	H	H	*sec*-Bu	H	H	5.75
37	H	*sec*-Pe	H	H	H	6.96
38	cyclo-Pe	H	H	H	H	5.96
39	H	cyclo-Pe	H	H	H	5.82
40	H	H	cyclo-Pe	H	H	4.57
41	Pr	H	H	H	H	5.27
42	*iso*-Bu	H	H	H	H	5.64
43	H	F	H	H	H	4.07
44	H	H	F	H	H	3.64
45	H	Cl	H	H	H	4.30
46	H	Br	H	H	H	4.89
47	H	H	Br	H	H	4.00
48	H	I	H	H	H	5.16
49	H	H	I	H	H	4.06
50	Cl	Cl	H	H	H	4.32
51	Cl	H	Cl	H	H	4.85
52	Cl	H	H	Cl	H	4.30
53	Cl	H	H	H	Cl	2.89
54	H	Cl	Cl	H	H	4.72
55	H	Cl	H	Cl	H	4.92
56	NO_2	H	H	H	H	2.30
57	H	NO_2	H	H	H	2.70
58	H	H	NO_2	H	H	2.52
59	NO_2	Me	H	H	H	3.70
60	NO_2	H	Me	H	H	3.89
61	NO_2	H	H	Me	H	4.80
62	H	NO_2	Me	H	H	3.50
63	S-Me	H	H	H	H	5.40
64	H	Et	NO_2	H	H	3.70
65	H	*iso*-Pr	NO_2	H	H	5.55
66	H	H	S-Et	H	H	4.25
67	H	H	S-Pr	H	H	4.92

68	H	H	S- *iso*-Pr	H	H	5.05
69	H	H	S-Bu	H	H	5.43
70	cyclo-Hex	H	H	H	H	5.85
71	H	cyclo-Hex	H	H	H	5.70
72	H	H	cyclo-Hex	H	H	5.05
73	H	*iso*-Pr	S-Me	H	H	7.00
74	H	*iso*-Pr	H	H	S-Me	6.75
75	H	S-allyl	H	H	H	5.44
76	H	H	S-allyl	H	H	5.07

3. Methods and formulas

To identify the synergistic effect of two substituents X and Y, grafted in the *i* and *j* positions on a common skeleton, one needs the value of the bioactivity of the non-substituted molecule, the bioactivity of molecule substituted by X_i, the bioactivity of molecule substituted by Y_j and the bioactivity of molecule substituted by X_i *and* Y_j. If all these four molecules are available the synergy can by calculated using the intuitive formulas (2a/2b).

$$S = A_{X+Y} - A_0 - \left[A_X - A_0 + A_Y - A_0 \right] \quad (2a)$$

Therefore:

$$S = A_{X+Y} - A_X - A_Y + A_0 \quad (2b)$$

where:

A_{X+Y} is the bioactivity of the molecule, di-substituted by X_i *and* Y_j (or including the fragments X
and Y)
A_X is the bioactivity of the molecule, mono-substituted by X_i (or including X only)
A_Y is the bioactivity of the molecule, mono-substituted by Y_j (or including Y only)
A_0 is the bioactivity of the non-substituted molecule

The synergy S in formula (2b) is the synergy of two specific substituents grafted on certain positions or the synergy of two included molecular fragments. In theory, the value of S in formula (2b) is within the $\{-\infty, +\infty\}$ range, in practice this range is {- 3, + 3}.

Sometimes, one or the other of the above four molecules are missing and/or the molecules in the analyzed set include X_i, Y_j and/or other substituents (fragments). In such situations, we used the molar refractivity MR_i and MR_j of substituents X_i and Y_j (or percentages in weight P_i and P_j of the two *i* and *j* fragments) for *all* the molecules in the

group, the value of bioactivity of *all* the molecules in group, and the proposed statistical formula (3a).

$$S_{ij} = T_{whole} - T_{sum} \quad (3a)$$

where:

$T_{whole} = r^2_{i+j}$
r_{i+j} is the correlation between the values of the sum $MR_i + MR_j$ (or $P_i + P_j$) and the values of bioactivity
$T_{sum} = r^2_i + r^2_j - r^2_{ij} \bullet \min (r^2_i , r^2_j)$
r_i is the correlation between the values of MR_i (or P_i) and the values of bioactivity
r_j is the correlation between the values of MR_j (or P_j) and the values of bioactivity
r_{ij} is the correlation between the values of MR_i (or P_i)and the values of MR_j (or P_j) (intercorrelation)
$\min (r^2_i , r^2_j)$ is the minimum value in the pair (r^2_i , r^2_j)

The synergy S_{ij} in formula (3a) is the average synergy of *any two* substituents grafted on the common skeleton in positions *i* and *j* or the average synergy of *certain two* fragments. In formula (3a) we used the sum, not the product or other non-linear combinations, because of presumptive additive character of synergy. Here, the two analyzed variables MR_i and MR_j (or P_i and P_j) have the same physical attributes. If the two variables had different physical attributes the physical meaning of the sum is debatable.

The value of S_{ij} in formula (3a) is within the {- 1, + 1} range. Here, we used the corrected value in formula (3b) because the value computed by formula (3a) is frequently close to zero.

$$S_{correct} = S^{1/3} \quad (3b)$$

The value of linear correlation *r* is within the {- 1, + 1} range. If the value of correlation between the values of (any) *A* variable and the values of (any) *B* variable is far from zero, i.e. r^2 is close to 1, then there are three possibilities:

- *A* is cause of *B*
- *B* is cause of *A*
- *A* and *B* are the effects of cause *C*

The value and sign of *r* cannot identify the cause *A, B* or *C* of a certain phenomenon, but can measure the level and sign of cause-effect relation.

The sign and values of r_i, r_j and $S_{correct}$ in formula 3a/3b should be helpful in drug design because they emphasize the sign and level of the cause-effect relationship between the presence of certain substituents (fragments) and the bioactivity and synergy of these substituents (fragments). The pair of substituents (fragments) with r_i, r_j and $S_{correct}$ values very

different from zero should be valuable in drug design. Therefore, we propose the *DrugDesignFocus* index, see formula (4). The value of DDF is within the range {- 3, + 3}.

$$DDF = r_i + r_j + S_{\text{correct}} \tag{4}$$

Formula (3a) and Figure 3 attempt to emphasize the idea that *synergy is the difference between the whole and the sum of its parts.*

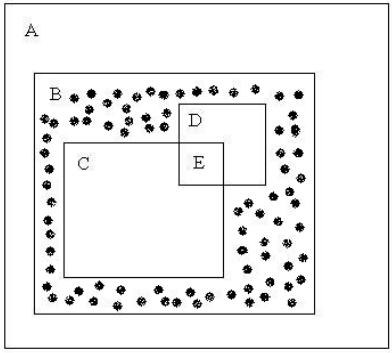

Figure 3 Graphical suggestion of synergy

The area of square A in Figure 3 represents the maximum possible value of T_{whole}, i.e. 1. The area of square B represents T_{whole}. The area of square C is r^2_i. The area of square D is r^2_j. The ratio (area of square E) / (area of square D) is r^2_{ij}. The area of square D represents here min (r^2_i , r^2_j). Therefore, the area of square E is equal with the value of the product r^2_{ij} • min (r^2_i , r^2_j). The synergy is represented by area of the airbrushed surface in square B.

The values for the molar refractivity of substituents MR are inferred from the observed (experimental) values of MR for various molecules and can be computed, as an additive property, using specific software [30]. For instance, the computed MR of ethane, divided by 2, is the MR of the CH_3 group, the computed MR of biphenyl, divided by 2, is the MR of the C_6H_5 group, the computed MR of nitro-methane, minus the MR of the CH_3 group, is the MR of the NO_2 group etc.

The values for the molar refractivity MR of substituents listed in Table 1, 2 and 3, used in our synergy and QSAR computations, are presented in Table 4. For molecule **39** in Table 1 we used, for substituents in positions 7 and 8, MR = 1.21.

Table 4 Molar refractivity of some substituents

Substituent	MR	Substituent	MR	Substituent	MR	Substituent	MR
H	0.11	C_6H_5	2.54	OCH_3	0.73	O-*iso*-propyl	1.65
F	0.08	*iso*-propyl	1.49	SCH_3	1.36	S-*iso*-propyl	2.30
Cl	0.57	*iso*-butyl	1.97	SC_2H_5	1.83	OH	0.27
Br	0.85	*sec*-butyl	1.96	SC_3H_7	2.30	NH_2	0.51
I	1.38	*tert*-butyl	1.95	SC_4H_9	2.77	NO_2	0.74
CH_3	0.57	*sec*-amyl	2.43	SC_6H_{13}	3.71		
C_2H_5	1.03	cyclo-pentyl	2.21	S-allyl	2.26		
C_3H_7	1.50	cyclo-hexyl	2.67	$N(CH_3)_2$	1.43		

Two linked heavy atoms (different from hydrogen) are included within the same fragment (together with the attached hydrogen atoms) if $B > 1.051$, where B is the bond order of the link [31]. Therefore, any suitable software can identify the fragments in an analyzed molecule [32]. For example, the group $–CH_2–O–CH_2–$ always includes three fragments. The group Ar–CO–O– includes, depending on conjugation and B value, can include only one fragment ArCOO, two fragments ArCO + O / Ar + COO or three fragments Ar + CO + O.

The Figure 4 presents, as example, the nine identified fragments in 5,3',4'-trihydroxy-6,7-dimethoxy-flavone, i.e. molecule **56** in Table 1, having the highest value of logK. In Figure 4 the hydrogen atoms and the single bonds ($B < 1.051$) are invisible. One can see two CH_3 fragments, two O fragments, and the C_6HO_2, CO, C_2H, C_6H_3OH and OH fragments.

After the identification of fragments, the percentage in weight of each fragment in each molecule can be computed. Some identified fragments are included only in one or two molecules of the analyzed group.

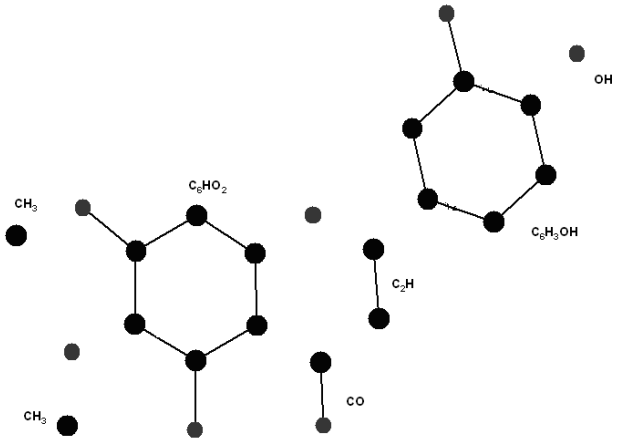

Figure 4 Nine identified molecular fragments in molecule **56** in Table 1

Before the QSAR computations we used the molecular mechanics program PCModel [33] for virtual building of the molecules and geometry optimization. Afterwards, we performed geometry optimization using the MOPAC software [34], more precisely the included semi-empirical quantum-mechanics PM6 method [35].

Afterwards, we used the PRECLAV software [36, 37] to calculate, for each molecule, 475 *whole molecule* descriptors, percentages of molecular fragments and 625 3D descriptors, specific to this program. Using the DESCRIPT software [34] we computed 44 aromaticity descriptors. In addition, we used 1664 descriptors, logP included, computed using DRAGON [38-40]. The sums $MR_i + MR_j$ were computed by hand, using the values in Table 4.

The PRECLAV software was used for statistical computations.

According to PRECLAV algorithm, the 'significant' descriptors fulfill criterion (5).

$$r^2 > \frac{\ln N}{N} \tag{5}$$

where

N is the number of molecules in the calibration set

r^2 is the square linear correlation descriptor/bioactivity

If the number of the 'significant' descriptors is huge (> 1000) then program selects only 1000 descriptors, according to correlations and intercorrelations.

The square linear intercorrelation r^2_{ij} of descriptors in any set has to be low enough, see criterion (6).

$$r_{ij}^2 < \frac{4}{N^{0.5}} \tag{6}$$

where

N is the number of molecules in the calibration set

r^2_{ij} is square linear intercorrelation descriptor *i* /descriptor *j*

PRECLAV computes type (7) multilinear QSARs obtained by a forward stepwise (selection) procedure of descriptors.

$$A = C_0 + \sum_{i=1}^{k} C_i \cdot D_i \tag{7}$$

where

A is (the value of) bioactivity

C_0 is the free term (intercept)

C_i are coefficients (weighting factors)

D_i are (the values of) significant descriptors

k is the number of descriptors

The maximum value of k in QSARs has to be low enough, see criterion (8).

$$k < \log_2 N \tag{8}$$

The square of Pearson linear correlation r^2 of observed/computed values, the Fisher function F, the standard error of estimation SEE and the quality function Q are criteria for the quality of prediction.

$$F = \frac{r^2}{1-r^2} \cdot \frac{N-p}{p} \tag{9}$$

$$SEE = \left[\frac{\sum \Delta^2}{N-1} \right]^{1/2} \tag{10}$$

$$Q = r^2 \left(1 - \frac{p}{N} \right) \tag{11}$$

where:

p is the number of descriptors
N is the number of molecules in the calibration set
Δ is the difference A_{obs}-A_{calc}

Within a forward stepwise selection procedure, the value of p in formula (11) increases and the value of Q increases, reaches a maximum and then decreases. Therefore, the value of Q and criterion (8) are criteria for stopping the calculations.

The total number of sets of descriptors is $10^9 - 10^{27}$. Actually, criteria (5), (6), (8), (11) and the 'forward selection' are used for a drastic selection of descriptors and sets of descriptors within a heuristic 'mathematical competition'.

The descriptors included in the best (by Q) QSAR are 'winners' in the 'mathematical competition' and are named 'predictors'. The description of bioactivity is not made from a certain predictor, but by using 'the best set' of predictors, *as a whole*.

In principle, certain descriptors can also act as predictors if the correlation with bioactivity (area of square C in Figure 3) is high enough *and* the intercorrelation with other descriptors (area of square E in Figure 3) is low enough. QSAR methodology can verify if certain intramolecular synergy descriptors are predictors. There is no conspicuous relationship between the quality of a certain descriptor as predictor and its synergy (airbrushed area in square B in Figure 3) with other descriptors.

The relative utility U of predictors was computed using formula (12). After calculating U, the values of U were normalized by the highest of them, i.e. the highest value becomes

1000. The predictors 'with high relative utility' (U > 600) are useful because they correlate well with the observed values of bioactivity and present low correlation with other predictors. Generally, each 'useful' predictor explains a lot of the bioactivity variation and also gives a unique description as compared to other predictors.

$$U = \frac{R^2 - r^2}{1 - r^2} \tag{12}$$

where:

R^2 is the square of linear correlation between the observed and computed values of bioactivity (using the QSAR including *p* predictors)
r^2 is the square of linear correlation between the observed and computed values of bioactivity (using the QSAR including *p-1* predictors, i.e. the equation *without* the analyzed predictor)

PRECLAV calculates the square of cross-validated linear correlation r^2_{CV} using the *Leave Half Out* method. However, this method is only applied here after ordering of molecules in the calibration set using the observed values of bioactivity. A low value (< 0.4) of r^2_{CV} means that '*the QSAR equation for molecules having high values of bioactivity and the QSAR equation for molecules having low values of bioactivity include the same descriptors, but very different weighting factors*'. Consequently, the function r^2_{CV} should be viewed here, in our opinion, as a measure of the homogeneity of the calibration set from the point of view of the structure-activity relationship, not as a drastic internal validation test.

4. Commented results

Table 5, (see Supplementary Data [47]) includes the value of synergy, calculated for the available substituents in flavones, using formula (2b), where $A_0 = 0$. The substituent 3-OCH_3 frequently seems to involve positive synergy while substituents 6-F and 6-OH, negative synergy.

Table 6, (see Supplementary Data) includes the value of synergy, calculated for the available substituents in phenols using formula (2b), where $A_0 = 0$. There is no very high synergy of substituents in the phenol derivatives analyzed.

Table 7, (see Supplementary Data) includes the value of synergy, calculated for the available substituents in urethanes using formula (2b), where $A_0 = 3.70$. There are two high synergy values (positive for the pair 2-NO_2/4-CH_3 and negative for the pair 2-Cl/3-Cl) of substituents in the urethane derivatives analyzed.

Table 8, (see Supplementary Data) includes the value of synergy and DDF, rounded to four decimals, calculated for flavones using formulas (3b) and (4). Almost all synergies in Table 8 are computed as antagonistic. There are eight positive DDF values in Table 8, i.e. there are eight possible interesting pairs (*i, j*) from the point of view of flavones' drug design. However, the positive values of DDF in Table 8 are low.

Table 9, (see Supplementary Data) includes the value of synergy and DDF, rounded to four decimals, calculated for phenols using formulas (3b) and (4). We noted in Table 9 six positive DDF values, i.e. six possible interesting (*i, j*) pairs from the point of view of phenols' drug design. The pair (5, 6) presents a high value of DDF, although there are no molecules in Table 2 having substituents in positions 5 *and* 6.

Table 10, (see Supplementary Data) includes the value of synergy and DDF, rounded to four decimals, calculated for urethanes using formulas (3b) and (4). There are six positive DDF values in Table 10, i.e. six possible interesting pairs (*i, j*) from the point of view of urethanes' drug design. The pair (2, 3) presents highest value of DDF although there are only two molecules in Table 3 having substituents in positions 2 *and* 3.

There are 26 identified molecular fragments and thus 325 fragment pairs in flavones, 16 identified fragments and 120 pairs in phenols and 22 identified fragments and 231 pairs in urethanes.

The highest (positive) and lowest (negative) four values for the synergy and DDF of fragments are presented in Table 11.

Table 11 maximum/minimum Synergy and DDF of fragments

Group	**Synergy/(Pair)**	**DDF(Pair)**
flavones	0.3373/($C_7H_4O_4$ and CO)	0.7313/(C_6H_4O and $C_7H_4O_4$)
	0.3207/($C_7H_4O_4$ and C_2H)	0.7233/($C_7H_4O_4$ and C_2H)
	0.3061/(C_6H_3 and C_8H_6O)	0.7211/(C_6H_4O and $C_7H_4O_2$)
	0.2906/(Br and Cl)	0.6861/($C_7H_4O_4$ and CO)
	-0.6000/(NO_2 and $C_7H_4O_4$)	-1.0126/(NO_2 and C_6H_4)
	-0.6258/(Br and C_6H_4O)	-1.0639/(C_6H_3 and Br)
	-0.6264/(C_6H_3 and C_6H_4O)	-1.1235/(NO_2 and Br)
	-0.6370/(NO_2 and C_6H_4O)	-1.1778/(NO_2 and C_6H_3)
phenols	0.4240/(C_6H_4 and C_6H_3)	1.0430/(Br and Cl)
	0.3494/(Br and Cl)	0.8898/(Cl and C_6H_2)
	0.3457/(HO and F)	0.8565/(C_6H_3O and C_6H_2)
	0.3316/(C_6H_4 and C_6H_5O)	0.8454/(Br and C_6H_2)
	-0.7018/(HO and C_6H_2)	-1.205/(HO and CH_3)

	-0.7089/(HO and Cl)	-1.2871/(HO and C_6H_5O)
	-0.7164/(HO and Br)	-1.4438/(HO and C_6H_3)
	-0.7506/(HO and C_6H_3O)	-1.4944/(HO and C_6H_4)
urethanes	0.3828/(CH_3 and CH_2)	1.1049/(CH_3 and CH_2)
	0.3215/(CH and C)	1.0825/(CH and C)
	0.2849/(C_6H_4 and C_6H_3)	0.8231/(CH_3 and C_2H_3)
	0.2455/(CH_3 and C_2H_3)	0.5669/(CH and CH_3)
	-0.7599/(CH and Cl)	-1.1408/(O and Cl)
	-0.7845/(CH and NHCO)	-1.1644/(NO_2 and C_6H_4)
	-0.8001/(CH_3 and NO_2)	-1.2320/(NO_2 and NHCO)
	-0.8283/(CH and NO_2)	-1.3581/(NO_2 and O)

The highest (positive) and lowest (negative) values of DDF for substituted positions, (see Table 8-10 in Supplementary Data) and for fragments in certain positions, (see Table 5-7 in Supplementary Data) suggest which fragments should be used and which positions should be substituted to increase/decrease the bioactivity, toxicity or other biochemical properties. Next we present our treatment of results.

In flavones, the OH *and* CH_3O chemical groups attached in positions 3 *and* 5 or 7 *and* 2' should increase the value of logK. The NO_2 *and* Br substituents used in positions 6 *and* 8 or 6 *and* 3' should decrease the value of logK.

In phenols, three substituents Br *and* Cl attached in positions 2, 4, 5 *and* 6 should increase the value of toxicity. One or two alkyl groups used in positions 3 *and* 5 should decrease the value of toxicity.

In urethanes, alkyl groups attached in positions 2 *and* 3 or 5 *and* 6 or 3 *and* 5 should increase the value of toxicity. The chemical groups NO_2, Cl *and* O should decrease the value of toxicity, regardless of substituted positions.

Before any QSAR calculations, one can observe in Tables 1, 2 and 3 few very marked *activity cliffs*, from the point of view of the structure-activity relationship.

For instance, the group **(38**, **56)** in Table 1 includes two molecules having very similar chemical structures, but very different values of logK. On the contrary, groups **(54, 65)** and **(20, 42)** in Table 1 include molecules having very different chemical structures, but very similar values of logK.

Molecules **1** and **48** in Table 2 include the same molecular fragments, but present very different values of toxicity. Groups **(22, 34)** in Table 2 include molecules having very different chemical structures, but very similar values of toxicity.

Groups (**12**, **48**, **49)** and (**53, 55**) in Table 3 include molecules having similar chemical structures, but different values of bioactivity.

Actually, the presence of cliffs emphasizes the violation of QSAR (opposite) axiom for some molecules, at least from the point of view of the chemical structure - bioactivity similarity. This violation can be the effect of intramolecular synergy and does not exclude the observance of QSAR axiom from the point of view of size, shape and other molecular features.

QSAR Studies Results

QSAR study #1

Calibration set: flavones in Table 1
Used descriptors: molar refractivity of substituents
percentage of molecular fragments

Number of outliers: 2 (molecules **40** and **42**)

After elimination of outliers:

Number of significant descriptors according to criterion (5): 134
The best QSAR according to criterion (11), see formula (7):

C_0 = - 0.2236
C_1 = - 0.0765
D_1 is the sum of percentages of NO_2 and $C_{10}H_6$ fragments (U = 1000)
C_2 = 12.6757
D_2 is $MR_7 \bullet MR_{4'}$ product (U = 786)
C_3 = - 0.0535
D_3 is the sum of percentages of C_8H_6O and Cl fragments (U = 708)
C_4 = 1.5295
D_4 is $MR_5 + MR_{2'}$ sum (U = 319)
C_5 = - 0.0286
D_5 is the sum of percentages of Br and $C_6H_2O_2$ fragments (U = 599)
C_6 = - 4.2619
D_6 is $MR_6 \bullet MR_{6'}$ product (U = 395)

Minimum correlation predictor/activity: for predictor D_6 (r^2 = 0.0611)
Maximum intercorrelation of predictors: for pair (D_5, D_6) (r^2 = 0.0507)

Predictive power of QSAR #1:
r^2 = 0.7625; F = 37.5; SEE = 0.503; Q = 0.7023; r^2_{CV} = 0.4278

The predictive quality of the best QSAR is satisfactory only after (and because of) the elimination of outlier molecules. In the absence of other descriptors, the molar refractivity of substituents is not quite suitable to describe the bioactivity of studied derivatives of flavone.

There are four synergy predictors in QSAR #1.

The homogeneity of molecules in Table 1, from the point of view of predictors in QSAR #1, is high enough.

QSAR study #2

Calibration set: flavones in Table 1
Used descriptors: molar refractivity of substituents
percentage of molecular fragments
PRECLAV *whole molecule* descriptors
PRECLAV 3D descriptors
DESCRIPT aromaticity descriptors
DRAGON descriptors

Number of outliers: 1 (molecule **31**)

After elimination of outlier:

Number of significant descriptors according to criterion (5): 676
The best QSAR according to criterion (11), see formula (7):

$C_0 = 27.6992$
$C_1 = - 0.3854$
D_1 is Variation coefficient of distances to geometric center (peripheral atoms) (U = 1000)
$C_2 = 0.0163$
D_2 is sum of percentages of C_6H_4O and $C_7H_4O_4$ fragments (U = 291)
$C_3 = - 2.3700$
D_3 is Complementary Information Content 0-order [38, 41] (U = 574)
$C_4 = 16.2003$
D_4 is $MR_7 \bullet MR_{4'}$ product (U = 833)
$C_5 = - 0.0381$
D_5 is sum of percentages of C_8H_6O and Cl fragments (U = 503)
$C_6 = - 0.4313$
D_6 is Gravitational index G2 [42] (U = 965)

Minimum correlation predictor/activity: for predictor D_6 ($r^2 = 0.0948$)
Maximum intercorrelation of predictors: for pair (D_1, D_3) ($r^2 = 0.4081$)

Predictive power of QSAR #2:
$r^2 = 0.8519$; F = 68.1; SEE = 0.393; Q = 0.7855; $r^2_{CV} = 0.2662$

The use of PRECLAV and DRAGON descriptors increases the predictive quality of QSAR.

There are two synergy predictors in QSAR #2.

The homogeneity of molecules in Table 1, from the point of view of predictors in QSAR #2, is low.

Figure 5 in Supplementary Data presents the scatter-plot of computed / observed values for QSAR #2 and Table 12 presents the computed / observed values for QSAR #2.

QSAR study #3

Calibration set: phenols in Table 2
Used descriptors: molar refractivity of substituents
percentage of molecular fragments

Number of outliers: 1 (molecule **38**)

After elimination of outlier:

Number of significant descriptors according to criterion (5): 73
The best QSAR according to criterion (11), see formula (7):

$C_0 = 4.2855$
$C_1 = - 0.2122$
D_1 is the percentage of OH fragment (U = 996)
$C_2 = - 0.0463$
D_2 is the sum of percentages of C_6H_3O and C_6H_4O fragments (U = 920)
$C_3 = - 0.0461$
D_3 is the sum of percentages of OH and C_6H_5O fragments (U = 1000)
$C_4 = - 0.0125$
D_4 is the sum of percentages of Br and I fragments (U = 504)

Minimum correlation predictor/activity: for predictor D_4 ($r^2 = 0.1404$)
Maximum intercorrelation of predictors: for pair (D_1, D_2) ($r^2 = 0.4610$)

Predictive power of QSAR #3:
$r^2 = 0.9055$; F = 107.8; SEE = 0.203; Q = 0.8316; $r^2_{CV} = 0.4520$

The predictive power of QSAR #3 is high.

The homogeneity of molecules in Table 2, from the point of view of predictors in QSAR #3, is high enough.

There are three synergy predictors in QSAR #3.

All molecules in Table 2 include one (and only one) hydroxyl group. Therefore, the presence of descriptor D_1 as predictor in QSAR #3 emphasizes the influence of molecular mass on the toxicity of phenols.

QSAR study #4

Calibration set: phenols in Table 2
Used descriptors: molar refractivity of substituents
Percentage of molecular fragments
PRECLAV *whole molecule* descriptors
PRECLAV 3D descriptors
DESCRIPT aromaticity descriptors
DRAGON descriptors

Number of outliers: 1 (molecule **38**)

After elimination of outlier:

Number of significant descriptors according to criterion (5): 1000
The best QSAR according to criterion (11), see formula (7):

C_0 = - 0.6691
C_1 = 1.3322
D_1 is the Broto-Moreau autocorrelation lag3 weighted by atomic polarizabilities [43] (U = 1000)
C_2 = 136.3052
D_2 is the Sum of attraction forces on probe atom #65 (U = 952)
C_3 = - 0.0289
D_3 is the Difference between walks and paths on topological graph [44] (U = 553)

Minimum correlation predictor/activity: for predictor D_3 (r^2 = 0.1706)
Maximum intercorrelation of predictors: for pair (D_1, D_3) (r^2 = 0.2998)

Predictive power of QSAR #4:
r^2 = 0.9493; F = 286.8; SEE = 0.149; Q = 0.8911; r^2_{CV} = 0.2567

Molecules **36, 37** and **38** in Table 2 are very similar from any point of view. However, the toxicity of molecule **38** is much higher (violation of *QSAR axiom*). Maybe the observed (experimental) value of toxicity for this molecule is erroneous and explains the outlier feature of this molecule from point of view of QSAR #3 and QSAR #4.

The use of PRECLAV and DRAGON descriptors increases the predictive quality of QSAR.

There are no synergy predictors in QSAR #4.

The homogeneity of molecules in Table 2, from the point of view of predictors in QSAR #3 and QSAR #4, is low.

Figure 6 in Supplementary Data presents the scatter-plot of computed / observed values for QSAR #4 and Table 13 presents the computed / observed values for QSAR #4.

QSAR study #5

Calibration set: urethanes in Table 3
Used descriptors: molar refractivity of substituents
percentage of molecular fragments

Number of outliers: 1 (molecule **53**)

After elimination of outlier:

Number of significant descriptors according to criterion (5): 72
The best QSAR according to criterion (11), see formula (7):

$C_0 = 5.4436$
$C_1 = 0.0903$
D_1 is the sum of percentages of C and CH fragments (U = 432)
$C_2 = - 0.0607$
D_2 is the sum of percentages of NO_2 and NHCO fragments (U = 1000)
$C_3 = 0.2295$
D_3 is $MR_3 + MR_5$ sum (U = 220)
$C_4 = 0.3113$
D_4 is $MR_2 + MR_3$ sum (U = 398)
$C_5 = - 0.0179$
D_5 is the sum of percentages of NO_2 and Cl fragments (U = 212)
$C_6 = 0.0170$
D_6 is the sum of percentages of CH_3 and C_6H_3 fragments (U = 536)

Minimum correlation predictor/activity: for predictor D_6 ($r^2 = 0.1009$)
Maximum intercorrelation of predictors: for pair (D_2, D_5) ($r^2 = 0.2888$)

Predictive power of QSAR #5:
$r^2 = 0.8129$; F = 50.0; SEE = 0.480; Q = 0.7478; $r^2_{CV} = 0.3697$

The predictive power of QSAR #5 is high enough.

All predictors are synergy predictors.

The homogeneity of molecules in Table 3, from the point of view of predictors in QSAR #5, is somehow low.

QSAR study #6

Calibration set: urethanes in Table 3
Used descriptors: molar refractivity of substituents
Percentage of molecular fragments
PRECLAV *whole molecule* descriptors
PRECLAV 3D descriptors
DESCRIPT aromaticity descriptors
DRAGON descriptors

Number of outliers: 0

Number of significant descriptors according to criterion (5): 1000
The best QSAR according to criterion (11), see formula (7):

C_0 = - 3.1617
C_1 = 0.0477
D_1 is the Rugosity of molecular surface (U = 624)
C_2 = - 0.1060
D_2 is the sum of percentages of NO_2 and NHCO fragments (U = 1000)
C_3 = 3.5609
D_3 is 3D-MoRSE signal #30 weighted by atomic polarizabilities [45] (U = 265)
C_4 = 19.4393
D_4 is R maximal autocorrelation of lag7 weighted by electronegativities [46] (U = 256)
C_5 = - 60.9592
D_5 is the Sum of attraction forces on probe atom #38 (U = 807)
C_6 = - 1.2350
D_6 is the Complementary Information Content 4-order [38, 41] (U = 420)

Minimum correlation predictor/activity: for predictor D_6 (r^2 = 0.1554)
Maximum intercorrelation of predictors: for pair (D_1, D_4) (r^2 = 0.4563)

Predictive power of QSAR #6:
r^2 = 0.8758; F = 82.2; SEE = 0.399; Q = 0.8066; r^2_{CV} = 0.4295

The rugosity of molecular surface (predictor D_1) is the variance of distances of peripheral atoms to surface of the circumscribed ellipsoid.

The use of PRECLAV and DRAGON descriptors increases the predictive quality of QSAR.

There is one synergy predictor in QSAR #6.

The homogeneity of molecules in Table 3, from the point of view of predictors in QSAR #6, is somehow high.

Figure 7 in Supplementary Data presents the scatter-plot of computed / observed values for QSAR #6 and Table 14 presents the computed / observed values for QSAR #6.

5. Conclusions

The following can be concluded from the results of synergy, QSAR calculations, and the high diversity of the analyzed molecules.

There is an intramolecular reciprocal influence of fragments (or substituents grafted on a common skeleton) and subsequent bioactivity depends on this synergy or antagonism.

The intramolecular synergy of substituents and molecular fragments can be computed using the molar refractivity MR of substituents, the percentages in weight of fragments, and the value of bioactivity of molecules.

The proposed formula is useful in drug design for the identification of the substituted positions and fragments having a value of the DrugDesignFocus index very different from zero.

The QSAR methodology emphasizes the existence of intramolecular synergistic or antagonistic effects of substituents (fragments) only if the synergy descriptors are predictors. The proposed synergy descriptors (sums of MRs or percentages of fragments) are frequently part of the usual QSAR predictors. Some proposed synergy descriptors are predictors above all because of their low intercorrelation with other predictors, not because of the high value of synergy or DrugDesignFocus index.

References

[1] according to ChemIDplus database, http://www.chem.sis.nlm.nih.gov/chemidplus/ accessed in February 2015.

[2] R. B. Fuller, E. J. Applewhite, *Synergetics: Explorations in the Geometry of Thinking*, Macmillan, New York, 1975.

[3] M. E. Henderson, W. L. Miranker, Synergy in parallel algorithms, *Parallel Computing* **11** (1989) 17–35.

[4] Q. Liu, H. Zhou, L. Liu, X. Chen, R. Zhu, Z. Cao, Multi-target QSAR modelling in the analysis and design of HIV-HCV co-inhibitors:an in-silico study, *BMC Bioinformatics* **12** (2011) #294.

[5] R. Singh, M. E. Sobhia, Synergistic application of target structure-based alignment and 3D-QSAR study of protein tyrosine phosphatase 1B (PTP1b) inhibitors, *Med. Chem. Res.* **20** (2011) 714–725.

[6] P. Agarwal, S. Dubey, Z. Mirza, A. K. Pathak, 2D-QSAR study of phenothiazine derivatives as antitubercular agents, *Int. J. Pharm. Biomed. Res.* **4** (2013) 114–119.

[7] A. Wani, S. Mahajan S., Synthesis, Antibacterial activity and preliminary QSAR studies of diaryl sulfones, *Am. J. PharmTech Res.* **4** (2014) 1056–1066.

[8] T. C. Chou, P. Talalay, Quantitative analysis of dose-effect relationship: the combined effects of multiple drugs or enzyme inhibitors, *Adv. Enzyme Regul.* **22** (1984) 27–55.

[9] E. Kosman, Y. Cohen, Procedures for calculating and differentiating synergism and antagonism in action of fungicide mixtures, *Phytopathology* **11** (1996) 1263–1272.

[10] T. C. Chou, Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies, *Pharm. Rev.* **58** (2006) 621–681.

[11] J. C. Boik, B. Narasimhan, An R package for assessing drug synergism/antagonism, *J. Stat. Softw.* **34** (2010) 1–18.

[12] T. C. Chou, Drug combination study and their synergy quantification using the Chou-Talalay method, *Cancer Res.* **70** (2010) 440–446.

[13] R. J. Tallarida, Quantitative methods for assessing drug synergism, *Genes Cancer* **2** (2011) 1003–1008.

[14] Y. Zhang, P. Smolen, D. A. Baxter, J. H. Byrne, Computational analysis of synergism in small molecular network motifs, *PLoS Comput. Biol.* **10** (2014) #e1003524.

[15] D. Tian, Z. Lin, Quantitative structure activity relationships (QSAR) for binary mixtures at non-equitoxic ratios based on toxic ratios-effects curves, *Dose Response* **11** (2013) 255–269.

[16] S. S. Mahajan, V. R. Kamath, S. S. Ghatpande, Synergistic antimalarial activity of ketones with rufigallol and vitamin C, *Parasitology* **11** (2005) 459–466.

[17] S. C. Peterangelo, P. G. Seybold, Synergistic interactions among QSAR descriptors, *Int. J. Quantum Chem.* **96** (2004) 1–9.

[18] H. F. Freitas, M. P. Postigo, A. D. Andricopulo, M. S. Castilho, Descriptor-and fragment-based QSAR models for a series of Schistosoma mansoni purine nucleoside inhibitors, *J. Braz. Chem. Soc.* **22** (2011) 1718–1726.

[19] M. A. Efroymson, *Multiple Regression Analysis*, Wiley, New York, 1960.

[20] R. R. Hocking, The analysis and selection of variables in linear regression, *Biometrics* **32** (1976) 1–49.

[21] N. Draper, H. Smith, *Applied Regression Analysis*, Wiley, New York, 1981.

[22] L. Richter, C. de Graaf, W. Sieghart, Z. Varagic, M. Mörzinger, I. J. de Esch, G. F. Ecker, M. Ernst, Diazepam–bound $GABA_A$ receptor models identify new benzodiazepine binding–site ligands, *Nat. Chem. Biol.* **8** (2012) 455–464.

[23] P. R. Duchowicz, M. G. Vitale, E. A. Castro, J. C. Autino,G. P. Romanelli, D. O. Bennardi, QSAR modeling of the interaction of flavonoid with GABA(A) receptor, *Eur. J. Med. Chem.* **43** (2008) 1593–1602.

[24] V. K. Agrawal, B. Shaik, P. V. Khadikar, S. Singh, Modeling of the interaction of flavanoids with GABA (A) receptor using PRECLAV (Property-evaluation by class variables), *Pharmacology & Pharmacy* **2** (2011) 271–281.

[25] L. H. Hall, T. A. Vaughn, QSAR of phenol toxicity using electrotopological state and kappa shape indices, *Med. Chem. Res.* **7** (1997) 407–416.

[26] K. Roy, G. Ghosh G., Introduction of extended topochemical atom (ETA) indices in the valence electron mobile (VEM) environment as tools for QSAR/QSPR studies, *Int. Elec. J. Mol. Des.* **2** (2003) 599–620.

[27] R. L. Metcalf, T. R. Fukuto, Effects of chemical structure on intoxication and detoxication of phenyl-N-methyl carbamates in insects, *J. Agr. Food. Chem.* **13** (1965) 220–231.

[28] R. L. Metcalf, T. R. Fukuto, C. F. Wilkinson, M. H. Fahmy, A. E. Azziz, E. R. Metcalf, Mode of action of carbamate synergists, *J. Agr. Food. Chem.* **14** (1966) 555–562.

[29] R. L. Metcalf, T. R. Fukuto, Effects of molecular structure upon anticholinesterase and insecticidal activity of substituted phenyl N-methylcarbamates, *J. Agr. Food. Chem.* **15** (1967) 1022–1029.

[30] for instance ACD/ChemSketch software,
see http://www.acdlabs.com/resources/freeware/chemsketch/,
accessed in January 2015.

[31] L. Tarko, Aromatic molecular zones and fragments, *ARKIVOC* **XI** (2008) 24–45.

[32] L. Tarko, Virtual fragmentation of molecules and similarity evaluation, *Rev. Chim. (Bucuresti)* **55** (2004) 539–546.

[33] J. J. Gajewski, K. E. Gilbert, PCModel; Serena Software, Box 3076, Bloomington, IN, USA.

[34] http://www.openmopac.net/, accessed in October 2014.

[35] J. J. P. Stewart, Optimization of parameters for semiempirical methods: modification of NDDO approximations and application to 70 elements, *J. Mol. Model.* **13** (2007) 1173–1213.

[36] L. Tarko, C. T. Supuran, QSAR studies regarding sulfamate and sulfamide inhibitors targeting human carbonic anhydrase isozymes I, II, IX and XII, *Bioorg. & Med. Chem.* **21** (2013) 1404–1409.

[37] PRECLAV and DESCRIPT programs are available from Center of Organic Chemistry – Bucharest – Romanian Academy; ltarko@cco.ro; tarko_laszlo@yahoo.com .

[38] DRAGON program is available from Talete srl., via V. Pisani, 13-20124, Milano, Italy; http://www.talete.mi.it, November 2014.

[39] I. Moriguchi, S. Hirono, I. Nakagome, H. Hirano, Comparison of reliability of logP values for drugs calculated by several methods, *Chem. Pharm. Bull.* **42** (1994) 976–978.

[40] a) A. K. Ghose, G. M. Crippen, Atomic physicochemical parameters for three-dimensional structure-directed quantitative structure-activity relationships I. Partition coefficients as a measure of hydrophobicity, *J. Comput. Chem.* **7** (1986) 565–577. b) A. K. Ghose, V. N. Viswanadhan, J. J. Wendoloski, Prediction of hydrophobic (lipophilic) properties of small organic molecules using fragmental methods: An analysis of ALOGP and CLOGP methods, *J. Phys. Chem. A* **102** (1998) 3762–3772.

[41] D. Bonchev, O. Mekenyan, N. Trinajstic, Isomer discrimination by topological information approach, *J. Comput. Chem.* **2** (1981) 127–148.

[42] A. R. Katritzky, L. Mu, V. S. Lobanov, M. Karelson, Correlation of boiling points with molecular structure. A training set of 298 diverse organics and a test set of 9 simple inorganics, *J. Phys. Chem.* **100** (1996) 10400–10407.

[43] P. Broto, G. Moreau, C. Vandycke, Molecular structures: Perception, autocorrelation descriptor and SAR studies. Autocorrelation descriptor, *Eur. J. Med. Chem.* **19** (1984) 66–70.

[44] see DRAGON documentation and
G. Rücker, C. Rücker, Counts of all walks as atomic and molecular descriptors, *J. Chem. Inf. Comput. Sci.* **33** (1993) 683–695.

[45] J. H. Schuur, P. Selzer, J. Gasteiger, The coding of the three-dimensional structure of molecules by molecular transforms and its application to structure-spectra correlation and study of biological activity, *J. Am. Chem. Soc.* **36** (1996) 334–344.

[46] V. Consonni, R. Todeschini, M. Pavan, P. Gramatica, Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 2. Application of the novel 3D molecular descriptors to QSAR/QSPR studies, *J. Chem. Inf. Comput. Sci.* **42** (2002) 693–705.

[47] Supplementary Data are available from author by E-mail or from http://www.cco.ro/articles/tarko/TarkoMATCHSupplementaryData.pdf .