

Universitatea Babeş-Bolyai
Facultatea de Chimie si Inginerie Chimica
Departamentul de Chimie

STUDIUL QSAR AL UNOR COMPUȘI CU OXIGEN

Rezumat

CONDUCĂTOR ȘTIINȚIFIC
Prof.Dr. MIRCEA V. DIUDEA

DOCTORAND
HARȘA ALEXANDRA MARIA

Cluj-Napoca
2015

Comisia de doctorat:

Președinte:

Prof. univ. dr. MUREȘAN Liana-Maria-Universitatea Babeș-Bolyai, Cluj-Napoca

Conducător științific :

Prof.Univ.Dr. MIRCEA V. DIUDEA -Universitatea Babes-Bolyai, Cluj-Napoca

Referenți :

Conf.Dr. BUNEA Andrea -Universitatea de Științe Agricole și Medicină Veterinară, Cluj

Conf. univ. dr. MEDELEANU Mihai -Universitatea Politehnica, Timișoara

Prof. univ. dr. PÂRV Bazil -Universitatea Babeș - Bolyai, Cluj-Napoca

Data susținerii publice: 16 Octombrie 2015

Cuprins

Mulțumiri	4
INTRODUCERE	5
II.CONTRIBUȚII PROPRI	6
1. Studiul QSAR și de andocare moleculară a derivaților de flavonoide pe baza clusterilor de similaritate	6
1.2.Andocarea moleculară a derivaților de flavonoide.....	14
2. Studiul QSAR a derivaților de testosteron.....	18
Sarcină parțială (cazul 2).....	22
3.Studiul QSAR și de andocare moleculară a derivaților de antrachinonă.....	26
Sarcină parțială (log P).....	29
Sarcină parțială (LD50 bazat pe similaritate topologică).....	31
4.Studiul QSAR și de andocare moleculară ale derivaților de resveratrol	32
4.1.Studiul QSAR ale derivaților de resveratrol.....	32
4.2.Studiul andocări moleculare	37
CONCLUZII GENERALE.....	38
BIBLIOGRAFIE.....	38



Investește în oameni !

FONDUL SOCIAL EUROPEAN

Programul Operațional Sectorial Dezvoltarea Resurselor Umane 2007 – 2013

Axa prioritară 1. „Educația și formarea profesională în sprijinul creșterii economice și dezvoltării societății bazate pe cunoaștere”

Domeniul major de intervenție 1.5. „Sprijin pentru doctoranzi și cercetători postdoctorat”

Titlul proiectului: Programe doctorale și postdoctorale - suport pentru creșterea competitivității cercetării în domeniul Științelor exacte

Cod Contract: POSDRU/159/1.5/S/ 137750

Beneficiar- Universitatea „Alexandru Ioan Cuza” din Iași, Partener 2-Universitatea Babeș-Bolyai

Mulțumiri

Îi multumesc conducătorului de doctorat, prof. dr. Mircea V. Diudea, pentru ajutorul și îndrumarea oferită în timpul activității mele pentru îndeplinirea acestui obiectiv.

Îmi exprim întreaga considerație și mulțumire față de membrii comisiei de doctorat pentru amabilitatea cu care au acceptat să facă parte din comisia de doctorat și pentru evaluarea tezei.

Aș dori să le mulțumesc comisiei de îndrumare: Conf. univ. dr. MAJDIK Cornelia , Conf. univ. dr. KATONA Gabriel ,Prof. univ. dr. ing. DĂRĂBANȚU Mircea .

Exprim de asemenea o deosebită recunoștință fratelui meu Aurel, care m-a ajutat la instalarea programelor și surori Teodora pentru sprijinul acordat. Mulțumesc, de asemenea, la grupurile de cercetare unde mi-am desfășurat activitatea de cercetare, precum și colegilor de facultate pentru prezența lor, sfaturile și sprijinul oferite atunci când a fost necesar. Aș dori să-mi exprim recunoștința parintilor mei pentru susținerea și sprijinul oferit în acești ani.

Această lucrare este rezultatul cercetării doctorale ce a fost posibilă prin sprijinul

financiar oferit prin Programul Operațional Sectorial Dezvoltarea Resurselor Umane 2007-2013, cofinanțat prin Fondul Social European, în cadrul proiectului POSDRU/159/1.5/S/137750, cu titlul “Programe doctorale și postdoctorale – suport pentru creșterea competitivității cercetării în domeniul Științelor exacte”.

Cuvinte cheie: QSAR, flavonoide, resveratrol, antrachinone, testosteron, similaritate, indici topologici, docking molecular, leave-one-out, hiper molecula, log P, LD50.

INTRODUCERE

În cadrul tezei de doctorat sunt prezentate aspectele teoretice și contribuțiile proprii referitoare la studiul QSAR al compușilor organici cu grupări cu oxigen, implicând conceptul de hiper moleculă, o procedură similară cu cea de aliniere a medicamentelor la receptorul biologic.

Astfel teza de doctorat face parte dintr-un domeniu important de cercetare în chimie, aducând contribuții la dezvoltarea conceptelor teoretice privind relațiile structură-proprietăți biologice ale moleculelor, cu aplicații în industria farmaceutică la descoperirea de noi medicamente pe baza similarității moleculare astfel reducându-se costul de producție.

Studiul QSAR și de andocare moleculară s-a realizat pe următoarele clase de compuși: flavonoide, testosteron, antrachinone și resveratrol datorită proprietăților biologice deosebite, aceste clase de compuși au fost descărcate din baza de date PubCHEM. Pentru setul de date s-a calculat indici topologici cu ajutorul programului TOPO Cluj, iar optimizarea structurilor s-a realizat cu programul Gaussian.

Aceste metode QSAR/QSPR se bazează pe analiza statistică de corelație dintre proprietăți cunoscute și descriptori topologici pentru care nu există relații funcționale cunoscute. Astfel activitatea biologică sau toxicitatea, reactivitatea unei molecule poate fi determinată dependent de diferitele proprietăți după legi mai simple (regresie liniară) sau mai complexe (regresie multiplă). Astfel se poate presupune o dependență multiliniară a reactivității, activității biologice, toxicității de anumite proprietăți chimice, fizice, biologice măsurabile

sau calculabile. Se pot stabili astfel pe cale statistică ce proprietăți au influență statistică semnificativă asupra activității biologice (toxicității) și ce proprietăți pot fi excluse, fiind nerelevante acestea sunt eliminate.

Modelarea QSAR prin aplicarea algoritmului propus de Diudea, iar proprietățile supuse modelării sunt: log P și LD50. Modelarea s-a realizat pe setul școală, iar cele mai bune modele care descriu activitatea biologică și toxicitatea a acestui set de molecule este validat prin procedura leave-one-out, în setul extern test și printr-o nouă versiune de predicție folosind clusteri de similaritate moleculară.

Moleculele care sunt similare din punct de vedere structural au proprietăți similare acest concept este folosit în descoperirea de noi medicamente, după compararea moleculelor .

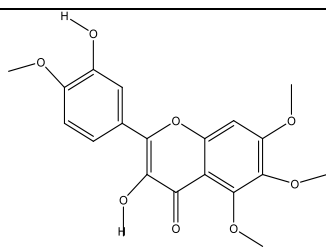
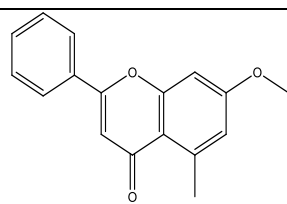
În studiul modelării moleculare, andocarea (docking) este metoda care prezice modul de legare a proteinei de ligand pentru a forma un complex stabil. Programul de andocare utilizat este AutoDock VINA , iar pe baza energiei de legare s-a construit farmacoforul.

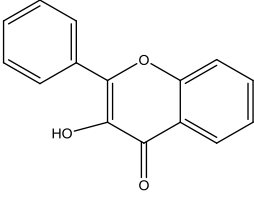
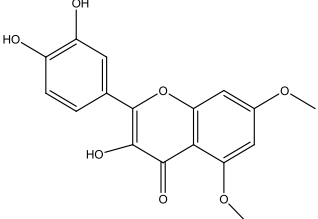
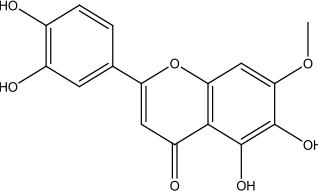
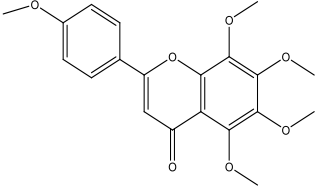
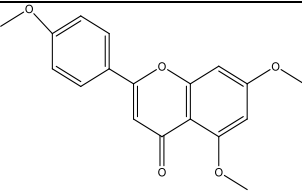
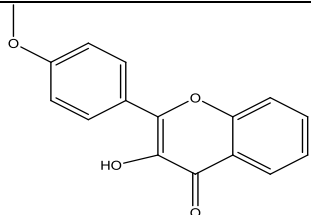
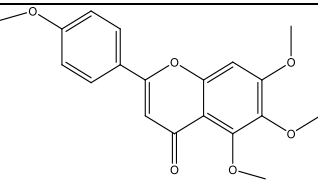
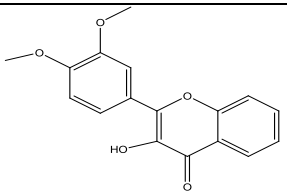
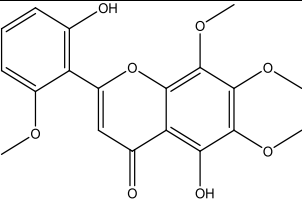
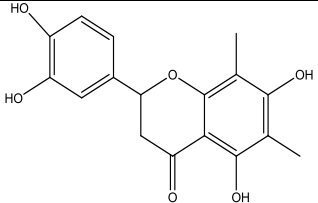
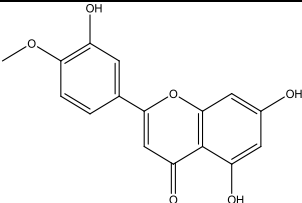
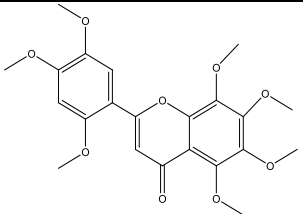
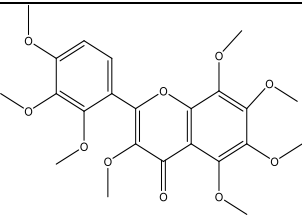
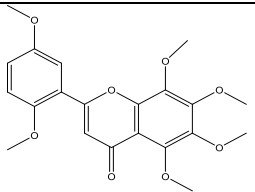
II.CONTRIBUȚII PROPRI

1. Studiul QSAR și de andocare moleculară a derivaților de flavonoide pe baza clusterilor de similaritate

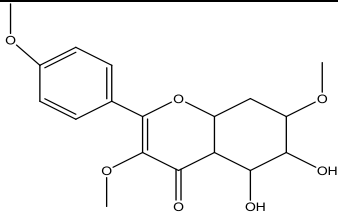
Un set de 40 de flavonoide au fost descărcate din baza de date PubChem [93](Tabelul 1) a fost împărțit în două seturi: set de școală (30 molecule) și set test (10 molecule), luate aleatoriu. Proprietățile alese pentru modelare sunt log P-(calcularea) coeficientului parțial dintre n-octanol și apă, măsură a hidrofobicității, implicat în transportul pasiv a moleculelor de medicamente prin membrana celulară (vezi tabelul 1) și LD50 (pe șoareci, administrată oral)[94].

Tabelul 1. Structura moleculelor de flavonoide și log P (luate din PubChem)

	Structura	log P	Structura	log P	
1		2.8	2		3.5

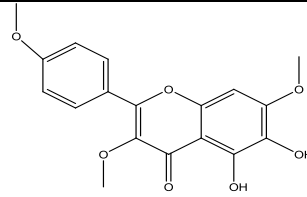
3		3.4	4		1.6
5		2.3	6		3
7		2.2	8		3.4
9		3.1	10		3.3
11		2.9	12		2.8
13		1.7	14		3
15		3.2	16		3

17



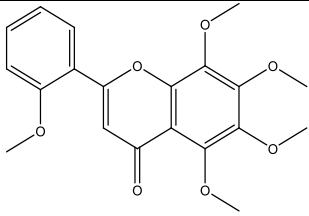
2.9

18



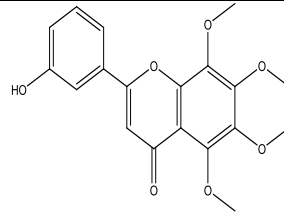
3.1

19



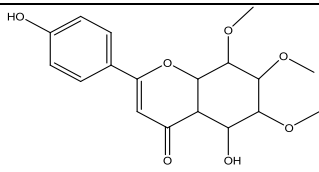
3

20



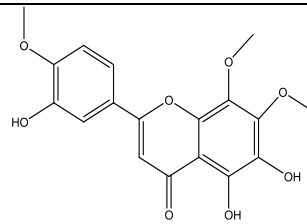
2.7

21



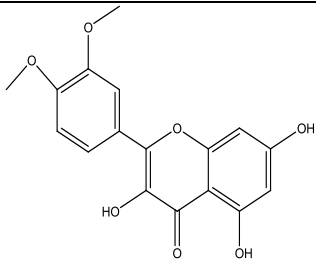
3.1

22



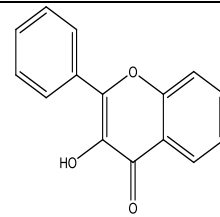
2.6

23



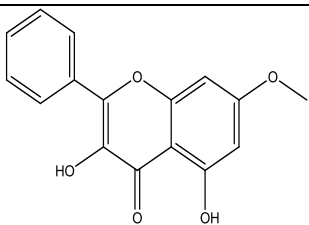
2.2

24



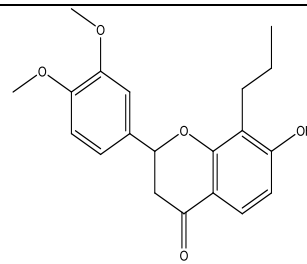
2.6

25



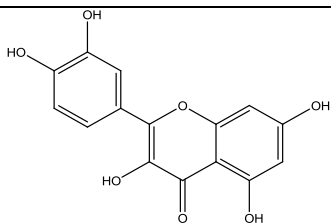
2.6

26



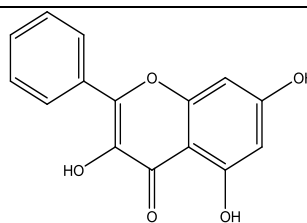
3.8

27

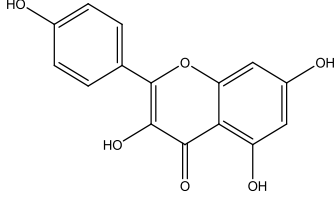
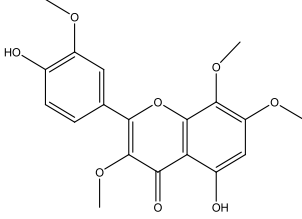
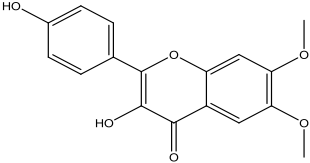
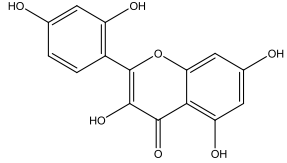
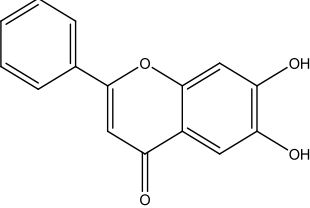
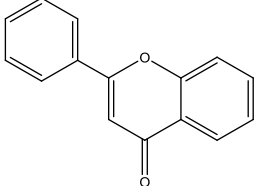
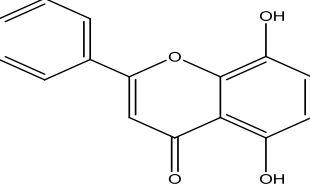
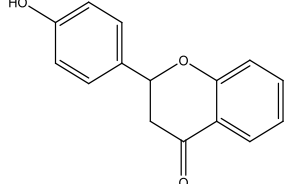
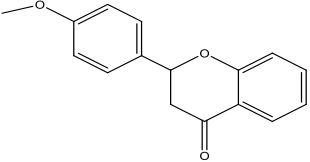
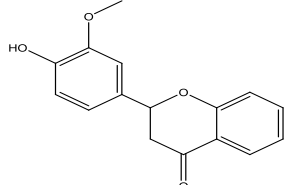
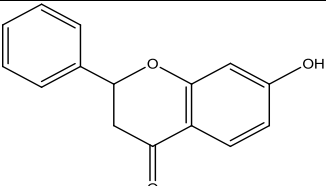
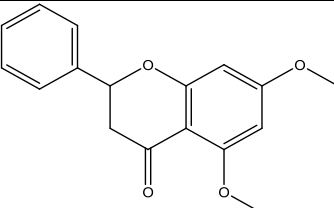


1.5

28



2.3

29		1.8	30		3.1
31		2.6	32		1.5
33		1.7	34		3.2
35		2.1	36		2.8
37		3.1	38		2.5
39		2.7	40		2.8

(b) Validare externă

Valoarea log P pentru setul test a derivaților de flavonoide este calculat folosind ecuația 19 din Tabelul 4. Datele sunt prezentate în Tabelul 6 și corelația monovariantă este:

$\log P = 0.804 \times \log P_{\text{calc.}} + 0.983$; $n=15$; $R^2=0.883$; $s=0.219$; $F=98.164$ este dată în Fig. 9. Numărul descriptorilor celui mai bun model este limitat la cinci, conform Topoliss și Costello [110].

Tabelul 6. Valorile calculate log P pentru moleculele din setul test (Tabelul 1)

Molecule	log P	log P _{calc.}
2	3.5	3.935
4	1.6	2.357
5	2.3	2.578
11	2.9	3.386
13	1.7	2.375
16	3	3.309
20	2.7	3.238
21	2.8	2.948
22	2.6	2.908
24	2.6	3.334
27	1.5	2.217
29	1.8	2.131
32	1.5	2.407
35	2.1	2.560
38	2.5	3.125

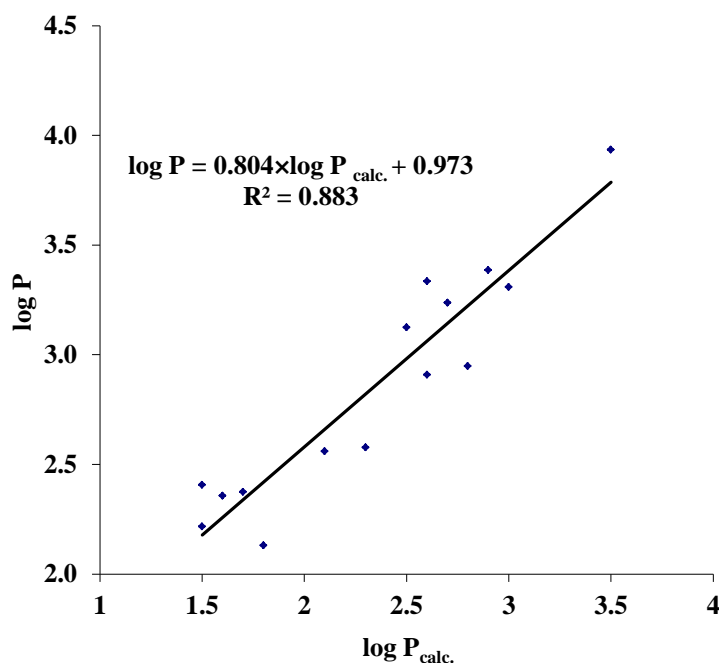


Figura 9. Reprezentarea grafică a log P vs. log P_{calc.} pentru setul test (validarea externă)

(c) Validarea pe clusteri de similaritate

Validarea s-a efectuat prin calcularea log P pentru moleculele din setul test prin folosirea clusterilor de similaritate: fiecare dintre cele 15 molecule este lider în propriul cluster, selectat prin similaritate 2D din cele 25 de structuri din setul școală initial. Valoarea log P_{calc.} a fost calculate cu 15 ecuații noi (liderul fiind lăsat afară) cu aceiași descriptori ca în ecuația 11, Tabelul 4. Datele sunt prezentate în Tabelul 7 și corelația monovariantă este: $\log P = 0.740 \times \log P_{\text{calc.}} + 1.037$; $n=15$; $R^2=0.907$; $s=0.195$; $F=126.972$ este dată în Figura 10.

De asemenea, validarea poate fi efectuată prin calcularea log P pentru moleculele din setul test prin folosirea clusterilor de similaritate și utilizarea descriptorilor ca în ecuația 19, Tabelul 4. Datele sunt prezentate în Tabelul 8 și corelația monovariantă este: $\log P = 0.7898 \times \log P_{\text{calc.}} + 0.715$; $n=15$; $R^2=0.932$; $s=0.167$; $F=176.737$ este dată în Figura 11.

Chiar și numărul descriptorilor per probă este mai mare decât în cazul validării externe, libertatea de alegere a structurilor în general este compensată de o predicție mai bună: comparând $R^2=0.883$ (validare externă) și $R^2=0.932$ (validarea prin clusteri de similaritate), cu același număr de descriptori. În cazul datelor din Tabelul 7 este îndeplinit criteriul dat de Topliss și Costello [27] în cazul utilizării a trei variabile urmate din nou de predicție ($R^2 = 0.907$) aceasta fiind mai bună decât în cazul validării externe clasice al modelului. O măsură de similaritate, care cuantifică gradul de asemănare structurală între structura țintă și fiecare dintre structurile din baza de date, se bazează pe suprapunerea maximă a substructurilor moleculare, la nivel topologic (2D). Volumul fiecărui din cele 15 clusteri este de 12 molecule.

Tabelul 7. Valorile calculate a log P prin clusterii de similaritate, pentru moleculele din setul test ca în ecuația 11

Molecule	log P	log P _{calc.}
2	3.5	3.77
4	1.6	2.33
5	2.3	2.58
11	2.9	3.17
13	1.7	2.35
16	3	3.18
20	2.7	3.01
21	2.8	2.91
22	2.6	2.83
24	2.6	3.18
27	1.5	2.13
29	1.8	2.20
32	1.5	2.27
35	2.1	2.56
38	2.5	3.13

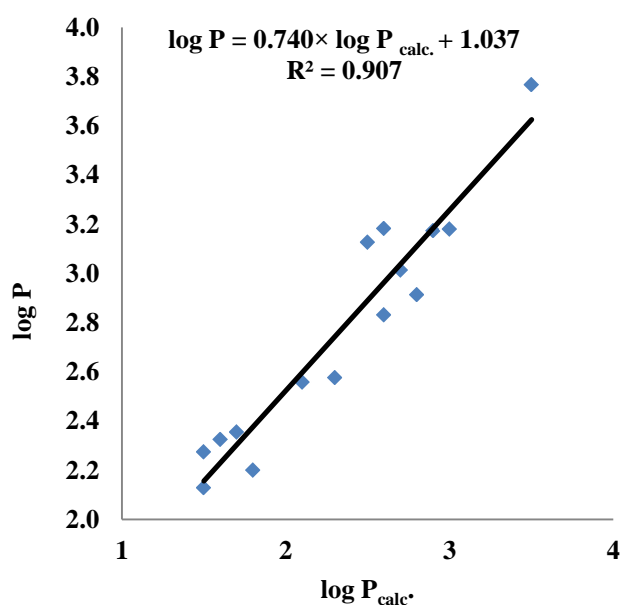


Figura 10. Graficul log P vs. log P_{calc.} prin clusteri de similaritate

Tabelul 8. Valorile calculate a log P prin clusterii de similaritate, pentru moleculele din setul test ca în ecuația 19

Molecule	log P	log P _{calc.}
2	3.5	3.469
4	1.6	2.036
5	2.3	2.353
11	2.9	2.898
13	1.7	1.974
16	3	3.214
20	2.7	3.005
21	2.8	2.885
22	2.6	2.861
24	2.6	2.739
27	1.5	1.748
29	1.8	2.057
32	1.5	2.124
35	2.1	2.554
38	2.5	2.530

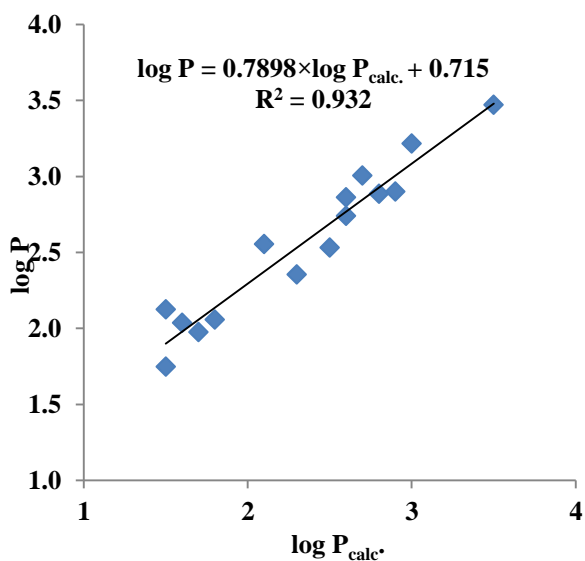


Figura 11. Graficul log P vs. log P_{calc.} prin clusterii de similaritate

○ **Cazul LD50**

(b) Validarea externă

Valorile LD50 pentru moleculele din setul test au fost calculate prin folosirea ecuației intrarea 19, Tabelul 10. Datele sunt prezentate în Tabelul 12 și corelația monovariată:

$LD50 = 1.036 \times LD50_{calc.} + 293.19$; $n=10$; $R^2=0.637$ $s=183.325$; $F=14.057$ este reprezentată în Figura 12.

Tabelul 12. Valorile calculate a LD50 pentru moleculele din setul test (Tabelul 1)

Mol.	LD50	LD50 _{calc.}
1	10	41.948
2	300	540.105
3	56	758.130
8	300	634.108
9	300	374.332
11	200	514.784
19	1000	1289.931
27	18	195.801
29	300	970.905
30	115	305.196

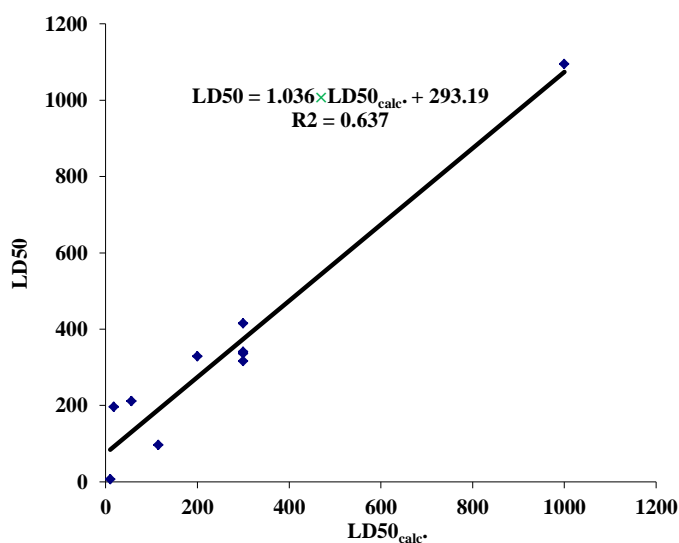


Figura 12. Reprezentarea grafică LD50 vs. LD50_{calc.} pentru setul test (validarea externă)

(c) Validarea pe clusteri de similaritate

Validarea poate fi efectuată prin calcularea LD50 pentru moleculele din setul test prin folosirea clusterilor de similaritate, ca în Secțiunea (b); valorile LD50_{calc.} au fost calculate cu aceiași descriptorii ca în ecuația 11, Tabelul 10. Datele sunt prezentate în Tabelul 13 și corelația monovariată: $LD50 = 1.013 \times LD50_{calc.} + 63.337$; $n=10$; $R^2=0.922$; $s=85.228$; $F=94.053$ este reprezentată în Figura 13.

Tabelul 13. Valorile calculate a LD50 prin clusteri de similaritate, pentru moleculele din setul test cu ecuația 11

Molecules	LD50	LD50 _{calc.}
1	10	22.10
2	300	332.54
3	56	299.08
8	300	367.49
9	300	349.53
11	200	247.45
19	1000	1058.68
27	18	-26.88
29	300	484.81
30	115	131.96

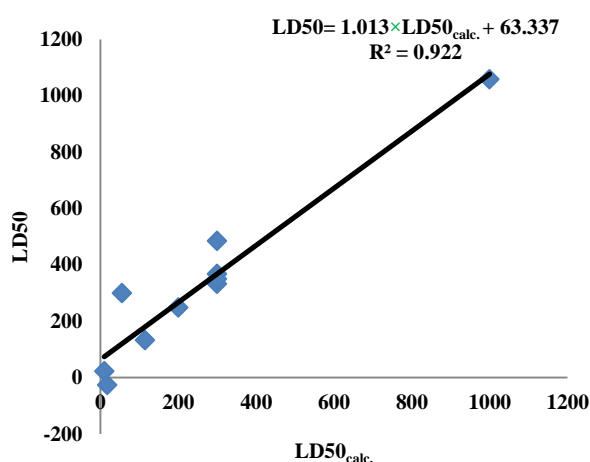


Figura 13. Graficul LD50 vs. LD50_{calc.} pe cluster de similaritate

De asemenea, validarea poate fi efectuată prin calcularea LD50 pentru moleculele din setul test cu aceiași descriptorii ca în ecuația 19, Table 10. Datele sunt prezentate în Tabelul 14 și corelația monovariată: $LD50 = 0.9998 \times LD50_{calc.} + 74.227$; $n=10$; $R^2=0.945$; $s=71.347$; $F=137.629$ este reprezentată în Figura 14.

Tabelul 14. Valorile calculate LD50 prin clusteri de similaritate pentru moleculele din setul test ca în ecuația 19

Molecules	LD50	LD50 _{calc.}
1	10	6.178
2	300	340.495
3	56	211.207
8	300	315.652
9	300	335.965
11	200	329.016
19	1000	1095.035
27	18	195.801
29	300	414.866
30	115	96.605

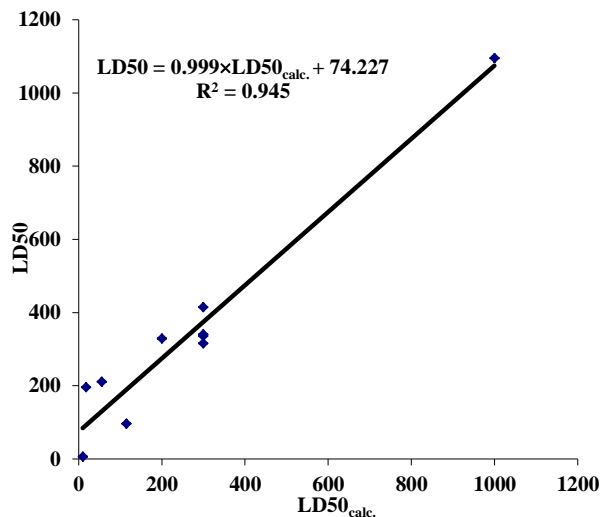


Figura 14. Graficul LD50 vs. LD50_{calc.} pe cluster de similaritate

Ca în cazul de mai sus, predicția LD50 este mai bună când folosim clusteri de similaritate decât prin clasică validare externă a modelului. Interesul nostru nu este să asigurăm un model general ci o mai bună predicție pentru cazul studiat.

1.2. Andocarea moleculară a derivaților de flavonoide

Un set de 30 derivați de flavonoide au fost descărcate din baza de date PubChem (Smiles cod, Tabelul 15).

Tabelul 15. Structura moleculară a derivaților de flavonoide, în cod SMILES (luate din PubChem).

Mol.	Canonical SMILES :
1	<chem>COC1=C(C=C(C=C1)C2=C(C(=O)C3=C(C(=C(C=C3O2)OC)OC)O)O)O</chem>
2	<chem>COC1=CC2=C(C=C1)C(=O)C=C(O2)C3=CC=CC=C3</chem>
3	<chem>C1=CC=C(C=C1)C2=C(C(=O)C3=CC=CC=C3O2)O</chem>
4	<chem>COC1=CC(=C2C(=C1)OC(=C(C2=O)O)C3=CC(=C(C=C3)O)O)OC</chem>
5	<chem>COC1=C(C(=C2C(=C1)OC(=CC2=O)C3=CC(=C(C=C3)O)O)O)O</chem>
6	<chem>COC1=CC=C(C=C1)C2=CC(=O)C3=C(O2)C(=C(C(=C3OC)OC)OC)OC</chem>
7	<chem>COC1=CC=C(C=C1)C2=CC(=O)C3=C(C=C(C=C3O2)OC)OC</chem>
8	<chem>COC1=CC=C(C=C1)C2=C(C(=O)C3=CC=CC=C3O2)O</chem>
9	<chem>COC1=CC=C(C=C1)C2=CC(=O)C3=C(C(=C(C=C3O2)OC)OC)OC</chem>
10	<chem>COC1=CC=CC(=C1C2=CC(=O)C3=C(O2)C(=C(C(=C3O)OC)OC)OC)O</chem>
11	<chem>CC1=C(C2=C(C(=C1O)C)OC(CC2=O)C3=CC(=C(C=C3)O)O)O</chem>
12	<chem>COC1=C(C=C(C=C1)C2=CC(=O)C3=C(C=C(C=C3O2)O)O)O</chem>
13	<chem>COC1=CC=C(C=C1)C2=C(C(=O)C3=C(C(=C(C=C3O2)OC)O)O)OC</chem>

14	<chem>COC1=CC=CC=C1C2=CC(=O)C3=C(O2)C(=C(C(=C3O)OC)OC)OC</chem>
15	<chem>COC1=C(C=C(C=C1)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)OC</chem>
16	<chem>C1=CC=CC(=C1)C3=C(C(C2=CC=CC=C2C3)=O)O</chem>
17	<chem>COC1=CC(=C2C(=C1)OC(=C(C2=O)O)C3=CC=CC=C3)O</chem>
18	<chem>C1=CC(=C(C=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O)O</chem>
19	<chem>C1=CC=C(C=C1)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O</chem>
20	<chem>C1=CC(=CC=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O</chem>
21	<chem>COC1=C(C=CC(=C1)C2=C(C(=O)C3=C(O2)C(=C(C=C3O)OC)OC)OC)O</chem>
22	<chem>COC1=C(C=C2C(=C1)C(=O)C(=C(O2)C3=CC=C(C=C3)O)O)OC</chem>
23	<chem>C1=CC(=C(C=C1O)O)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O</chem>
24	<chem>C1=CC=C(C=C1)C2=CC(=O)C3=CC(=C(C=C3O2)O)O</chem>
25	<chem>C1C(OC2=CC=CC=C2C1=O)C3=CC=CC=C3</chem>
26	<chem>C1=CC=C(C=C1)C2=CC(=O)C3=C(C=CC(=C3O2)O)O</chem>
27	<chem>C1C(OC2=CC=CC=C2C1=O)C3=CC=C(C=C3)O</chem>
28	<chem>COC1=CC=C(C=C1)C2CC(=O)C3=CC=CC=C3O2</chem>
29	<chem>COC1=C(C=CC(=C1)C2CC(=O)C3=CC=CC=C3O2)O</chem>
30	<chem>C1C(OC2=C(C1=O)C=CC(=C2)O)C3=CC=CC=C3</chem>

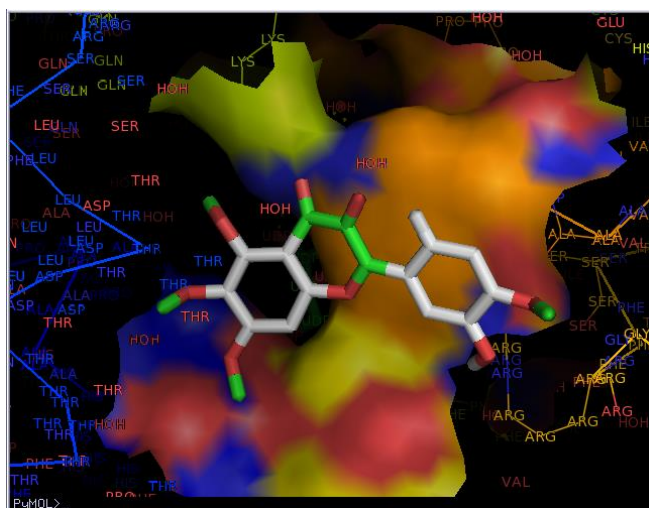


Figura 16: Interacțiunea dintre flavonoide cu
Flavonoid 3-O-glucosiltransferaza

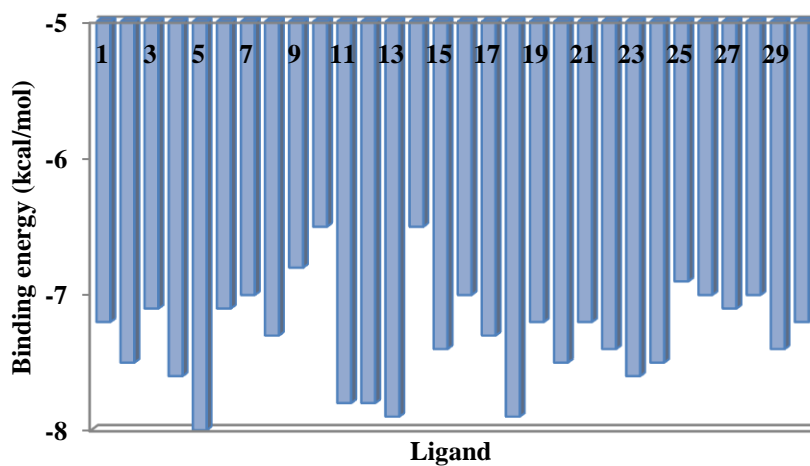


Figura 17: Energia de andocare obținută la apropierea de situsul activ al proteinei

Pentru a obține modelul farmacoforului pentru receptorul Flavonoid 3-O-glucosiltransferaza, cel mai bun conformer (pentru docking) a fost ales: liganzii 5, 27, 18 și 2 (cu energia de legare între -8 și -7.5). Pe baza acestor liganzi, am construit farmacoforul. Farmacoforul rezultat este prezentat în Figura 18.

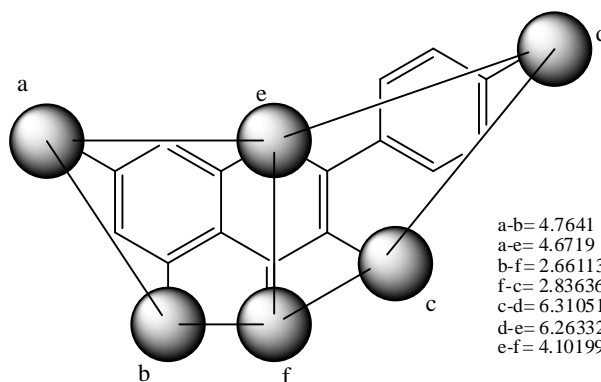
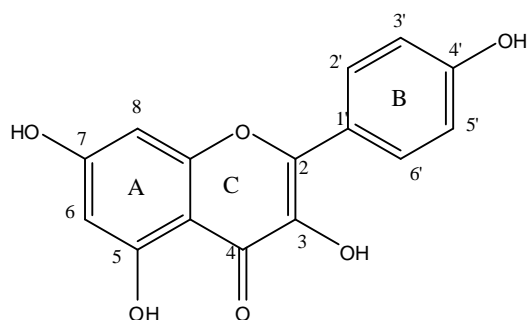


Figura 18(a): Modelul farmacofor pentru receptorul Flavonoid 3-O-glucosiltransferaza

Figura 18(b): Distanța dintre farmacofori Å).

1.2.3.2. Studiul QSAR pe baza energiei de andocare

(b) Validare externă

Valorile LD50 pentru setul test au fost calculate cu ecuația 15 din Tabelul 20. Datele sunt prezentate în Table 22 iar corelația monovariată: $LD50 = 0.787 \times LD50_{calc.} + 351.25$; $n = 10$; $R^2 = 0.787$; $s = 354.005$; $F = 19.292$ graficul este dat în Figura 20 [27].

Tabelul 22. Valoriile calculate LD50 pentru moleculele din setu test (vezi Tabelul 1)

Mol.	LD50	LD50 _{calc.}
2	300	408.58
4	1410	1040.06
5	2000	2254.1
11	800	1170.1
12	300	318.26
13	1000	1124.03
18	18	756.95
20	300	1023.86
23	555	528.23
24	300	383.83

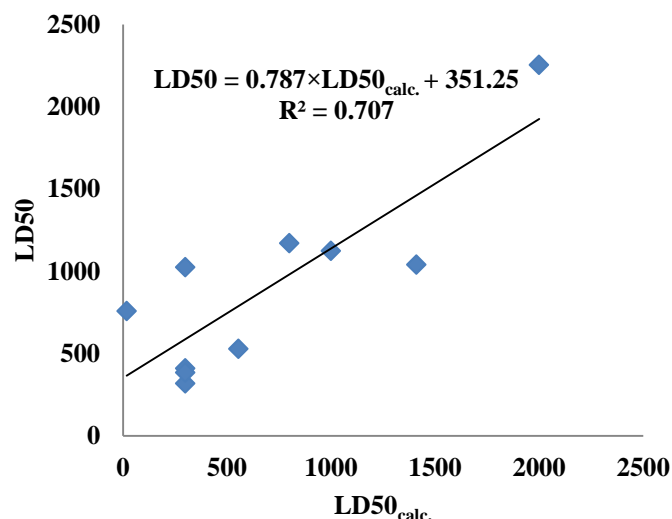


Figura 20. Graficul LD50 vs. LD50_{calc.} pentru setul test (validare externă)

Din tabelul 22 și figura 20, se poate vedea că modelele noastre nu sunt deosebit de bune în predicție, deoarece modelele sunt concepute pentru a da informații despre cei mai buni descriptorii.

(c) Validarea prin cluster de similaritate

Validarea se poate realiza de asemenea prin cluster de similaritate: fiecare dintre cele 10 molecule din setul test (ales cu energia minimă de andocare) este liderul propriului cluster, selectând similaritate 2D între 20 structuri a setului școală (fiecare cluster cuprinde aproximativ 15-17 molecule). Valorile LD50_{calc} sunt prezise cu 10 ecuații noi (liderul fiind lăsat afară) cu aceiași descriptorii ca în Ec. 15, Tabelul 18. Datele sunt prezentate în Tabelul 23 iar corelația monovariată este: $LD50 = 0.996 \times LD50_{calc.} + 58.008$; $n = 10$; $R^2 = 0.928$; $s = 175.605$; $F = 102.912$ graficul este dat în Figura 21 [26].

Tabelul 23. Valorile calculate LD50 pe clusters de similaritate, pentru moleculele din setul test

Mol.	LD50	LD50 _{calc}
2	300	333.02
4	1410	1265.67
5	2000	2281.45
11	800	595.18
12	300	342.23
13	1000	1130.75
18	18	363.89
20	300	408.81
23	555	542.53
24	300	272.08

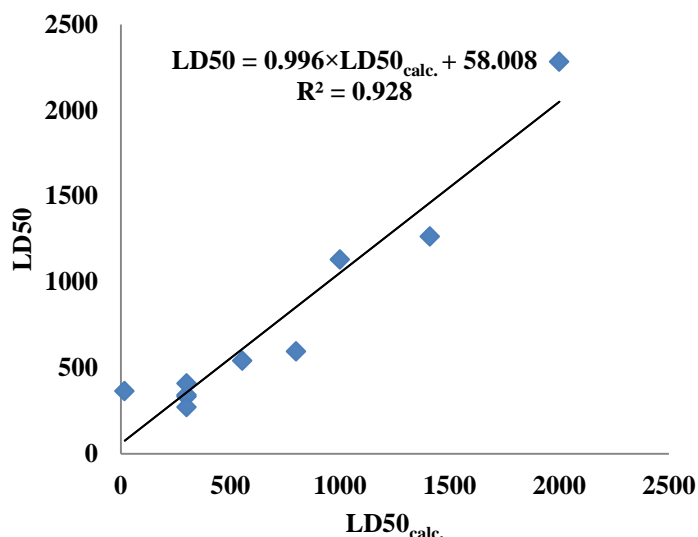


Figura 21. Graficul LD50 vs. LD50_{calc.} pe clusters de similaritate

Predicția LD50 este mult îmbunătățită folosind clusteri de similaritate ($R^2 = 0.928$) decât prin validarea externă caldică ($R^2 = 0.707$).

Energia de legătură mai mică este corelată cu LD50_{calc.} pentru setul test ($R^2 = 0.317$) această corelație fiind nesemnificativă din punct de vedere statistic, energia de andocare mai mică este folosită doar pentru a alege setul test.

2. Studiul QSAR a derivaților de testosteron

Un set de 40 molecule de testosteron au fost descărcate din baza de date PubChem (tabelul 22) și s-au împărțit în set de școală (25 molecule) și set de test (15 molecule), luate la întâmplare. Proprietatea aleasă pentru modelare este: log P (calculate) coeficientul de partiție între n-octanol și apă (a se vedea tabelul 24) și LD50 (pe șoareci, administrată oral)[119].

Tabelul 24. Structurile moleculare a testosteronului (in SMILES code) cu log P și LD50 (luate din PubChem)

Nr. Crt	Canonical SMILES	Log P	LD50
1	<chem>CC12CCC3C(C1CCC2O)CCC4=CC(=O)CCC34C</chem>	3.3	5000
2	<chem>CCC(=O)OC1CCC2C1(CCC3C2CCC4=CC(=O)CCC34C)C</chem>	4.4	1000
3	<chem>CC12CCC(=O)C=C1CCC3C2CCC4(C3CCC4(C)O)C</chem>	3.4	2500
4	<chem>CC12CCC(CC1CCC3C2CCC4(C3CCC4=O)C)O</chem>	3.7	980
5	<chem>CCCCCCC(=O)OC1CCC2C1(CCC3C2CCC4=CC(=O)CCC34C)C</chem>	6.3	1000
6	<chem>CC12CCC3C(C1CCC2OC(=O)CCC4=CC=CC=C4)CCC5=CC(=O)CCC35</chem>	5.1	595
7	<chem>CCCCC(=O)OC1CCC2C1(CCC3C2CCC4=CC(=O)CCC34C)C</chem>	5.3	980
8	<chem>CC(=O)OC1CCC2C1(CCC3C2CCC4=CC(=O)CCC34C)C</chem>	3.9	980

9	CC12CCC3C(C1CCC2(C#C)O)CCC4=CC(=O)CCC34	3	2000
10	CC12CCC3C(C1CCC2=O)CC(=C)C4=CC(=O)C=CC34C	3.1	980
11	CC(=O)OC1(CCC2C1(CCC3C2CCC4=CC(=O)CCC34)C)C#C	3.5	980
12	CC1CC2C(CCC3(C2CCC3(C(=O)C)OC(=O)C)C)C4(C1=CC(=O)CC4)C	4.1	6400
13	CC(=O)OC1CCC2C1(CCC3C2CCC4=C(C(=O)CCC34)C)C	4.7	980
14	CCC12CCC3C(C1CCC2(C#C)O)CCC4=CC(=O)CCC34	3.3	5010
15	CC12CCC3C(C1CCC2OC(=O)C4=CC=CC=C4)CCC5=CC(=O)CCC35C	5.6	980
16	CC12CCC(=O)C=C1CCC3C2CCC4(C3CCC4(C#C)O)C	3.5	980
17	CCC12CC(=C)C3C(C1CCC2(C#C)O)CCC4=CC(=O)CCC34	3.3	980
18	CC12CCC3C(C1CCC2O)CCC4=CC(=O)C=CC34C	3.5	980
19	CC1CC2C3CCC(C3(CC(C2C4(C1=CC(=O)CC4)C)O)C)C(=O)C	2.7	980
20	CC12CC(C3C(C1CCC2(C(=O)CO)O)CCC4=CC(=O)C=CC34C)O	1.6	250
21	CC12CCC(=O)C=C1CCC3C2C(CC4(C3CCC4(C(=O)CO)O)C)O	1.6	5000
22	CCC(=O)C1(C(CC2C1(CC(C3C2CCC4=CC(=O)C=CC34C)O)C)C)C	3.5	980
23	CC1CC2C3CCC(C3(CC(C2C4(C1=CC(=O)C=C4)C)O)C)(C(=O)CO)O	1.9	4000
24	CC1CC2C3CCC(C3(CC(C2C4(C1=CC(=O)C=C4)C)O)C)(C(=O)COC(=O)C)O	2.7	10000
25	CC12CCC(=O)C=C1CCC3C2CCC4(C3CCC4(C#C)O)C	3.5	980
26	CC12CCC3C(C1CCC2(C)O)CCC4C3(COC(=O)C4)C	3.7	10000
27	CC12C=CC3=C4CCC(=O)C=C4CCC3C1CCC2(CC=C)O	2.8	980
28	CC12CCC3C(C1CCC2OC(=O)C4=CC=CC=C4)CCC5=CC(=O)CCC35C	5.6	980
29	CC12CCC(=O)C=C1CCC3C2C(CC4(C3CCC4(C(=O)O)O)C)O	1.6	980
30	CC12CCC3C(C1CCC2(C)O)CCC4=CC(=O)C=CC34C	3.6	1000
31	CC1CC(=O)CC2C1(C3CCC4(C(C3CC2)CCC4O)C)C	4.1	980
32	CC1=CC2C(CCC3(C2CCC3(C(=O)C)OC(=O)C)C)C4(C1=CC(=O)CC4)C	3.1	980
33	CC12CCC3C(C1CCC2C(=O)COC(=O)C(C)C)CCC4=CC(=O)CCC34C	4.5	980
34	CCCCC(=O)OC1(CCC2C1(CCC3C2CCC4=CC(=O)CCC34)C)C(=O)C	5.7	980
35	CC1CC2(C(CCC3C2CCC4(C3CCC4O)C)CC1=O)C	4.2	980
36	CC12CCC(=O)C=C1CCC3C2CCC4(C3CCC4(C#C)O)C	3.5	2000
37	CC12CCC3C(C1CCC2(C#C)O)CCC4=C3CCC(=O)C4	2.1	980
38	CCC(=O)OC1CCC2C1(CCC3C2CCC4C3(CC(C(=O)C4)C)C)C	5.3	980
39	CC1=CC(=O)CC2C1(C3CCC4(C(C3CC2)CCC4OC(=O)C)C)C	4.4	4000
40	CC12CCC(=O)C=C1CCC3C2CCC4(C3CCC4(C)O)C)O	3.2	980

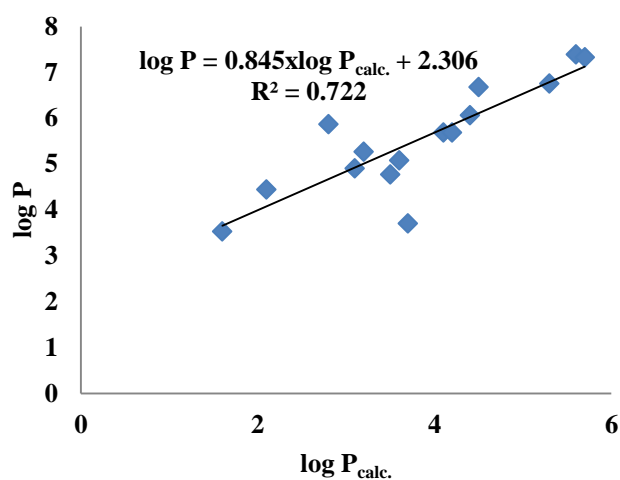
Fragmente de masă (cazul 1)

(b) Validarea externă

Valorile log P pentru setul test a derivaților de testosteron (ultimele 15 structurile din tabelul 24), au fost calculate folosind ecuația cea mai bună din tabelul 27, ecuația 9. Datele sunt prezentate în Tabelul 29 și corelația monovariate: $\log P = 0.845 \times \log P_{calc.} + 2.306$; $n=15$; $R^2=0.722$; $s=0.652$; $F=33.8$ în Figura 23.

Tabelul 29. Valoarea calculată log P pentru molecule din setul test (Tabelul1)

Mol.	logP	log P _{calc.}
26	3.7	3.70
27	2.8	5.86
28	5.6	7.39
29	1.6	3.53
30	3.6	5.08
31	4.1	5.68
32	3.1	4.90
33	4.5	6.68
34	5.7	7.33
35	4.2	5.68
36	3.5	4.77
37	2.1	4.44
38	5.3	6.75
39	4.4	6.06
40	3.2	5.27



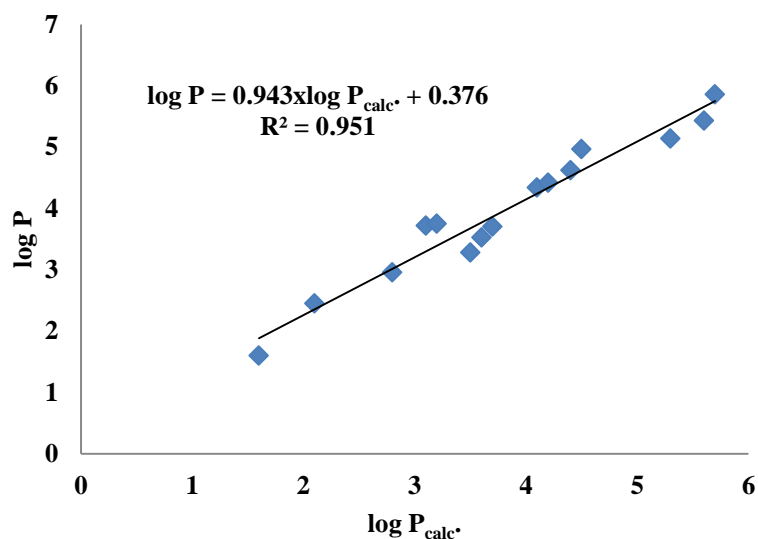
Figurea 23. Grafic log P vs. log P_{calc.} pentru setul test (validare externă)

(c) Validarea prin Clustere de Similaritate

Validarea poate fi efectuată prin calcularea log P pentru moleculele din setul test stabilit cu ecuația de învățare pe grupuri de similaritate: fiecare din cele 15 molecule este lider în propriul său grup, selectate prin (2D) similaritate între cele 25 de structurile din setul școală și 15 structuri din setul test. Valorile log P_{calc.} pentru fiecare din cele 15 de moleculele din setul test au fost calculate cu 15 ecuații noi (liderul nu rămâne în afara) cu aceleași descriptorii din ecuația. 9, Tabelul 27. Datele sunt prezentate în Tabelul 30 și corelarea monovariate: $\log P = 0.943 \times \log P_{calc.} + 0.376$; $n=15$; $R^2=0.951$; $s=0.273$; $F=254.62$ iar graficul este dat în Figure 24.

Tabelul 30. Valoarea calculată a log P pentru moleculele din setul test (Tabelul 1)

Mol.	logP	logPcalc.
26	3.7	3.70
27	2.8	2.96
28	5.6	5.43
29	1.6	1.60
30	3.6	3.52
31	4.1	4.34
32	3.1	3.72
33	4.5	4.97
34	5.7	5.86
35	4.2	4.42
36	3.5	3.28
37	2.1	2.45
38	5.3	5.14
39	4.4	4.62
40	3.2	3.75



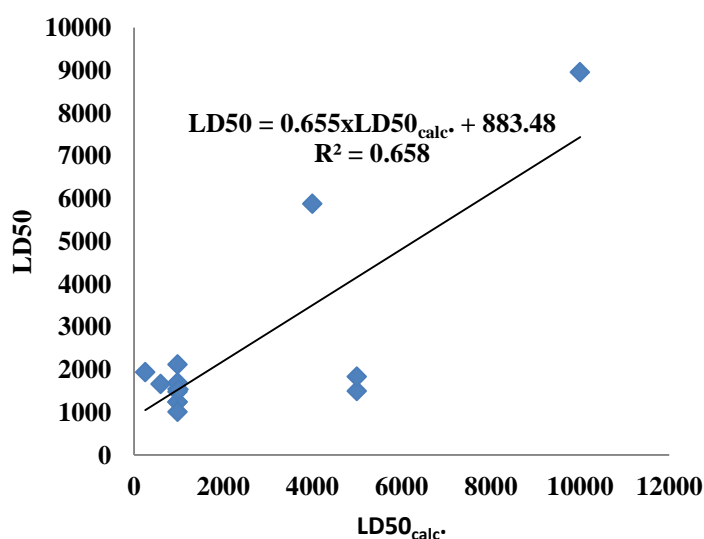
Figurea 24. Grafic log P vs. log P calc. pentru setul test (pe custer de similaritate)

(b) Validare externă

Valoarea LD50 pentru setul test a derivaților de testosteron (Tabelul 22, ultimele 15 structuri) sunt calculate folosind ce-a mai bună ecuație din Tabelul 31, intrarea 14. Datele sunt prezentate în Tabelul 33 și corelția monovariantă: $LD50 = 883.48 + 0.655 \times LD50_{calc.}$; $n=15$; $R^2=0.658$; $s=1618.663$; $F= 24.96$ graficul este dat în Figure 25.

Tabelul 33. Valoarea calculate a LD50 pentru moleculele din setul test (Tabelul 1)

Mol.	LD50	LD50calc.
1	5000	1497.73
6	595	1660.89
8	980	1240.10
15	980	1240.63
16	980	1694.42
17	980	1009.40
18	980	1544.38
19	980	2120.44
20	250	1940.07
21	5000	1829.83
22	980	1483.04
23	4000	5876.90
24	10000	8954.71
25	980	1694.42
30	1000	1536.21



Figurea 25. Grafic LD50 vs. LD50calc. pentru setul test (validare externă)

(c) Validarea clustere de similaritate

Validarea se realizează prin calcularea LD50 pentru moleculele din setul test. Valoarea LD50calc. este calculată cu aceiași descriptori ca în ecuația 14 din Tabelul 31. Datele sunt prezentate în Tabelul 34 și corelația monovariată este: $LD50 = 0.877 \times LD50_{calc.} + 142.53$; $n=15$; $R^2=0.935$; $s=706.157$; $F=186.44$ is plotted in Figura 26.

Tabelul 34. Valoarea calculată LD50 prin clusteri de similaritate pentru moleculele din setul test

Mol.	LD50	LD50calc.
1	5000	2507.05
6	595	881.54
8	980	1194.66
15	980	1082.21
16	980	954.23
17	980	985.43
18	980	922.69
19	980	922.69
20	250	612.35
21	5000	4839.34
22	980	944.58
23	4000	4498.85
24	10000	9413.14
25	980	954.23
30	1000	952.63

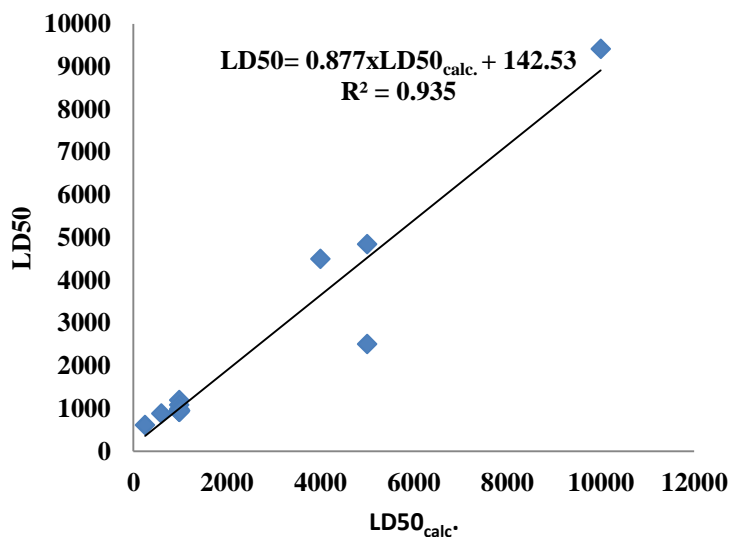


Figura 26. Graficul LD50 vs. LD50calc. prin clustes de similaritate

Sarcină parțială (cazul 2)

Validarea modelului (pentru log P)

(a) Validarea Externă

Valoarea log P pentru setul test de testosteron (Tabelul 24), este calculat folosind ecuația 10 Tabelul 35. Datele sunt prezentate în Tabelul 37 corelația monovariantă este: $\log P = 0.809 \times \log P_{calc.} + 0.722$; $n=15$; $R^2=0.938$; $s=0.347$; $F=195.43$ graficul în Figura 27.

Tabelul 37. Valoarea calculată log P pentru moleculele din setul test (Tabelul 24)

Mol.	log P	log Pcalc.
1	3.3	3.83
2	4.4	4.55
3	3.4	3.58
4	3.7	3.40
5	6.3	5.94
18	3.5	3.44
19	2.7	2.51
20	1.6	1.78
21	1.6	2.26
22	3.5	3.20
23	1.9	2.54
37	2.1	2.71
38	5.3	5.15
39	4.4	4.06
40	3.2	3.07

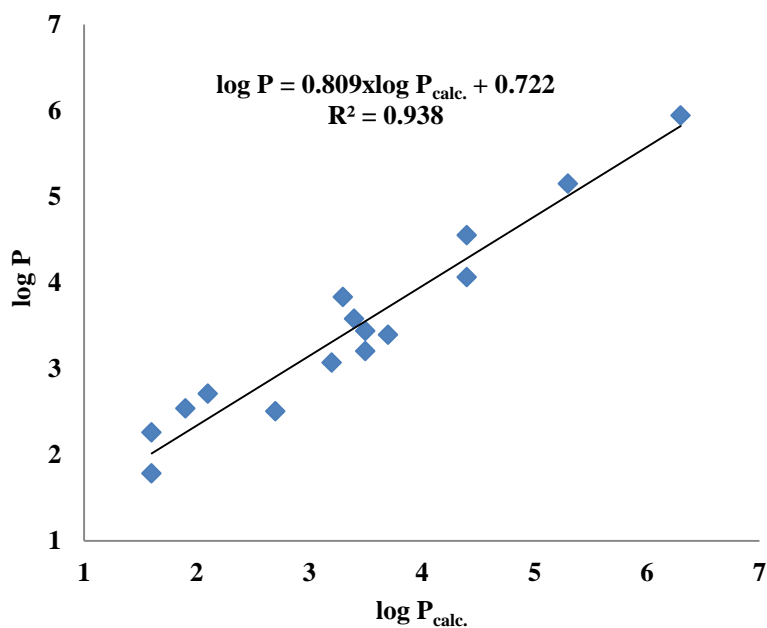


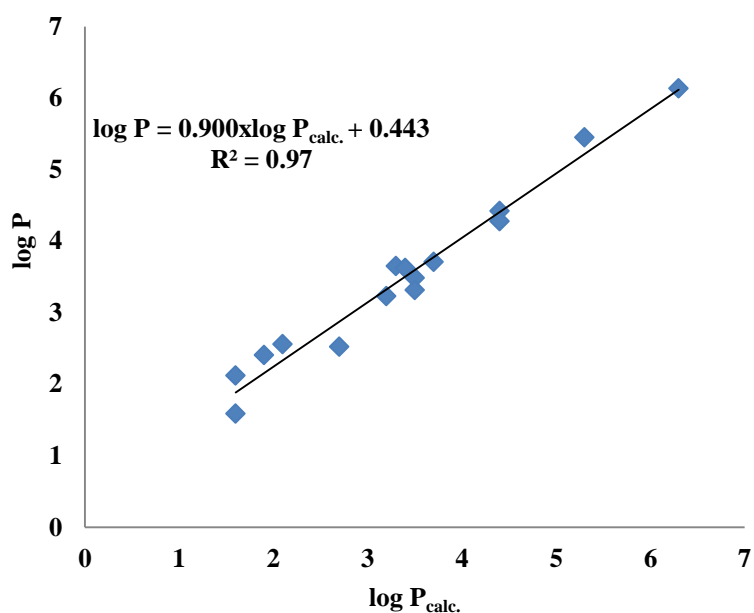
Figura 27. Graficul log P vs. log P calc. pentru setul test (validare externă)

(c) Validarea prin clustere de similaritate

Validarea se efectuează prin calcularea log P pentru moleculele din setul test. Valoarea log Pcalc. este calculată cu aceiași descriptori ca în ecuația 10, Tabelul 35. Datele sunt prezentate în Tabelele 38 și corelația monovariată este: $\log P = 0.900 \times \log P_{calc.} + 0.443$; $n=15$; $R^2=0.97$; $s=0.240$; $F=420.87$ graficul este dat în Figura 28.

Tabelul 38. Valorile calculate a log P, pe clusteri de similaritate, pentru moleculele din setul test

Mol.	log P	log P _{calc.}
1	3.3	3.65
2	4.4	4.42
3	3.4	3.62
4	3.7	3.71
5	6.3	6.13
18	3.5	3.48
19	2.7	2.52
20	1.6	1.59
21	1.6	2.12
22	3.5	3.31
23	1.9	2.41
37	2.1	2.56
38	5.3	5.45
39	4.4	4.27
40	3.2	3.23



Figurea 28. Graficul Log P vs. log P_{calc.} prin clusters de similaritate

Validarea modelului (pentru LD50)

(b) Validare externă

Valoarea LD50 pentru setul test de derivați de testosteron (Tabelul 24), este calculat folosind cea mai bună ecuația 10 din Tabelul 39. Datele sunt prezentate în Tabelul 41 și corelație monovariantă este: $LD50 = 0.830 \times LD50_{\text{calc.}} + 839.36$; $n=15$; $R^2=0.840$; $s=1039.906$; $F=68.31$ graficul este dat în Figura 29.

Tabelul 41. Valoarea calculată LD50 pentru moleculele din setul

test		
Mol.	LD50	LD50calc.
1	5000	2967.99
2	1000	1745.51
3	2500	3699.88
4	980	530.19
5	1000	1834.74
6	595	1619.14
7	980	1563.68
8	980	1751.13
9	2000	2467.30
22	980	3051.90
23	4000	2819.61
24	10000	10376.48
25	980	1666.63
37	980	1973.52
38	980	1886.35

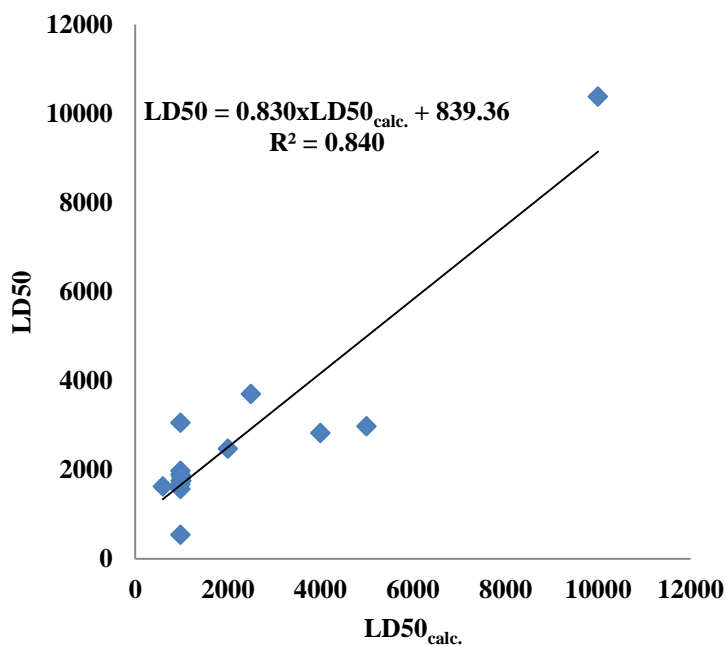


Figura 29. Graficul LD50 vs. LD50calc. (validare externă)

(c) Validarea prin clustere de similaritate

Validarea s-a realizat prin calcularea LD50 pentru moleculele din setul test. Valoarea LD50calc. este calculată cu aceiași descriptori din ecuația 10, Tabelul 39. Datele sunt prezentate în Tabelul 42 și corelația monovariată este: $LD50 = 464.29 + 0.832 \times LD50_{calc.}$; $n=15$; $R^2=0.943$; $s=622.801$; $F= 213.68$ reprezentat grafic în Figura 30.

Tabelul 42. Valoarea calculată LD50 prin clusteri de similaritate pentru moleculele din setul test (Tabelul 1)

Mol.	LD50	LD50 _{calc.}
1	5000	3632.01
2	1000	1382.42
3	2500	3207.06
4	980	1225.48
5	1000	860.77
6	595	350.34
7	980	1372.07
8	980	1282.10
9	2000	2463.65
22	980	1554.78
23	4000	2913.17
24	10000	9339.85
25	980	1722.29
37	980	1303.88
38	980	1758.71

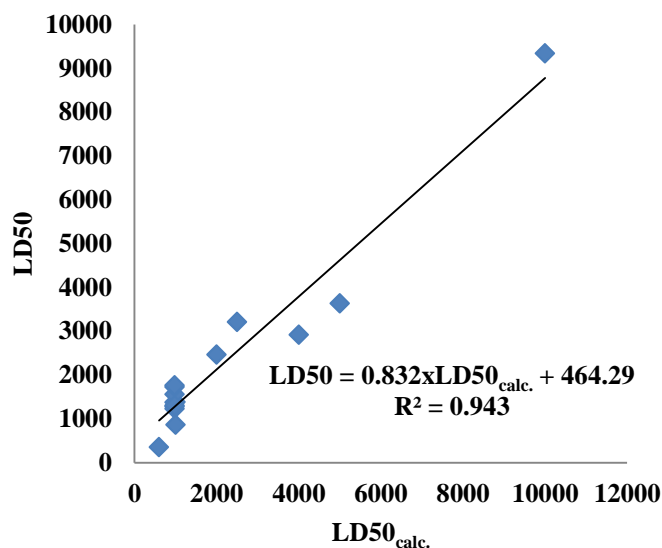


Figura 30. Graficul LD50 vs. LD50 calc. pentru setul test prin cluster de similaritate

3. Studiul QSAR și de andocare moleculară a derivaților de antrachinonă

Un set de 40 antrachinone a fost descărcat din baza de date PubChem [122] (Table 43) și a fost împărțită în set școală (25 molecule) și set test (15 molecule), luate randomly. Proprietatea aleasă pentru modelare sunt log P (calculat) coeficientul de partiție dintre n-octanol și apă, LD50 (șobolani, ruta de administrare orală) (vezi Tab. 44).

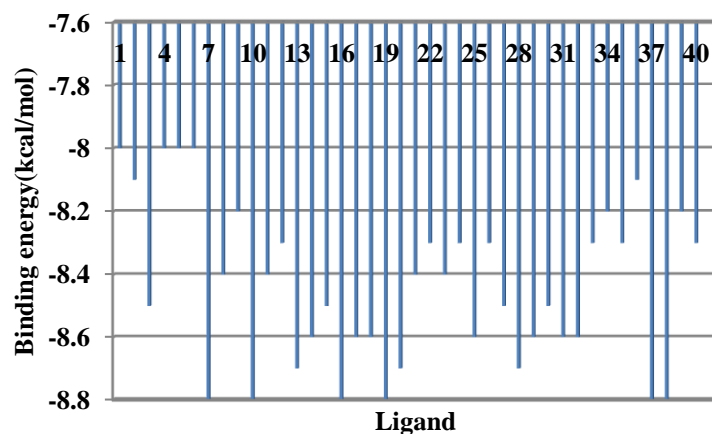
Tab. 43. Antrachinone structura moleculară și log P (luate din PubChem)

Nr. Crt.	Canonical SMILES	CID	log P
1	<chem>C1=CC=C2C(=C1)C(=O)C3=CC=CC=C3C2=O</chem>	6780	3.4
2	<chem>C1=CC=C2C(=C1)C(=O)C3=C(C2=O)C=C(C=C3)O</chem>	11796	3
3	<chem>CC1=CC(=C2C(=C1)C(=O)C3=C(C2=O)C(=CC=C3)O)O</chem>	10208	3.5
4	<chem>C1C2=C(C(=CC=C2)O)C(=O)C3=C1C=CC=C3O</chem>	2202	3.2
5	<chem>C1=CC2=CC3=C(C(=CC=C3)O)C(=C2C(=C1)O)O</chem>	10187	3.9
6	<chem>C1=CC2=C(C(=C1)O)C(=O)C3=C(C2=O)C=CC=C3O</chem>	2950	3.2
7	<chem>CC1=C(C2=C(C=C1)C(=O)C3=C(C2=O)C=CC(=C3)O)O</chem>	442756	3.3
8	<chem>CC1=C(C=C2C(=C1)C(=O)C3=CC=CC=C3C2=O)O</chem>	10889963	2.9
9	<chem>C1=C(C=C2C(=C1O)C(=O)C3=C(C=C(C=C3C2=O)O)O)O</chem>	3016789	2
10	<chem>CC1=CC(=C2C(=C1)C(=O)C3=CC(=C(C(=C3C2=O)O)O)O)O</chem>	12548	2.4
11	<chem>CC1=CC(=C2C(=C1)CC3=CC(=CC(=C3C2=O)O)O)O</chem>	122635	3.2
12	<chem>CC1=CC2=C(C=C1)C(=O)C3=C(C2=O)C(=CC=C3)O</chem>	155237	3.9
13	<chem>CC1=C(C2=C(C=C1)C(=O)C3=C(C2=O)C=CC(=C3)O)O</chem>	124063	3.1
14	<chem>CC1=C(C2=C(C=C1)C(=O)C3=CC(=C(C=C3C2=O)O)O)O</chem>	25202820	2.7
15	<chem>CC1=C(C(=C2C(=C1)C(=O)C3=C(C2=O)C=CC=C3)O)O</chem>	12322346	3.3
16	<chem>CC1=CC2=C(C=C1)C(=O)C3=C(C2=O)C=CC(=C3)O</chem>	5319503	3.1
17	<chem>CC1=CC=CC2=C1C(=O)C3=C(C2=O)C=C(C=C3)O</chem>	57536669	3.1

18	<chem>CC1=C(C(=C2C(=C1)C(=O)C3=CC=CC=C3C2=O)O)O</chem>	429241	3.1
19	<chem>CC1=CC(=C2C(=C1)C(=O)C3=C(C2=O)C(=C(C=C3)O)O)O</chem>	12313148	2.7
20	<chem>CC1=C(C(=C2C(=C1)C(=O)C3=C(C2=O)C(=CC=C3)O)O)O</chem>	442759	2.7
21	<chem>C1=CC2=C(C(=C1)O)C(=O)C3=CC(=C(C=C3C2=O)O)O</chem>	11196140	2.8
22	<chem>C1=CC2=C(C(=C1)O)C(=O)C3=C(C2=O)C(=C(C=C3)O)O</chem>	436367	3.4
23	<chem>C1=CC2=C(C=C1O)C(=O)C3=C(C2=O)C(=C(C=C3)O)O</chem>	65739	2.4
24	<chem>C1=CC=C2C(=C1)C(=O)C3=C(C2=O)C(=C(C=C3)O)O</chem>	6293	3.2
25	<chem>C1=CC2=C(C=C1O)C(=O)C3=C(C2=O)C=CC(=C3O)O</chem>	1320	2.4
26	<chem>CC1=C2C(=CC(=C1O)O)C(=O)C3=CC=CC=C3C2=O</chem>	11391150	2.5
27	<chem>C1=CC(=C(C2=C1C(=O)C3=C(C2=O)C=CC(=C3O)O)O)O</chem>	69440	2.5
28	<chem>CC1=C(C=C2C(=C1)C(=O)C3=C(C2=O)C(=CC=C3)O)O</chem>	71368906	3.1
29	<chem>C1=CC(=C(C2=C1C(=O)C3=C(C2=O)C(=C(C=C3)O)O)O)O</chem>	22643725	2
30	<chem>CC1=C(C=C2C(=C1C)C(=O)C3=C(C2=O)C=CC=C3O)O</chem>	57745748	3.4
31	<chem>CC1=C(C2=C(C=C1)C(=O)C3=CC(=C(C=C3C2=O)O)O)O</chem>	25203424	2.4
32	<chem>CC1=C(C(=C2C(=C1)C(=O)C3=CC(=CC(=C3C2=O)O)O)O)O</chem>	11818503	2.4
33	<chem>CC1=C2C(=CC(=C1)O)C(=O)C3=C(C2=O)C(=CC=C3)O</chem>	3085033	3.1
34	<chem>C1=CC2=C(C(=C1)O)C(=O)C3=C(C2=O)C=CC(=C3)O</chem>	14886011	3.2
35	<chem>C1=CC2=C(C(=C1)O)C(=O)C3=C(C2=O)C=C(C=C3)O</chem>	12628831	3.2
36	<chem>C1=CC=C2C(=C1)C(=O)C3=CC(=C(C=C3C2=O)O)O</chem>	11031986	2.7
37	<chem>CC1=C(C=C2C(=C1)C(=O)C3=C(C2=O)C=C(C=C3)O)O</chem>	10060853	2.5
38	<chem>CC1=CC(=C(C2=C1C(=O)C3=C(C=C(C=C3C2=O)O)O)O)O</chem>	9817337	2.9
39	<chem>C1=CC(=C(C2=C1C(=O)C3=C(C=CC(=C3C2=O)O)O)O)O</chem>	5004	2.5
40	<chem>C1=C2C(=CC(=C1O)O)C(=O)C3=CC(=C(C=C3C2=O)O)O</chem>	44300874	1.4

În Figura 33 sunt prezentate conformația liganzilor cea mai favorabilă cu cea mai mică energie de legare .

Figura 33: Reprezentarea grafică arată energia de legare (kcal / mol) al liganzilor.



Obținerea farmacoforului pe baza toxicității model pentru 3 Beta glicogen kinaza a fost ales conformerul cu cea mai favorabilă interacțiune cu receptorul rezultat la andocare. Ligandul 7, 10, 16, 19, 37 și 38 au cea mai mică energie de legare -8.8 kcal/mol, bazat pe acești compuși s-a construit farmacoforul. Farmacoforul este arătate în Figura 34(a).

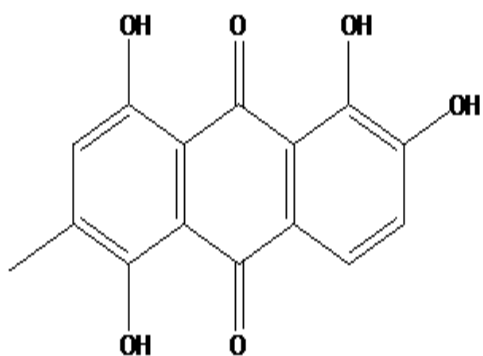


Figura 34(a): Modelul farmacofor

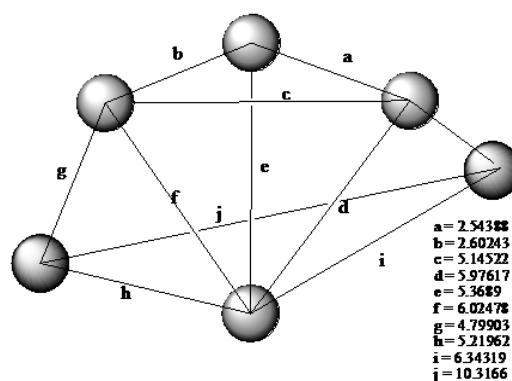


Figura 34(b): Distanțele în farmacofor

Farmacoforul în șase puncte, cum se arată în figura 34 (b), a fost construit în conformitate cu conformația cu energia de legare minimă a acestor molecule.

(b) Validarea externă

Valorile $\log P$ pentru setul test de antrachinone este calculat cu ecuația 11 din Tabelul 49. Datele sunt prezentate în Tabelul 51 și corelația monovariată este: $\log P = 0.934 \times \log P_{calc.} + 0.298$; $n=15$; $R^2=0.754$; $s=0.201$; $F=39.749$ în graficul din Figura 35.

Tabelul 51. Valorile calculate $\log P$ pentru moleculele din setul test (Tabelul 39)

Mol.	$\log P$	$\log P_{calc.}$
1	3.4	3.69
2	3	2.92
4	3.2	3.64
5	3.9	3.64
8	2.9	3.12
11	3.2	3.39
12	3.9	4.06
13	3.1	3.32
14	2.7	2.77
16	3.1	3.23
17	3.1	3.13
18	3.1	3.23
24	3.2	3.01
26	2.5	2.80
36	2.7	2.45

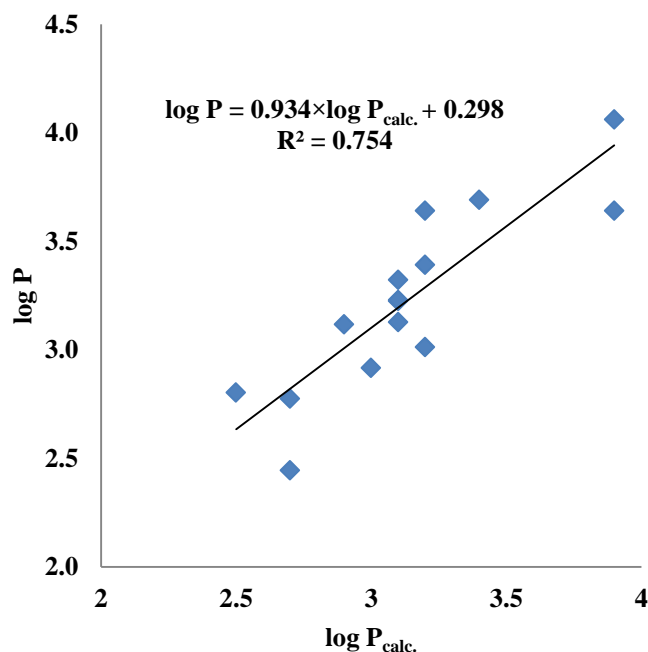


Figura 35. Graficul $\log P$ vs. $\log P_{calc.}$ pentru moleculele din setul test (validare externă)

(c) Validarea prin clustere de similaritate

Valorile log P pentru setul test a derivaților de antrachinone sunt calculate folosind ecuația 11 din Tabelul 49. Datele sunt prezentate în Tabelul 52 corelația monovariată: $\log P = 1.039 \times \log P_{calc.} - 0.042$; $n=15$; $R^2=0.961$; $s=0.080$; $F=317.747$ în graficul din Figura 36.

Chiar și numărul de descriptorii per probă este mai mare decât în cazul validării externe, libertatea în general este compensată printr-o mai bună predicție: comparând $R^2=0.754$ (validarea externă) și $R^2=0.961$ (validarea prin clustere de similaritate), cu același număr de descriptorii.

Tabelul 52. Valorile calculate of log P

Pe clustere de similaritate pentru
moleculele din setul test

Mol.	log P	log P _{calc.}
1	3.4	3.47
2	3	2.91
4	3.2	3.35
5	3.9	4.06
8	2.9	3.04
11	3.2	3.32
12	3.9	4.00
13	3.1	3.29
14	2.7	2.75
16	3.1	3.19
17	3.1	3.13
18	3.1	3.20
24	3.2	3.13
26	2.5	2.65
36	2.7	2.69

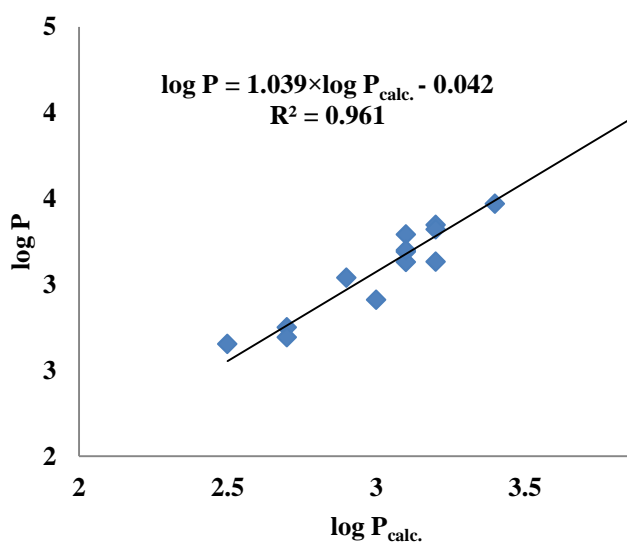


Figura 36. Graficul log P vs. log P_{calc.} pe
clustere de similaritate

Sarcină parțială (log P)

(b) Validarea externă

Valorile log P pentru setul test a fost calculat folosind ecuația 11 din Tabelul 53. Datele sunt prezentate în Tabelul 55, iar corelația monovariată este $\log P = 0.896 \times \log P_{calc.} + 0.307$; $n=15$; $R^2=0.719$; $s=0.214$; $F=33.338$ în graficul din Figura 37.

Tabelul 55. Valorile calculate log P pentru moleculele din setul test (Tabelul 43)

Mol.	log P	log P _{calc.}
1	3.4	3.78
2	3	3.05
4	3.2	3.26
5	3.9	3.48
8	2.9	3.09
11	3.2	2.99
12	3.9	3.93
13	3.1	3.26
14	2.7	2.47
16	3.1	3.14
17	3.1	2.98
18	3.1	3.15
24	3.2	2.98
26	2.5	2.75
36	2.7	2.43

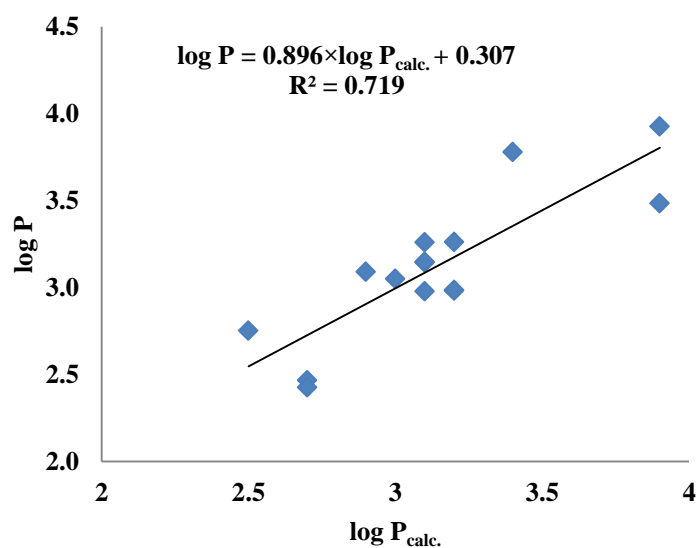


Figura 37. Graficul log P vs. log P_{calc.} pentru setul test (validarea externă)

(c) Validarea prin clustere de similaritate

Valorile log P pentru setul test sunt calculate cu ecuația 11 din Tabelul 53. Datele sunt prezentate în Tabelul 56, iar corelația monovariată este: $\log P = 0.964 \times \log P_{\text{calc.}} + 0.112$; $n=15$; $R^2=0.969$; $s=0.071$; $F=403.421$ este reprezentat în Figure 38. O mai bună predicție comparând $R^2 = 0.969$ sarcină parțială față de $R^2 = 0.961$ când folosim fragmente de masă.

Tabelul 56. Valorile calculate log P pe clustere de similaritate, pentru moleculele din setul test

Mol.	log P	log P _{calc.}
1	3.4	3.39
2	3	3.02
4	3.2	3.26
5	3.9	3.81
8	2.9	2.98
11	3.2	3.23
12	3.9	3.94
13	3.1	3.16
14	2.7	2.75
16	3.1	3.11
17	3.1	3.02
18	3.1	3.08
24	3.2	3.03
26	2.5	2.47
36	2.7	2.72

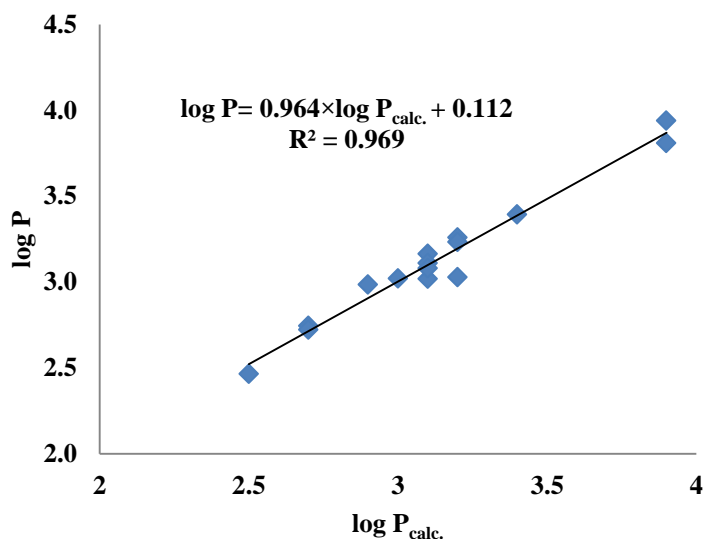


Figura 38. Graficul log P vs. log P_{calc.} pe clusters de similaritate

Sarcină parțială (LD50 bazat pe similaritate topologică)

(b) Validare externă

De asemenea, valorile LD50 calc. pentru fiecare din cele 12 molecule din setul test au fost alese pe baza celei mai mici energii de andocare și sunt calculate cu aceiași descriptori din ecuația de mai jos:

$$LD50 = 11832.36 + 1.017 \times SD_{LD50} + 18.889 \times CjDe - 52.112 \times CfDe$$

Datele sunt prezentate în Tabelul 59 și corelația monovariată este:

$$LD50 = 0.866 \times LD50_{calc.} + 545.6; n=12; R^2=0.904; s=477.245; F=95.201$$

graficul din Figura 39.

Tabelul 59. Valorile calculate LD50 pentru moleculele din setul test

Mol.	LD50	LD50 _{calc.}
3	2500	2586.67
7	1230	1850.72
13	4000	3709.78
14	2000	2484.99
16	2795	1980.80
18	5000	4816.41
19	35	688.14
20	308	366.56
25	1870	2448.36
28	2795	3423.44
32	3950	4365.44
38	2795	3185.28

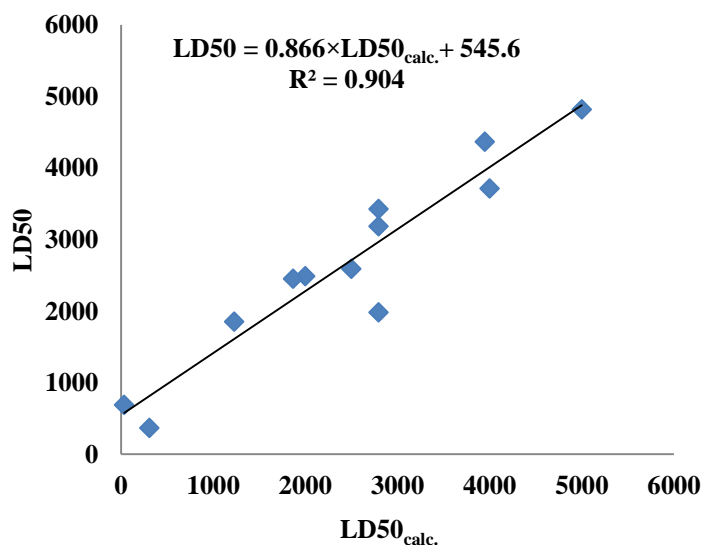


Figura 39. Grafic LD50 vs. LD50_{calc.} pentru setul test
(validarea externă)

(c) Validarea prin clustere de similaritate

Valorile LD50 pentru setul test sunt calculate folosind ecuația 10 din Tabelul 57. Datele sunt prezentate în Tabelul 60 și corelația monovariată este: $LD50 = 0.861 \times LD50_{calc.} + 506.19$; $n=7$; $R^2=0.959$; $s=314.696$; $F=231,948$ graficul în Figura 40.

Tabelul 60. Valorile calculate LD50, prin clustere de similaritate, pentru moleculele din setul test

Mol.	LD50	LD50 _{calc.}
3	2500	2528.503
7	1230	1622.526
13	4000	3463.02
14	2000	2377.399
16	2795	2796.373
18	5000	4810.405
19	35	594.5576
20	308	366.6446
25	1870	2432.604
28	2795	3328.103
32	3950	4081.971
38	2795	2872.895

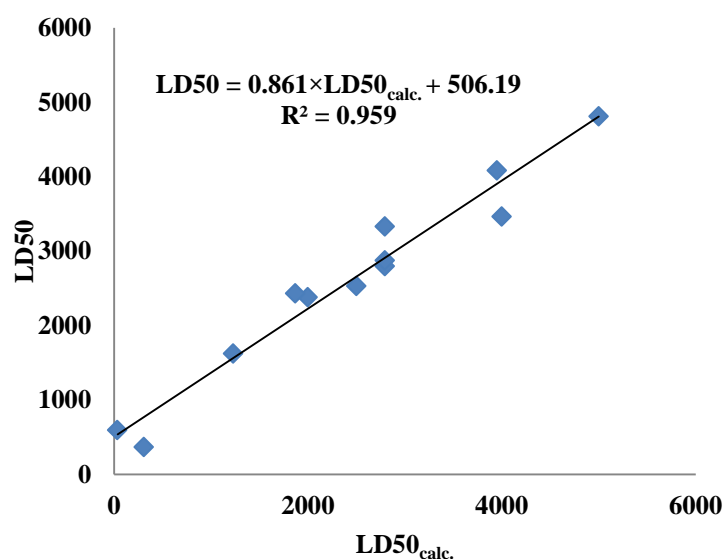


Figura 40. Grafic LD50 vs. LD50_{calc.} prin clustere de similaritate

4. Studiul QSAR și de andocare moleculară ale derivaților de resveratrol

4.1. Studiul QSAR ale derivaților de resveratrol

Un set de 40 derivați de resveratrol, descărcate din baza de date PubChem [129] (Tabelul 61), este împărțit în set de învățare (25 molecule) și set test (15 molecule), luate randomy; proprietatea modelată este log P.

Tabelul 61. Derivați de resveratrol structura moleculară (în cod SMILES) și log P (din PubChem).

Mol.	Canonical SMILES	log P	CID
1	<chem>C1=CC(=CC=C1CCC2=CC(=CC(=C2)O)O)O</chem>	3.1	185914

2	CC(=CC1=CC=C(C=C1)O)C2=CC(=CC(=C2)O)O	3.7	75071272
3	C1=CC(=CC(=C1)O)CCC2=CC(=CC(=C2)O)O	3.1	21574990
4	C1=CC=C(C=C1)CCC2=CC(=CC(=C2)O)O	3.4	442700
5	CC(CC1=CC(=CC(=C1)O)O)C2=CC=C(C=C2)O	3.4	58892268
6	COC1=C(C=CC(=C1)C=CC2=CC(=CC(=C2)O)O)O	3.2	5318650
7	COC1=C(C=C(C=C1)CC(C2=CC(=C(C(=C2)OC)OC)OC)O)O	2.6	335929
8	C1=CC(=CC=C1C=CC2=CC(=CC(=C2)O)O)O	3.1	445154
9	C1=CC(=CC(=C1)O)CCC2=CC=C(C=C2)O	3.5	181511
10	C1=CC=C(C=C1)COC2=CC=C(C=C2)O	3.4	7638
11	C1=CC=C(C=C1)C2C(O2)C3=CC=CC=C3	2.9	5742860
12	CCC(C1=CC=C(C=C1)O)C(CC)C2=CC=C(C=C2)O	5.2	3606
13	O(C1=CC(=CC(=C1)OC)\C=C(\C2=CC=C(OC)C=C2)[H])[H]C	4.1	5388063
14	COC1=CC=C(C=C1)C=CC2=CC(=C(C(=C2)OC)OC)OC	4.1	125922
15	COC1=C(C=C(C=C1)C(C(C2=CC(=C(C(=C2)OC)OC)OC)O)O)O)O	1.4	10247286
16	COC1=CC(=CC(=C1O)OC)C(CC2=CC(=C(C=C2)O)OC)OC	2.8	75149948
17	COC1=CC(=CC(=C1O)O)C(CC2=CC=C(C=C2)O)OC	2.5	74429419
18	CCOC(CC1=CC=C(C=C1)O)C2=CC(=C(C(=C2)OC)O)O	2.8	74429420
19	CC(C(=CC1=CC(=C(C=C1)OC)O)C2=CC(=C(C(=C2)OC)OC)OC)O)O	3.5	54586166
20	COC1=CC=C(C=C1)CC(C2=C(C(=C(C=C2)OC)OC)O)O	3.1	44429048
21	COC1=CC=C(C=C1)C(C(C2=CC(=C(C(=C2)OC)OC)OC)O)O)O	1.8	10592816
22	COC1=C(C=C(C=C1)CC(C2=CC(=CC(=C2)OC)OC)O)OC	2.5	66673695
23	COC1=CC=C(C=C1)CC(C2=CC(=C(C(=C2)OC)OC)OC)O	2.9	57423765
24	COC1=C(C(=C(C=C1)C(C(C2=CC(=C(C(=C2)OC)OC)OC)O)O)O)O)O	1.6	54129628
25	COC1=C(C=C(C=C1)C=C(CO)C2=CC(=C(C(=C2)OC)OC)OC)O	3.1	11078510
26	COC1=C(C=C(C=C1)C(CC2=CC(=C(C(=C2)OC)OC)OC)O)OC	2.9	356755
27	COC1=CC(=CC(=C1OC)OC)C(CC2=CC=CC=C2)O	2.9	353079
28	CC(=CC1=CC(=CC(=C1)OC)OC)C2=CC=C(C=C2)OC	4.7	75071221
29	COC1=CC=CC(=C1)C=CC2=CC(=CC(=C2)OC)OC	4.1	69452320
30	COC1=CC(=O)OC(C1)C=CC2=CC=CC=C2	2.5	5369129
31	COC1=CC(=CC(=C1)C=CC2=CC=CC=C2)OC	4.1	13556468
32	CC(=CC1=CC(=CC(=C1)OC)OC)C2=CC=CC=C2	4.8	68796507
33	O(C1=CC(=CC(=C1)OC)C=CC2=CC(=CC(=C2)OC)OC)C	4.1	67145168
34	COC1=CC(=CC(=C1)C=CC2=CC=C(C=C2)C=C)OC	4.9	70184295
35	CCOC1=CC=C(C=C1)C=CC2=CC(=CC(=C2)OC)OC	4.5	69899106
36	CC1=CC=C(C=C1)C=CC2=CC(=CC(=C2)OC)OC	4.5	58240360
37	CCOC1=CC=C(C=C1)C=CC2=CC(=CC(=C2)OCC)OCC	5.2	67435273
38	O(C2=C(C=CC1=CC(=CC(=C1)OC)OC)C=CC(=C2)OC)C	4.1	5491
39	COC1=CC(=CC(=C1)CC(=C)C2=CC=CC=C2)OC	4.8	69940018
40	CC(C)OC1=CC=C(C=C1)C=CC2=CC(=CC(=C2)OC)OC	4.9	66674282

(b) Validare externă

Valorile log P pentru setul test (Tabelul 61) au fost calculate folosind ecuația 10 din (cu trei variabile) Tabelul 63. Datele sunt prezentate în Tabelul 65 și corelația monovariată este: $\log P = 0.763 \times \log P_{calc.} + 0.876$; n=15; $R^2=0.859$; s=0.411; F=79.105 în graficul din Figura 42.

Tabelul 65. Valorile calculate

log P pentru moleculele din setul test		
Mol.	log P	log P _{calc.}
5	3.4	3.67
6	3.2	3.72
7	2.6	2.40
8	3.1	3.65
9	3.5	3.30
10	3.4	3.36
11	2.9	3.58
12	5.2	4.40
13	4.1	4.11
14	4.1	4.08
15	1.4	1.45
37	5.2	4.60
38	4.1	4.28
39	4.8	4.57
40	4.9	4.59

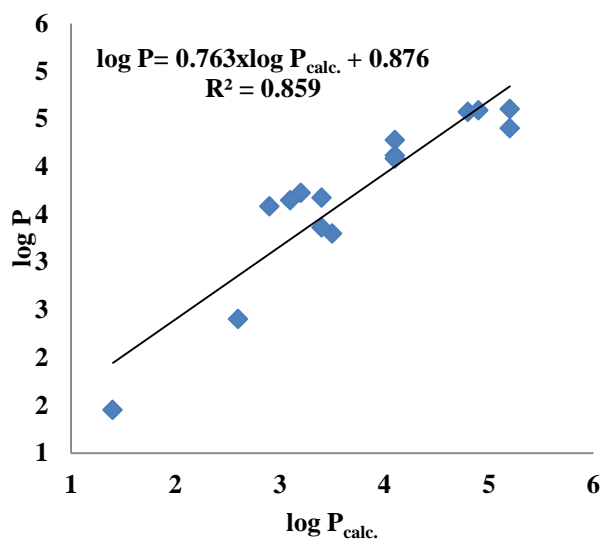


Figura 42. Graficul log P vs. log P_{calc.} pentru moleculele din setul test (validare externă)

(c) Validarea prin clustere de similaritate

Validarea poate fi realizată prin calcularea log P pentru moleculele din setul test cu ecuația de învățare pe clustere de similaritate: fiecare dintre cele 15 molecule este lider în propriul cluster, selectat prin similaritate (2D) printre cele 25 structuri a setului inițial. Valorile log P_{calc.} pentru fiecare dintre cele 15 molecule din setul test sunt calculate cu 15 ecuații noi (liderul fiind lăsat afară) cu același descriptor din ecuația 10, Tabelul 63. Datele sunt prezentate în Tabelul 66 și corelația monovariată este: $\log P = 0.923 \times \log P_{calc.} + 0.288$; n=15; $R^2=0.979$; s=0.157; F=622.623 este dat în Figura 43.

Tabelul 66. Valorile calculate log P prin clustere de similaritate, pentru moleculele din setul test (Tabelul1)

Mol.	log P	log P _{calc.}
5	3.4	3.31
6	3.2	3.31
7	2.6	2.57
8	3.1	3.28
9	3.5	3.32
10	3.4	3.39
11	2.9	3.15
12	5.2	4.88
13	4.1	4.12
14	4.1	4.08
15	1.4	1.52
37	5.2	5.08
38	4.1	4.27
39	4.8	4.60
40	4.9	5.03

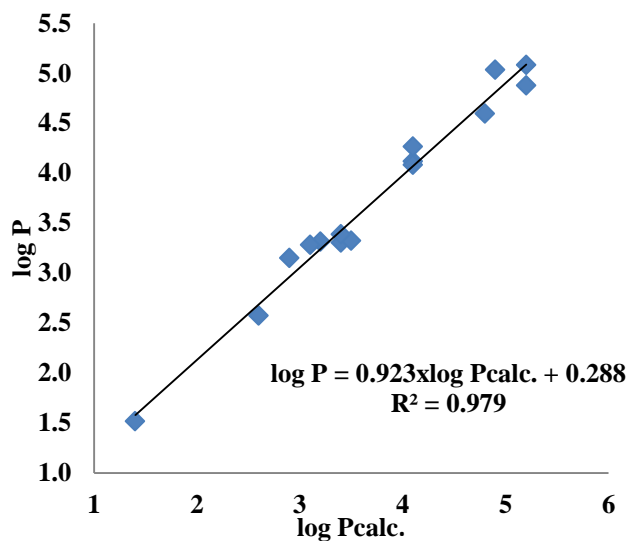


Figura 43. Graficul log P vs. log P calc. pentru setul test (clustere de similaritate)

Sarcină parțială (cazul 2)

Validarea modelului (pentru log P)

(b) Validare externă

Valorile log P pentru setul test de derivați de resveratrol (Tabelul 61), sunt calculate folosind cea mai bună ecuație 11 din Tabelul 67. Datele sunt prezentate în Tabelul 69 și corelația monovariată este: $\log P = 1.031 \times \log P_{calc.} - 0.051$; $n=15$; $R^2=0.938$; $s=0.213$; $F=195.279$ graficul în Figura 44 .

Tabelul 69. Valorile calculate
log P pentru moleculele din setul
test (Tabelul 61)

Mol.	log P	log P _{calc.}
1	3.1	3.26
2	3.7	3.86
3	3.1	2.80
4	3.4	3.50
5	3.4	3.88
6	3.2	3.34
7	2.6	2.34
8	3.1	3.40
9	3.5	3.36
10	3.4	3.54
11	2.9	2.74
12	5.2	5.03
13	4.1	4.31
14	4.1	4.25
15	1.4	1.37

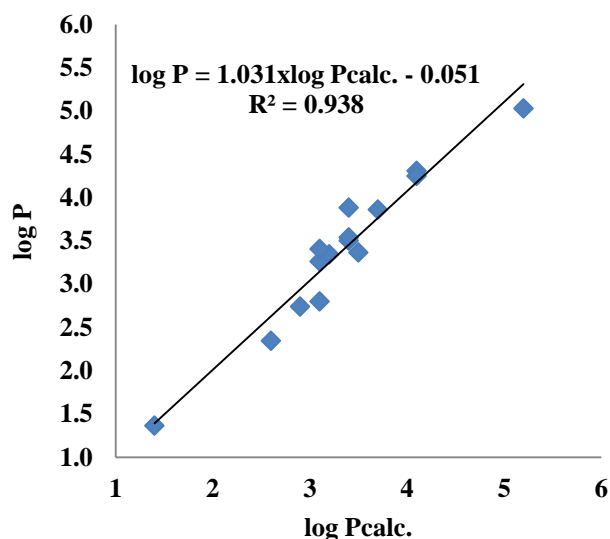


Figura 44. Graficul log P vs. log P_{calc.} pentru moleculele din setul
test (validare externă)

(c) Validarea prin clustere de similaritate

Valorile log P calc. pentru fiecare dintre cele 15 molecule din setul test sunt calculate cu același descriptori ca în ecuația 11, Tabelul 67. Datele sunt în Tabelul 70 și corelația monovariată este: $\log P = 1.020 \times \log P_{calc.} + 0.002$; $n=15$; $R^2=0.981$; $s=0.119$; $F=659.369$ graficul în Figura 45.

Tabelul 70. Valorile calculate log
P prin clustere de similaritate,
pentru moleculele din setul test
(Tabelul 1)

Mol.	log P	log P _{calc.}
1	3.1	3.29
2	3.7	3.84
3	3.1	2.90
4	3.4	3.53
5	3.4	3.58
6	3.2	3.34
7	2.6	2.69
8	3.1	3.13
9	3.5	3.42
10	3.4	3.55
11	2.9	2.97
12	5.2	5.14
13	4.1	4.31
14	4.1	4.23
15	1.4	1.34

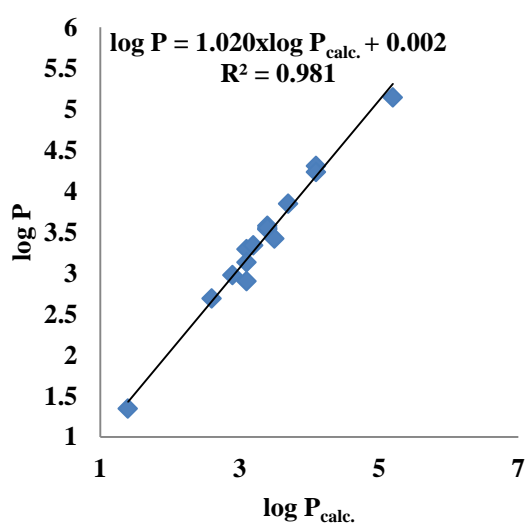


Figura 45. Graficul log P vs. log P calc. pentru setul test
(clustere de similaritate)

4.2. Studiul andocării moleculare

Studiul de andocare moleculară a fost efectuat pentru a explora modul de legare a derivaților de resveratrol (Tabelul 61) de stilben sintetaza și să înțelegem relația lor activitate- structură folosind software-ul de andocare AutoDock Vina [131, 132].

Pentru a studia interacțiunea dintre derivați de resveratrol și proteina 1Z1F (a se vedea figura 47), AutoDock Vina, un program de modelare moleculară a fost condusă, pentru a simula legarea între derivații de resveratrol și 1Z1F (Tabelul 68).

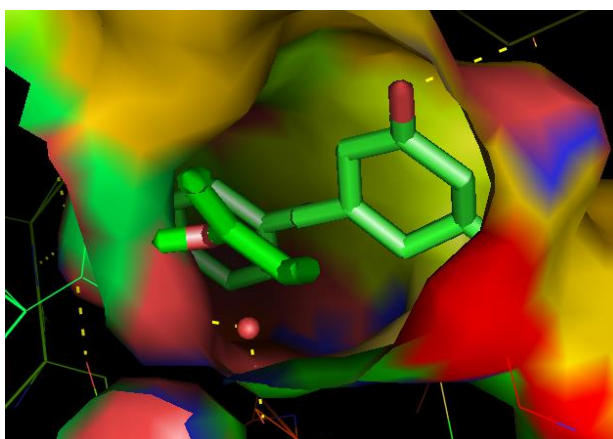


Figura 47: Interacțiunea resveratrolului cu stilben sintetaza

Pentru a obține un model de farmacofor pentru resveratrol receptorul stilben sintetaza a fost aleși conformeri cu cele mai favorabile interacțiuni cu receptorul care rezultă din andocare. Liganzii 8, 3, 32, 5, 31 au cea mai mică energie de andocare între -7.3 și -7.0, bazat pe acești compuși am construit farmacoforul. Farmacoforul rezultat este prezentat în Figura 49.

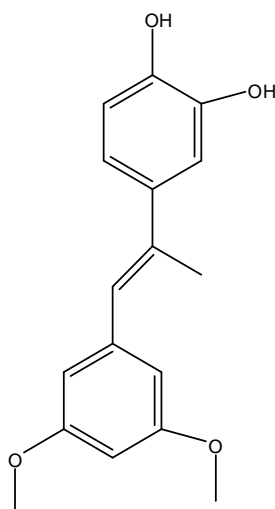


Figura 49: Modelul farmacofor pentru resveratrol

CONCLUZII GENERALE

În lucrarea de față s-a urmărit tratarea și obținere unor metode teoretice de obținere a unor noi compuși cu proprietăți biologice deosebite pentru a putea fii utilizați în industrie, obținerea de noi medicamente fără a fii nevoie de testarea lor pe animale astfel încât să se reducă costul de producție.

Docking-ul molecular permite aprecierea modului de legare a unei molecule de ligand în situsul activ al receptorului biologic proteina, astfel c a energia de legare să fie minimă. Andocarea moleculară necesită informații despre receptorul biologic, iar cu ajutorul metodei QSAR permite corelarea activității biologice manifestate de următoarele clase de compuși: flavonoide, testosteron, antrachinone și resveratrol.

O nouă abordare QSAR, bazate pe corelația masei sau a sarcini parțială a hipermoleculii, s-a realizat pe mai multe clase de compuși: flavonoide, testosteron, antrachinone, resveratrol care sau descărcat din baza de date PubCHEM. O nouă procedură similară cu „aliniamentul” moleculelor de medicament la receptorul biologic.

Setul de date a fost împărțit într-un set școală și set test, ultimul este folosit pentru validarea modelului, în așa numita “validarea setului extern”. De asemenea, validarea este realizată printr-o nouă versiune de predicție folosind clustererele de similaritate. Clustererele de similaritate permit realizarea unui set de structuri “quasi-congenere”, demonstrând o mai bună predicție decât în cazul clasic de validare externă.

BIBLIOGRAFIE

1. Cos P., Ying L., Calomme M., Jia P. Hu, Cimanga K., Van Poel B., Pieters L., Arnold J. Vlietinck, and Vanden D. Berghe, Structure-Activity Relationship and Classification of Flavonoids as Inhibitors of Xanthine Oxidase and Superoxide Scavengers, *J. Nat. Prod.* **1998**, 61, 71-76
2. Dixon R. A., , Isoflavonoids: biochemistry, molecular biology and biological functions, In U.Sankawa, *Comprehensive Natural Products Chemistry Elsevier*, **1999**, 1, 773-823.
3. Surpatin R., A theoretical study of 1,1-diphenyl-2-picrylhydrazyl (DPPH) radical scavenging activities of flavonoids using electrotopological state atoms (E-state) parameters. *Int. J. Pharm. Bio. Sci.*, **2012**, 3 (3), 543 – 550.

4. Aherne S.A., O'Brien N.M., Dietary Flavonols: Chemistry, Food Content, and Metabolism, Nutrition, **2002**, 18, 75-81.
5. Agarwal O.P., The antiinflammatory action of nepitrin, a flavonoid, Agents, Actions, **1982**, 12, 298-302.
6. Church M.K., Cromoglycate-like anti allergic drugs: a review, Drugs Today, **1978**, 14, 281-341.
7. Rasulev B.F., Abdullaev N. D., Syrov V.N. and Leszczynski J., A quantitative structure-activity relationship (QSAR) study of the antioxidant activity of flavonoids, QSAR & Comb. Sci., **2005**, 24, 1056-1065.
8. Free, S.M., Jr, Wilson, J.W. A mathematical contribution to structure-activity studies. J Med Chem. **1964**, 7, 395–399.
9. B. F. Rasulev, N. D. Abdullaev, V. N. Syrov and J. Leszczynski, A quantitative structure activity relationship (QSAR) study of the antioxidant activity of flavonoids, QSAR & Comb. Sci., **2005**, 24, 1056- 1065.
10. The RCSB Protein data bank. [<http://www.rcsb.org/pdb>].
11. Trott O., Olson A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, Journal of Computational Chemistry, **2010**, 31, 455-461.
12. Fernandez S. P., Wasowski C., Loscalzo L. M., R. E. Granger, G. A. R. Johnston, A. C. Paladini, M. Marder, Central nervous system depressant action of flavonoid glycosides. European Journal of Pharmacology, **2006**, 539, 168-176.
13. Cushnie T. P. T., , Lamb A. J., Antimicrobial activity of flavonoids. International Journal Of Antimicrobial Agents, **2005**, 26, 343-356.
14. Vagánek A., Rimarc̣ík J., Dropková K., Lengyel J., Klein E., Reaction enthalpies of OAH bonds splitting-off in flavonoids: The role of non-polar and polar solvent, Computational and Theoretical Chemistry, **2014**, 1050, 31–38.
15. Rice-Evans C. A., Miller N. J., Pagana G., Structure-antioxidant activity relationships of flavonoids and phenolic acids, Free Radical Biol. Med., **1996**, 20, 933–956.
16. Gülçin I., Antioxidant activity of food constituents: an overview, Arch. Toxicol., **2012**, 86, 345–391.
17. Yeon J. Y., Bae Y. J., Kim E. Y., Lee E. J., Association between flavonoid intake and

- diabetes risk among, *Clinica Chimica Acta*, **2015**, 439, 225–230.
18. Yang K. J., Lin S. C., Huang S. J., Ching W. M., Hung C. H., Tzou D. L. M., Solid-state NMR study of fluorinated steroids, *Steroids* **2014**, 80, 64–70.
 19. Kammerer R.C., Drug testing and anabolic steroids. In: Yesalis CE, editor. *Anabolic steroids in sports and exercise*, Champaign: human kinetics, **2000**, 2, 415–59.
 20. Shahidi N.T., A review of the chemistry, biological action, and clinical applications of anabolic–androgenic steroids. *Clinical Therapeutics* **2001**;23(9):1355–90.
 21. Liu, P.Y., Death, A.K., Handelsman, D.J., Androgens and cardiovascular disease. *Endocr. Rev.* **2003** 24, , 313—340
 22. Choong, K., Basaria, S., Emerging cardiometabolic complications of androgen deprivation therapy. *Aging Male* **2010**, 13 , 1—9.;
 23. Traish, A.M., Miner, M.M., Morgentaler, A., Zitzmann, M., Testosterone deficiency. *Am. J. Med.* **2011**, 124, 578—587.;
 24. Pereg D., Chan J., Russell E., Berlin T., Mosseri M., Seabrook J. A., Koren G., Van Uum S., Cortisol and testosterone in hair as biological markers of systolic heart failure, *Psychoneuroendocrinology*, **2013**, 38, 2875—2882
 25. Arnlöv, J., Pencina, M.J., Amin, S., Nam, B.H., Benjamin, E.J., Murabito, J.M., Wang, T.J., Knapp, P.E., D’Agostino Sr., R.B., Bhasin, S., Vasan, R.S., Endogenous sex hormones and cardiovascular disease incidence in men. *Ann. Intern. Med.* **2006**, 145, 176-184.
 26. Jankowska, E.A., Biel, B., Majda, J., Szklarska, A., Lopuszanska, M., Medras, M., Anker, S.D., Banasiak, W., Poole-Wilson, P.A., Ponikowski, P., Anabolic deficiency in men with chronic heart failure: prevalence and detrimental impact on survival. *Circulation* **2006**, 114, 1829—1837.
 27. Nishiyama T., Ph.D., M.D. Serum testosterone levels after medical or surgical androgen deprivation: A comprehensive review of the literature, *Urologic Oncology: Seminars and Original Investigations*, **2014**, 32, 38, 17–38.
 28. Weisman O., Zagoory-Sharon O., Feldman R., Oxytocin administration, salivary testosterone, and father–infant social behavior, *Progress in Neuro-*

- Psychopharmacology & Biological Psychiatry **2014**, *49*, 47–52.
29. Prado-Prado, O. Martinez V., Uriarte E., Ubeira F.M, Chou K.-C., González Díaz H., Unified QSAR approach to antimicrobials. Multi-target QSAR modeling and comparative multi-distance study of the giant components of antiviral drug–drug complex networks, *Bioorg. Med. Chem.* **2008**, doi:10.1016/j.bmc.2008.11.075.
 30. Thomson, R.H., Naturally Occurring Quinones IV. *Springer*, London, **1996**.
 31. Blum, M.S., Hilker, M., Chemical protection of insect eggs. In: Hilker, M., Meiners, T. (Eds.), *Chemoecology of Insect Eggs and Egg Deposition*. Blackwell Publishing, Berlin, Oxford, **2002**, 61–90.
 32. Matasyoh, J.C., Dittrich, B., Schueffler, A., Laatsch, H., Larvicidal activity of metabolites from the endophytic *Podospora* sp. against the malaria vector *Anopheles gambiae*. *Parasitol. Res.* **2011**. 108, 561–566.
 33. Izhaki, I., Emodin: a secondary metabolite with multiple ecological functions in higher plants. *New Phytol.* **2002**, 155, 205–217.
 34. Dragos D., Heghes A., Medeleanu M., Vlaia V., Seiman C., Kaycesa A., Ciubotariu D., Topological similarity/ Dissimilarity indicators: Application to cytochrome P450 inhibition by alcohols *Journal Medical Timisoara*, **2004**, 54, 2, 128-134.
 35. Koblitz H (1988). Anthraquinones. In: F Constabel and IK. Vasil eds. Cell culture and somatic cell genetics of plants. Orlando: Academic Press, p. 113.
 36. Fremont L., Biological effects of resveratrol. Life Sciences, Acta Zhengzhou University Overseas Edition. **2000**, 66, 8, 663.
 37. Larifa S., Salem C.B., Hmouda H., Bouraoui K., Journal of Molecular Graphics and Modelling, **2014**, 53, P.M. Kris-Etherton, and C.L. Keen, *Curr. Opin. Lipidol.*,**2002**, 13, 1, 41-49.
 38. Russo A., Palumbo M., Aliano C., Lempereur L., Scoto G., and Renis M., “Red wine micronutrients as protective agents in Alzheimer-like induced insult, Life Sciences, Acta Zhengzhou University Overseas Edition..., **2003**, 72, 21, 2369.
 39. Surh Y., Molecular mechanisms of chemopreventive effects of selected dietary and medicinal phenolic substances. *Mutat Res.*,**1999**, 428, 1-2, 305.
 40. Snyder, Scott A., Zografos, Alexandros L. and Lin, Y., Total Synthesis of Resveratrol-Based Natural Products: A Chemoselective Solution. *Angew. Chem. Int. Ed.*,**2007**, 46, 8186–8191. doi: 10.1002/anie.200703333

41. Alarcón de la Lastra, C. and Villegas, I., Resveratrol as an anti-inflammatory and anti-aging agent: Mechanisms and clinical implications. *Mol. Nutr. Food Res.*, **2005**, 49: 405–430. doi: 10.1002/mnfr.200500022
42. Wang, L.-X., Heredia, A., Song, H., Zhang, Z., Yu, B., Davis, C. and Redfield, R. Resveratrol glucuronides as the metabolites of resveratrol in humans: Characterization, synthesis, and anti-HIV activity. *J. Pharm. Sci.*, **2004**, 93: 2448–2457. doi: 10.1002/jps.20156
43. Dismukes W.E., *Clinical Infectious Diseases, Introduction to Antifungal Drugs*, **2000**, 30, 653-657.
44. Rogers D., and Hopfinger A. J., Application of Genetic Function Approximation to Quantitative Structure-Activity Relationships and Quantitative Structure-Property Relationships *J. Chem. Inf. Comput. Sci.*, **1994**, 34, 854, DOI: 10.1021/ci00020a020
45. Zhu, H., Shen, Z., Tang, Q., Ji, W., & Jia, L. Degradation mechanism study of organic pollutants in ozonation process by QSAR analysis. *Chemical Engineering Journal*, **2014**, 255, 431-436.
46. M.A. Johnson, G.M. Maggiora, *Concepts and Applications of Molecular Similarity*, Eds., John Wiley & Sons, New York, **1990**.
47. Moldovan C. D., Costescu A., Katona G. and Diudea M. V., A novel QSAR descriptors, *MATCH Commun. Math. Comput. Chem.*, **2008**, 60, 977-984.
48. Goa KL, Barradell LB. Fluconazole: an update of its pharmacodynamic and pharmacokinetic properties and therapeutic use in major superficial and systemic mycoses in immunocompromised patients. *Drugs* **1995**;50,658–90.
49. Nurettin M., Ozlem S., Yusif A., Safiye Sağ Erdem, and Metin B., Design and Synthesis of Pyrrolotriazepine Derivatives: An Experimental and Computational Study, *J. Org. Chem.*, **2013**, 78 (11), 5184–5195.
50. Hayes, A., 'Principles and Methods of Toxicology', Taylor and Francis, 4th Ed, 18, **2001**.
51. Sheridan R. P. and Kearsley S. K., Why do we need so many chemical similarity search methods?, *Drug Discovery Today*, **2002**, 7, 903.
52. Misik J., Vodakova E., Pavlikova R., Cabal R., Novotny L., Kamil Kuca., Acute toxicity of surfactants and detergent-based decontaminants in mice and rats., *Mil. Med. Sci. Lett. (Voj. Zdrav. Listy)*, **2012**, 81(4), p. 171-176
53. Willett P., Barnard J. M., Downs G. M., “Chemical similarity searching,” *Journal of Chemical Information and Computer Sciences*, **1998**, 38(6), 983–996.

54. Lyman W.J., Lyman In W.J. and Rosenblatt, Handbook of Chemical Property Estimation Methods: Environmental Behaviour of Organic Compounds, American Chemical Society, Washington, DC, **1990**, 1.
55. Doo Ho Cho, Sung Kwang Lee, Bum Tae Kim, Tai No K., Quantitative Structure-Activity Relationship (QSAR) Study of New Fluorovinylacetamides, Bull. Korean Chem. Soc. 2001, 22, 4.
56. Zollanvari A., Braga-Neto U., Dougherty E. R., Exact representation of the second-order moments for resubstitution and leave-one-out error estimation for linear discriminant analysis in the univariate heteroskedastic Gaussian model, Pattern Recognition **2012**, 45, 908.
57. Diudea, M. V. Layer Matrices in Molecular Graphs. J. Chem. Inf. Comput. Sci. **1994**, 34, 1064-1071;
58. Diudea, M. V.; Topan, M. I.; Graovac, A. Layer Matrices of Walk Degrees. J. Chem. Inf. Comput. Sci. **1994**, 34, 1072-1078;
59. M. Randić, W.L. Woodworth, A. Graovac, Unusual random Walks, Int. J. Quantum Chem., 1983, 24, 435-452.
60. M. Randić, Generalized molecular descriptors, J. Math. Chem., 1991, 7, 155-168
61. Balaban, A. T.; Motoc, I.; Bonchev, D.; and Mekenyan, O. "Topological Indices for Structure-Activity Correlations." Top. Curr. Chem. **114**, 21-55, 1983.
62. Diudea, M.V.; Horvath, D.; Graovac, A. 3-D Distance Matrices and Related Topological Indices. J. Chem. Inf. Comput. Sci. **1995**, 35, 129-135.
63. Mircea V. Diudea,, Gabriel Katona, István Lukovits, and Nenad Trinajsti, Detour and Cluj-Detour Indices, Croatica chemica acta, **1988**, 71 (3) 459-471 .
64. Bates, D. M. and Watts, D. G. , Nonlinear Regression: Iterative Estimation and Linear Approximations, in Nonlinear Regression Analysis and Its Applications, John Wiley & Sons, Inc., Hoboken, NJ, USA. **1988**, doi: 10.1002/9780470316757.ch2
65. McCormick D., Roach A., Measurement, Statistics and Computation, John Wiley & Sons, London, **1987**.
66. Cawley, G. C., & Talbot, N. L. C., Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers. Pattern Recognition, **2003**,36 (11), 2585-2592.
67. Kearns, M. J., & Ron, D., Algorithmic stability and sanity-check bounds for leave-one-out crossvalidation. Computational Learning Theory, **1997**, 152– 162.

68. Kohavi, R., A study of cross-validation and bootstrap for accuracy estimation and model selection. *IJCAI*, **1995**, 1137–1145.
69. Kohavi, R., A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence IJCAI*, **1995**, 1137–1145.
70. Zhang, T., Leave-one-out bounds for kernel methods. *Neural Computation*, **2003**, 15, 1397–1437.
71. Stone, M. Asymptotics for and against crossvalidation. *Biometrika*, **1997**, 64, 29–35.
72. DiMasi, Joseph A., Ronald W. Hansen, and Henry G. Grabowski. "The price of innovation: new estimates of drug development costs." *Journal of health economics* 22.2 **2003**, 151-185.
73. Holliday, John D., C. Y. Hu, and Peter Willett. "Grouping of coefficients for the calculation of inter-molecular similarity and dissimilarity using 2D fragment bit-strings." *Combinatorial chemistry & high throughput screening*, **2002**, 5.2, 155-166..
74. Barnard, John M., and Geoffrey M. Downs. "Clustering of chemical structures on the basis of two-dimensional similarity measures." *Journal of chemical information and computer sciences* 32.6 (1992): 644-649.
75. Poornima, C. S., and P. M. Dean. "Hydration in drug design. 2. Influence of local site surface shape on water binding." *Journal of computer-aided molecular design*, **1995**,9.6, 513-520.
76. Poroikov, V. V., et al. "Robustness of biological activity spectra predicting by computer program PASS for noncongeneric sets of chemical compounds." *Journal of chemical information and computer sciences*, **2000**, 40.6, 1349-1355.
77. Anderberg MR, *Clustering analysis for applications*. London, Academic Press, **1973**, 359-362.
78. Costescu, Adina, et al. "QSAR modeling of human catechol O-methyltransferase enzyme kinetics." *Journal of mathematical chemistry*, **2009**, 45.2, 287-294.
79. Klopmand, G., *Concepts and applications of molecular similarity*, by Mark A. Johnson and Gerald M. Maggiora, eds., John Wiley & Sons, New York, 1990, **1992**, 393-394. Price: \$65.00. *J. Comput. Chem.*, 13: 539–540. doi: 10.1002/jcc.540130415
80. Carbó, R., Leyda, L. and Arnau, M., How similar is a molecule to another? An electron density measure of similarity between two molecular structures. *Int. J. Quantum Chem.*,**1980**, 17, 1185–1189. doi: 10.1002/qua.560170612

81. Bergström, S., V. G. Bundoc, and A. G. Barbour. "Molecular analysis of linear plasmid-encoded major surface proteins, OspA and OspB, of the Lyme disease spirochaete *Borrelia burgdorferi*." *Molecular microbiology*, **1989**, 3.4, 479-486.
82. Makino, S. and Kuntz, I. D., Automated flexible ligand docking method and its application for database search. *J. Comput. Chem.*, **1997**, 18, 1812–1825.
83. Blaney, J.M. & Dixon, J.S., A good ligand is hard to find: automated docking methods. *Perspectives in Drug Design*, **1993**, 1, 301-319.
84. Gehlhaar, D.K., Verkhivker, C., Rejto, P.A., Fogel, D.B., Fogel, L.J. & Freer, S.T., Docking conformationally flexible small molecules into a protein binding site through evolutionary programming. In *Proceedings of the Fourth Annual Conferen*, **1995**.
85. Li, Z., et a/., & Cohen, F.E., Anti-malarial drug development using models of enzyme structure. *Chemistry & Biology*, **1994**, 1, 31-37.
86. Cehlhaar, D.K., Moerder, K.E., Zichi, D., Sherman, C.J., Ogden, R.C. & Freer, S.T. De nova design of enzyme inhibitors by Monte Carlo ligand generation. */. Med. Chem.*, **1995**, 38, 4661172.
87. Huang S. Y. and Zou X., Advances and Challenges in Protein-Ligand Docking, *Int. J. Mol. Sci.*, **2010**, 11,3016-3034.
88. Schneider G., Virtual screening: an endless staircase? *Nat Rev Drug Discov*, **2010**, 9, 273–6.
89. Leurs R., Smit M. J., and Timmerman H., "Molecular pharmacological aspects of histamine receptors," *Pharmacol Ther*, **1995**, 66, 413-463.
90. Dixon SL, Smodyrev AM, Rao SN. PHASE: a novel approach to pharmacophore modeling and 3d database searching. *Chemical Biology & Drug Design*, **2006**, 67, 370–372.
91. Fei J, Zhou L, Liu T, Tang XY. Pharmacophore modeling, virtual screening, and molecular docking studies for discovery of novel Akt2 inhibitors. *International Journal of Medical Sciences*, **2013**, 10, 265–275 DOI 10.7150/ijms.5344.
92. Guner OF. History and evolution of the pharmacophore concept in computer-aided drug design. *Current Topics in Medicinal Chemistry*, **2002**, 2, 1321–1332.
93. **PubChem** database, accessed 23.05. 2013.

94. Emad M. Atta¹, Adel A. Nassar , Nawal M. Hasan and Ahmad R. Hasan, New Flavonoid Glycoside and Pharmacological Activities of *Pteranthus dichotomus* Forssk., *Rec. Nat. Prod.*, **2013**, 7, 2, 69-79.
95. Balaban A.T., Chiriac A., Motoc I., and Simon Z., *Steric Fit in QSAR (Lectures Notes in Chemistry, Vol. 15)*, Springer, Berlin, **1980**, Chap. 6.21.
96. Balaban, A. T., Can topological indices transmit information on properties but not on structure?, *J. Comput. Aided Design*, **2005**, 19, 651–660
97. Berinde Z. and Drinkal, C., Molecular modelling of molar refraction for alkanes, *Chem. Bull. "Politehnica" Univ. (Timisoara)*, **2004**, 49 (63) 1-2, 4–7
98. Harsa, Alexandra. "A novel qsar approach in modeling hydrophobicity of a set of flavonoids." *Studia Universitatis Babes-Bolyai, Chemia*, **2014**, 59.1.
99. Randic, Milan. "Molecular shape profiles." *Journal of chemical information and computer sciences* , **1995**, 35.3, 373-382.
100. Diudea, Mircea V., Ivan Gutman, and Lorentz Jantschi. *Molecular topology*. Huntington, NY: Nova Science Publishers, **2001**.
101. **PubMed** database, accessed 23.05. 2014.
102. Balaban A. T., Chiriac A., Motoc I., and Simon Z., *Steric Fit in QSAR (Lectures Notes in Chemistry, Vol. 15)*, Springer, Berlin, **1980**.
103. **Gaussian 09**, Gaussian Inc Wallingford CT, Revision A.1, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam NJ, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ, **2009**.
104. Anne-Sophie Carpentier, Alessandra Riva, Pierre Tisseur, Gilles Didier, Alain Hénaut, The operons, a criterion to compare the reliability of transcriptome analysis tools: ICA is more reliable than ANOVA, PLS and PCA, *Computational Biology and Chemistry*, **2004**, 28, 1, 3–10

105. Johan A. Westerhuis, Theodora Kourti and John F. Macgregor, Analysis of multiblock and hierarchical PCA and PLS models, *Journal of Chemometrics*, **1998**, 12, 301–321
106. Alexandra M. HARSA, Teodora E. HARSA, Sorana D. BOLBOACĂ, Mircea V. DIUDEA, 2014, Qsar In Flavonoids By Similarity Cluster Prediction, *Current Computer-Aided Drug Design*, 10:115-128, Bentham Science (ISSN: 1875-6697 (online), doi:10.2174/1573409910666140410104542).
107. Hawkins D. M., Basak S. C., and Mills D., Assessing model fit by cross-validation, *J. Chem. Inf. Comput. Sci.*, **43**, **2003**, 579-586.
108. Bolboacă S. D., Jäntschi L., and Diudea M. V., Molecular Design and QSARs with Molecular Descriptors Family, *Current Computer-Aided Drug Design*, **9**(2), **2013**, 195-205.
109. Jäntschi L., LOO Analysis (LOO: leave one out), AcademicDirect Library of software, **2005**, Available at: http://l.academicdirect.org/Chemistry/SARs/MDF_SARs/loo/
110. Topliss J.G., Costello R.J., Change correlations in structure-activity studies using multiple regression analysis. *J. Med. Chem.*, **1972**, 15, 1066-1068.
111. O'Boyle N. M., Banck M., James C. A., Morley C., Vandermeersch T., Hutchison G. R., "[Open Babel: An open chemical toolbox](#)". *Journal of Cheminformatics* , **2011**, 3, 33.
112. Modolo L. V. , Li L. , Pan H. , Blount J. W., Dixon R. A. , Wang X., Crystal structures of glycosyltransferase UGT78G1 reveal the molecular basis for glycosylation and deglycosylation of (iso)flavonoids, *J.Mol.Biol*, **2009**, 392, 1292-1302.
113. Garrett M. M, David S. G., Robert S. H., Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *Journal of Computational Chemistry*, **1998**, 19, 14, 1639-1662 .
114. Dhananjayan K., Arunachalam S., Sivanandy P., Molecular Docking Studies of Anthraquinone Derivatives on the Crystal Structure of Glycogen Synthase Kinase 3 Beta, *Electronic Journal of Biology*, **2014**, Vol. 10, 1, 14-20.

115. Ursu O., Diudea M. V., "TOPOCLUJ software program", Babes-Bolyai University, Cluj, **2005**.
116. HyperChem 7.52, Hypercube, Inc., <http://www.hyper.com>.
117. PyMol 0.97, DeLano Scientific, <http://pymol.org>.
118. Alexandra M. Harsa · Teodora E. Harsa · L. Jantschi · Mircea V. Diudea, QSAR study on Testosterone derivatives March 2015 · Journal of Chemical and Pharmaceutical Research
119. Richmond N.J., Abrams C.A., Wolohan P.R., Abrahamian E., Willett P., Clark R. D., GALAHAD: 1. Pharmacophore identification by hypermolecular alignment of ligands in 3D, *J Comput Aided Mol Des*, **20**, **2006**, , 567–587.
120. Balaban A. T., Chiriac A., Motoc I., and Simon Z., *Steric Fit in QSAR*, Lectures Notes in Chemistry, Springer, Berlin, **15**, **1980**.
121. Harsa T. E., Harsa A. M., Szeffler B. QSAR of Caffeines by similarity cluster prediction, *Cent. Eur. J. Chem.*, **2014**, **12**, 365-376.
122. Alexandra M. Harsa, Teodora E. Harsa & Mircea V. Diudea, QSAR and docking studies of anthraquinone derivatives by similarity cluster prediction, *Journal of Enzyme Inhibition and Medicinal Chemistry*, **2015** DOI:10.3109/14756366.2015.1046061,
123. Gianoncelli, Alessandra. "design, synthesis and biochemical characterization of inhibitors of protein kinases." **2010**.
124. Trott, O. Olson A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *Journal of Computational Chemistry*, **2010**, **31** 455-461.
125. Jubie S., Kalirajan R. and Pavankumar Y., Design, Synthesis and Docking Studies of a Novel Ciprofloxacin Analogue as an Antimicrobial AGENT, *E-Journal of Chemistry*, **2012**, **9**, **2**, 980-987.
126. Sanner, M. F. Python: a programming language for software integration and development. *Journal of Molecular Graphics Modeling*, **1999**, **17**, 57–61.
127. Dhananjayan K., Kalathil K., Sumathy, A., Sivanandy P., A computational study on binding affinity of Bio-flavonoids on the crystal structure of 3-hydroxy-3-methyl-glutaryl-CoA reductase – An *insilico* molecular docking approach, *Der Pharma*

Chemica, **2014**, 6, 2, 378-387

128. Nagy Cs. L., Diudea M. V., Nano Studio software package, "Babes-Bolyai" University, Cluj, **2009**.
129. Harsa, Alexandra M., Teodora E. Harsa, and Mircea V. Diudea. "QSAR studies on derivatives of resveratrol." *Studia Universitatis Babes-Bolyai, Chemia*, **2014**, 59, 4.
130. Bolboacă S. D., Jäntschi L., and Diudea M. V., Molecular Design and QSARs with Molecular Descriptors Family, *Current Computer-Aided Drug Design*, **2013**, 9(2), 195.
131. Romano T. Kroemer, Structure-Based Drug Design: Docking and Scoring, *Current Protein and Peptide Science*, **2007**, 8, 312-328.
132. Bottegoni, G. et al., Four-dimensional docking: a fast and accurate account of discrete receptor flexibility in ligand docking. *J. Med. Chem.*, **2009**, 52, 397–406.
133. Ankur Dhanik_, John S McMurray and Lydia Kavraki, AutoDock-based incremental docking protocol to improve docking of large ligands, To appear in the 2012 Proceedings of IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW) **2012**, 1-8.
134. Žiemys A., Rimkutė L., Kulys J., Modeling the Enantioselective Enzymatic Reaction with Modified Genetic Docking Algorithm, *Nonlinear Analysis: Modelling and Control*, **2004**, 9, 4, 373–383
135. Jones G., Willett P., Glen R.C., Leach A.R., Taylor R., "Development and validation of a genetic algorithm to flexible docking", *J. Mol. Biol.*, **1997**, 267, 727– 748.
136. Halgren T.A., Murphy R.B., Friesner R.A., Beard H.S., Frye L.L., Pollard W.T. Banks J.L. "Glide: A New Approach for Rapid, Accurate Docking and Scoring.2. Enrichment Factors in Database Screening", *J. Med. Chem.*, **2004**, 47, 1750–1759.