

PCE ID-458

Biochimie Versus Biomatematică în Medicina Moleculară

- lucrare în extenso -

2007

Pentru etapa unică a anului 2007 obiectivele planificate au fost:

1. *Documentare la zi asupra metodelor de studiu relații structură-activitate*
2. *Elaborare model matematic*

Activitățile prevăzute a se desfășura au fost:

- **Obiectivul 1:**

1. *Identificare și colectare metode SAR*
2. *Dobândire competențe complementare prin participări la manifestări științifice/stagii de documentare-cercetare*
3. *Sumarizare și stocare metode SAR*
4. *Achiziție, instalare, testare și configurare aparatură suport.*

- **Obiectivul 2:**

1. *Dobândire competențe complementare prin participări la manifestări științifice/stagii de documentare-cercetare*
2. *Elaborare specificații și metodologie model matematic*
3. *Testare și validare model.*

Activitatea 1. *Identificare și colectare metode SAR*

Relațiile structură-activitate (SAR), structură-proprietate (SPR), și proprietate-activitate (PAR) au apărut odată cu studiile publicate de Louis Plack HAMMETT în 1937 [LP Hammett, The Effect of Structure upon the Reactions of Organic Compounds. Benzene Derivatives, J Am Chem Soc, 1937, 59(1), 96-103].

Un relativ recent review însumează cele mai importante contribuții ale lui LP Hammett [C Hansch, A Leo, RW Taft, A Survey of Hammett Substituent Constants and Resonance and Field Parameters, Chem Rev, 1991, 91, 165-195].

Relațiile cantitative (QSAR, QPSR, QPAR) intervin atunci când proprietatea și/sau activitatea care este subiect al investigației este o mărime cantitativă. Nu toate proprietățile și activitățile moleculelor chimice pot fi încadrate în categoria cantitativă. Câteva exemple sunt LD50 (Median Lethal Dose, 50% - doza necesară pentru a omorî jumătate din populația supusă testării), Sweetness (dulce - unul dintre cele 5 gusturi de baza, fiind aproape universal relatat ca o experiență a plăcerii) zaharurilor, care poate fi apreciată doar prin comparație (scară relativă), neexistând două referințe și o scară (așa cum este cazul la punctul de fierbere, care are ca referințe înghețul și evaporarea apei și intervalul divizat în 100).

Nici proprietățile unanim acceptate ca fiind cantitativ exprimate nu au același grad de acuratețe al exprimării. Din acest motiv în ultima vreme se evita a se folosi QSAR, QPRR, și QPAR, folosindu-se în schimb (Q)SAR, (Q)SPR, (Q)PAR, sau mai simplu SAR, SPR, PAR. În ce privește structura lucrurile stau ceva mai simplu. Astfel un atom, o legătură într-o moleculă poate să existe (și este atunci evidențiat prin analiza tranzițiilor electronice ale moleculei), sau nu (adică este o problemă de 0 și 1).

Nu chiar atât de simplu stau lucrurile în privința geometriei moleculare. Principiul lui Heisenberg (Werner HEISENBERG, 1901-1976, unul din fondatorii mecanicii cuantice, laureat Nobel) ne arată prin principiul incertitudinii că la nivel micro (nivelul atomic și molecular) domnește incertitudinea. Mai mult decât atât, geometria moleculară depinde de mediul în care molecula se află (vecinătatea acesteia), temperatura, presiune, etc, astfel a vorbi despre geometria moleculară este cel puțin o problemă de relativitate, dacă nu de incertitudine.

Aplicațiile soft actuale sunt capabile să folosească modele cuantice și/sau semi-empirice pentru a prezice geometria moleculară in vitro și/sau vivo, așa cum este cazul aplicației HyperChem [HyperCube Inc], pe care am folosit-o.

Însa și aici este o barieră care "separa apele". Astfel, proprietăți care depind de energetica moleculară (cum este exemplul tipic al energiei de legătură) concordă de la relativ bine la foarte bine cu geometria calculată, în timp ce alte proprietăți (cum este doza letală), și în special acele proprietăți

și/sau activități selective și/sau specifice (care depind foarte mult de geometrie și în mai mică măsură de topologie) concordă doar cel mult relativ bine cu modelul geometric calculat al moleculei.

În acest domeniu al relațiilor structură-proprietate-activitate (numite și SPAR) avem parte de certitudini (topologia moleculară), incertitudini (geometria moleculară), relativități (activitățile biologice) și evidențe (proprietățile fizico-chimice).

Sintetizarea de noi compuși chimici biologic activi, cu performante superioare și/sau costuri mai mici de producție care să servească la tratamentul medical aplicat oamenilor și/sau animalelor, obținerea de noi soiuri de plante transgenice și mai bună conservare a mediului ambiant este principalul obiectiv al *platformei tehnologice chimie durabilă* (<http://www.suschem.org>).

În ultima perioadă de timp, indicii structurali folosiți pentru modelarea de compuși biologic activi prin studii integrate structură-activitate (SAR – structure activity relationship) sunt tot mai frecvent calculați din parametri sterici (geometrici) și/sau electrostatici (sarcini parțiale) [1] în detrimentul parametrilor pur topologici [2]. Sunt preferate calcule semiempirice și cuantice cu programe ca Hondo95, Gaussian94, Gamess, Icon08, Tx90, Polyrate, Unichem/Dgauss, Allinger's MM3, Mopac93, Mozyme și HyperChem [3].

Analiza de regresie structură/activitate folosește metode clasice ca regresia liniară, liniară multiplă, neliniară, sisteme expert sau rețele neuronale pentru baze mari de date [4].

Elaborarea, implementarea și folosirea modelului molecular pentru obținerea de noi compuși biologic activi este referită frecvent în fluxul de informații. Iată un rezumat al acestor preocupări.

Ca metoda preliminară de analiză, unii autori aliniază setul de molecule. Mai mult, metoda CoMFA introduce un algoritm în 6 pași pentru analiza structură-activitate [5]:

1. construiește setul școală de molecule cu activitate biologică cunoscută și generează structura 3D a moleculelor (folosind programe ca: Mopac, Sybyl [6-7], HyperChem [8-9], Alchemy2000 [6], MolConn [10]);
3. alege o metodă de suprapunere (care poate fi suprapunerea de fragmente alese din molecule [5], [11,12] sau suprapunerea de grupuri farmacofore [13] și suprapune virtual coordonatele spațiale;
4. construiește un grid care înconjoară moleculele suprapuse la pasul (2) într-o formă standard sau modificată ([14]) și alege un atom de probă pentru interacția cu punctele gridului [15,16];
5. utilizează o metoda empirică (Hint [17]), un model specific (suprapunere farmacoforă [18]), energia potențială clasică (Lennard-Jones, Coulomb), energia legăturilor de hidrogen [19], câmpul generat de orbitalii moleculari [20,21] sau orice alt model definit de utilizator [15] și calculează valorile de interacțiune în gridul de la pasul (3) a câmpului de interacțiune ales cu atomul probă de la pasul (3);
6. folosește valorile calculate ale interacțiunii de la pasul (4) între punctele gridului și atomul de probă pentru a face estimarea activității cunoscute construind o relație structură-activitate;

7. folosește relația structură activitate obținută la pasul (5) pentru a face predicția activității pentru molecule cu aceeași suprapunere cu setul școala de la pasul (1).

Metoda CoMFA este un bun instrument pentru predicția unui variat tip de activități biologice cum sunt citotoxicitatea [22], inhibiția [16,21], sau proprietățile de formare [23,24]. Mai mult, metoda servește în modelarea compușilor cu efecte farmaceutice [13,25] și inhibitorilor HIV [26].

Revenind asupra investigațiilor structurale pe compuși biologic activi, o activitate foarte importantă este căutarea substructurilor active biologic din compușii biologic activi care produc cea mai mare parte a răspunsului biologic măsurat [17].

Una din modalitățile de realizare a căutării substructurilor active biologic este identificarea invarianților moleculari. În acest sens metoda WHIM (Weighted Holistic Invariant Molecular) calculează un set de indicatori statistici derivat din proprietățile sterice și electrostatice ale moleculelor [28-30]. O variantă a acestei metode, MS-WHIM (de la Molecular Surface) servește în analiza suprafeței moleculare [31]. MS-WHIM este o colecție de 36 indicatori statistici derivați din proprietățile sterice și electrostatice și este orientată spre parametrizarea suprafeței moleculare [32].

Relațiile cantitative structură-activitate QSAR (Q - quantitative) referă în general o activitate biologică măsurabilă iar tehnica de investigare ce folosește QSAR este o tehnica modernă folosită astăzi în multe domenii prioritare, incluzând farmaceutica, mediul, biotehnologia și microbiologia. Literatura de specialitate conține astăzi un număr mare de relații structură-activitate care evaluează impactul produselor chimice asupra mediului [33-37].

Modelul matematic folosit pentru obținerea relațiilor structură-activitate este cel mai frecvent bazat pe regresia liniară [38] și pe rețelele neuronale artificiale [39-41].

Surse citate in text:

[1] Computational approaches to structural and functional analysis of plastocyanin and other blue copper proteins, proteins, F. De Rienzo, R. R. Gabdouliline, R. C. Wade, M. Sola, and M. C. Menziani, Cellular and Molecular Life Sciences, Springer-Verlag, Issue: Volume 61, Number 10, Date: May 2004, Pages: 1123-1142.

[2] Atomistic Topological Indices Applied to Benzodiazepines using Various Regression Methods, Wikler D. A., Burden F. R., Watkins A. J. R, Quantitative Structure Activity Relationships, Elsevier, 1998, 17(1), p. 14-19.

[3] Computation of the influence of chemical substitution on the pKa of pyridine using semiempirical and ab initio methods, I-Jen Chen and Alexander D. MacKerell Jr, Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta), Springer-Verlag, Volume 103, Number 6, Date: April 2000, Pages: 483-494.

- [4] Development of an Expert System for the Remediation of Petroleum-Contaminated Sites, Z. Chen, G.H. Huang, C.W. Chan, L.Q. Geng and J. Xia, *Environmental Modeling and Assessment*, Springer Science+Business Media B.V., Volume 8, Number 4, Date: December 2003, Pages: 323-334.
- [5] CoMFA, HQSAR and molecular docking studies of butitaxel analogues with b-tubulin, Suzanne L. Cunningham, Albert R. Cunningham and Billy W. Day, *Journal of Molecular Modeling*, Volume 11, Number 1, Date: February 2005, Pages: 48-54.
- [6] Unity Program for SIMCA (Soft Independent Modeling Class Analogy); Tripos Associates, St. Louis, MO.
- [7] Alfred Merz, Didier Rognan, Gerd Folkers, 3D QSAR Study of N2-phenylguanines as Inhibitors of Herpes Simplex Virus Thymidine Kinase, Antiviral and Antitumor Research, <http://www.pharma.ethz.ch/text/research/tk/qsar.html>.
- [8] Gurba P. E., Parham M. E., Voltano J. R., Comparison of QSAR Models Developed for Acute Oral Toxicity (LD50) by Regression and Neural Network Techniques, Conference on Computational Methods in Toxicology - April, 1998, Holiday Inn/I-675, Dayton, Ohio, USA, abstract available at <http://www.ccl.net/ccl/toxicology/abstracts/abs9.html>.
- [9] HyperChem, Molecular Modelling System; Hypercube Inc.,
<http://hyper.com/products/Professional/>
- [10] Molconn-Z, <http://www.eslc.vabiotech.com/molconn>.
- [11] Waller C. L., Wyrick S. D., Park H. M., Kemp W. E., Smith F. T., Conformational Analysis, Molecular Modeling, and Quantitative Structure-Activity Relationship Studies of Agents for the Inhibition of Astrocytic Chloride Transport, *Pharm. Res.*, 1994, 11(1), p. 47-53.
- [12] Horwitz J. P., Massova I., Wiese T., Wozniak J., Corbett T. H., Sebolt-Leopold J. S., Capps D. B., Leopold W. R., Comparative Molecular Field Analysis of in Vitro Growth Inhibition of L1210 and HCT-8 Cells by Some Pyrazoloacridines, *J. Med. Chem.*, 1993, 36(23), p. 3511-3516.
- [13] McGaughey G. B., Mewshaw R. E., Molecular Modeling and the Design of Dopamine D2 Partial Agonists, (presented at the Charleston Conference; march; 1998), submitted in may 1998, Network Science, <http://www.netsci.org/Science/Compchem/feature20.html>.
- [14] Chuman H., Karasawa M., Fujita T., A Novel Three-Dimensional QSAR Procedure: Voronoi Field Analysis, *QSAR*, 1998, 17(4), p. 313-326.
- [15] Walter C. L., Kellogg G. E., Adding Chemical Information of CoMFA Models with Alternative 3D QSAR Fields.
- [16] Merz A., Rognan D., Folkers G., 3D QSAR Study of N2-phenylguanines as Inhibitors of Herpes Simplex Virus Thymidine Kinase, Antiviral and Antitumoral Research,
<http://www.pharma.ethz.ch/text/research/tk/qsar.html>.

- [17] Kellogg G. E., Semus S. F., Abraham D. J., HINT: a new method of empirical hydrophobic field calculation for CoMFA, *J. Comput.-Aided Mol. Des.*, 1991, 5(6), p. 545-552.
- [18] Myers A. M., Charifson P. S., Owens C. E., Kula N. S., McPhail A. T., Baldessarini R. J., Booth R. G., Wyrick S. D., Conformational Analysis, Pharmacophore Identification, and Comparative Molecular Field Analysis of Ligands for the Neuromodulatory σ_3 Receptor, *J. Med. Chem.*, 1994, 37(24), p. 4109-4117.
- [19] Kim K. H., Use of the hydrogen-bond potential function in comparative molecular field analysis (CoMFA): An extension of CoMFA.
- [20] Durst G. L., Comparative Molecular Field Analysis (CoMFA) of Herbicidal Protoporphyrinogen Oxidase Inhibitors using Standard Steric and Electrostatic Fields and an Alternative LUMO Field.
- [21] Waller C.L., Marshall G. R., Three-Dimensional Quantitative Structure-Activity Relationship of Angiotensin-Converting Enzyme and Thermolysin Inhibitors. II. A Comparison of CoMFA Models Incorporating Molecular Orbital Fields and Desolvation Free Energy Based on Active-Analog and Complementary-Receptor-Field Alignment Rules, *J. Med. Chem.*, 1993, 36, p. 2390-2403.
- [22] Wiese M., Pajeva I. L., A Comparative Molecular Field Analysis of Propafenone-type Modulators of Cancer Multidrug Resistance, *Quant. Struct.-Act. Relat.*, 1998, 17(4), p. 301-312.
- [23] Klebe G., Abraham U., On the Prediction of Binding Properties of Drug Molecules by Comparative Molecular Field Analysis, *J. Med. Chem.*, 1993, 36(1), p. 70-80.
- [24] Czaplinski K.H.A., Grunewald G. L., A Comparative Molecular Field Analysis Derived Model of Binding of Taxol Analogs to Microtubes, *Bioorg. Med. Chem. Lett.*, 1994, 4(18), p. 2211-2216.
- [25] Akagi T., Exhaustive Conformational Searches for Superimposition and Three-Dimensional Drug Design of Pyrethroids, *QSAR*, 1998, 17(6), p. 565-570.
- [26] Waller C.L., Oprea T.I., Giolitti A., Marshall G.R., Three-Dimensional QSAR of Human Immunodeficiency Virus. (I) Protease Inhibitors. 1. A determined Alignment Rules, *J. Med. Chem.*, 1993, 36(26), p. 4152-4160.
- [27] Thompson E., The Use of Substructure Search and Relational Databases for Examining the Carcinogenic Potential of Chemicals; Conference on Computational Methods in Toxicology – April, 1998, Holiday Inn/I-675, Dayton, Ohio, USA; abstract available at <http://www.ccl.net/ccl/toxicology/abstracts/tabs6.html>.
- [28] Todeschini R., Lasagni M., Marengo E., New Molecular Descriptors for 2D and 3D Structures. Theory *J. Chemometrics*, 1994, 8, p. 263-272.
- [29] Todeschini R., Gramatica P., Provenzani R., Marengo E., Weighted Holistic Invariant Molecular (WHIM) descriptors. Part2. Their Development and Application on Modeling Physico-chemical Properties of Polyaromatic Hydrocarbons, *Chemometrics and Intelligent Laboratory Systems*, 1995, 27, p. 221-229.

- [30] Todeschini R., Vighi M., Provenzani R., Finizio A., Gramatica P., Modeling and Prediction by Using WHIM Descriptors in QSAR Studies: Toxicity of Heterogeneous Chemicals on *Daphnia Magna*, *Chemosphere*, 1996, 8, p. 1527.
- [31] Zaliani A., Gancia E., MS-WHIM Scores for Amino Acids: A New 3D-Description for Peptide QSAR and QSPR Studies, *J. Chem. Inf. Comput. Sci.*, 1999, 39(3), p. 525-533.
- [32] Bravi G., Gancia E., Mascagni P., Pegna M., Todeschini R., Zaliani A., MS-WHIM., New 3D Theoretical Descriptors Derived from Molecular Surface Properties: A Comparative 3D QSAR Study in a Series of Steroids, *J. Comput.-Aided Mol. Des.*, 1997, 11, p. 79-92.
- [33] Niemelä J.R., Validation of the BIODEG Probability Program, *TemaNord Repost*, 1994, 589, p. 153-156.
- [34] Damborsky J., A mechanistic approach to deriving quantitative structure-activity relationship models for microbial degradation of organic compounds. SAR and QSAR in Environmental Research, Proceedings of the Satellite Workshop on Biodegradation accompanying the 6th International Workshop on QSAR in Environmental Sciences, Italy, September 12, 1994.
- [35] Blok J., Classification of biodegradability by growth kinetic parameters, *Ecotoxicology and Environmental Safety*, 1994, 27, p. 294-305.
- [36] Warne M. A., Ebbels T. M. D., Lindon J. C., Nicholson J. K., Semiempirical Molecular-Orbital Properties of Some Polycyclic Aromatic Hydrocarbons and Correlation with Environmental Toxic Equivalency Factors, *Polycyclic Aromatic Compounds*, 2003, 23, 23-74.
- [37] Schultz T. W., Cronin M. T. D., Walker J. D., Aptula A. O., Quantitative structure- activity relationships (QSARs) in toxicology: a historical perspective, *Journal of Molecular Structure: THEOCHEM*, 2003, 622(1), p. 1-22.
- [38] Schultz T. W., Cronin M. T. D., Netzeva T. I., The present status of QSAR in toxicology, *Journal of Molecular Structure: THEOCHEM*, 2003, 622, p. 23-38.
- [39] Sparks T. C., Crouse G. D., Durst G., Natural products as insecticides: the biology, biochemistry and quantitative structure-activity relationships of spinosyns and spinosoids, *Pest Management Science*, 2001, 10, 896-905.
- [40] Kompare B. Estimating environmental pollution by xenobiotic chemicals using QSAR (QSBR) models based on artificial intelligence, *Water Science and Technology*, 1998, 37(8), p. 9-18.
- [41] Marjan Vracko, Kohonen Artificial Neural Network and Counter Propagation Neural Network in Molecular Structure-Toxicity Studies, *Current Computer - Aided Drug Design*, 2005, 1(1), p. 73-78.

Familia de Descriptori Moleculari (MDF) pe relații structură-activitate (SAR) și -proprietate (SPR)

Modelul Matematic al MDF

Un membru al MDF este construit din 7 piese, fiecare piesa dând o literă în denumirea acestuia.

1► Prima literă este operatorul de linearizare. Acesta face legătura între nivelul micro și nivelul nano/macro (de exemplu pH-ul $\text{pH} = -\log[\text{H}^+]$ este mărimea (efectul) "macro" măsurată a fenomenului (cauzei) "micro" și anume prezența ionilor H^+ în soluții.

2► A doua literă este operatorul de suprapunere globală moleculară. Existența acestuia este justificată de multitudinea de efecte și interacțiuni care se manifestă între atomii ce compun moleculele.

3► A treia literă este criteriul de fragmentare. Justificarea acestuia află răspuns încă de la primele studii efectuate de Hammett, care a observat ca contribuția diferitelor părți ale moleculei (fragmentele) este esențială și net diferită de la unul la celălalt (substituent).

4► A patra literă este modelul de interacțiune. Justificarea rațiunii acestuia se află încă de la Thales (magnetism), Newton (gravitație), Coulomb (electrostatică), și culminând cu Schrödinger (ecuația Schrödinger, pentru care a primit premiul Nobel în 1933) care au pus în evidență câmpul, potențialul și energia, în aproximația sa terestră și în expresia sa exactă.

5► A cincea literă este formula descriptorului de interacțiune. Diferitele tipuri de potențial, câmpuri și forțe au expresii diferite și aceasta este un fapt.

6► A șasea literă este proprietatea atomică. Fiecare tip de atom are o serie de proprietăți care îl diferențiază și îl aseamănă cu altul, observație care a dus la realizarea tabelului lui Mendeleev.

7► A șaptea literă este operatorul de distanță. Așa cum s-a argumentat în preambul, geometria (3D - 3 dimensiuni) și topologia (2D - 2 dimensiuni) oferă deopotrivă avantaje și dezavantaje.

MDF folosește modele fizice cuantice

Astfel fiecare posibilitate pentru fiecare dintre cele 7 piese ale sale are cel puțin câte un model fizic concret asociat.

Lista acestor posibilități împreună cu semnificația acestora este disponibilă online: [Model Fizic MDF](#).

Numărul de posibilități concretizat în modelul fizic al MDF este redat mai jos:

- 1► 6 posibilități;
- 2► 19 posibilități;
- 3► 4 posibilități;
- 4► 6 posibilități;
- 5► 24 posibilități;
- 6► 6 posibilități;
- 7► 2 posibilități;

Astfel, pentru construcția MDF se calculează un număr de 787968 descriptori. Nu toți au sens fizic (cum ar fi logaritmul dintr-un număr negativ) și nu toți furnizează informație distinctă (calitatea ca 2

descriptori cu formule de calcul diferite sa aibă aceeași valoare se numește degenerare; degenerarea poate apare atât pentru o moleculă dată cât și pentru un întreg set de molecule luat în studiu). De asemenea nu toți sunt finiți (o împărțire cu 0 generează un număr infinit). Toate aceste posibilități ies din cadrul natural și sunt eliminați printr-o procedură numită de bias. Rămân doar aproximativ (depinde de la un set de molecule investigat la altul) 100000 de membrii MDF valizi.

MDF foloseste selectia naturala [algoritm genetic]

Experiența anterioară a arătat că aproape niciodată cel mai bun descriptor (cel care corelează cel mai mult cu proprietatea măsurată) nu se regăsește printre descriptorii ce compun perechea de descriptori a celei mai bune relații în 2 variabile. Astfel, MDF folosește perechi de descriptori pentru obținerea de relații structură-activitate/proprietate în mai multe variabile.

Metodologia MDF

Date de intrare:

- ▶ Formule chimice moleculare si/sau structurale pentru un set de compuși
- ▶ Valorile proprietății/activității (sau mai multe) măsurate pe setul de compuși

Date de ieșire:

- ▶ Relația structură-activitate/proprietate

Pași:

- 1▶ Construiește modelul topologic (2D) al fiecărei molecule din set cu ajutorul aplicației

HyperChem

- 2▶ Construiește modelul geometric (3D) al fiecărei molecule din set cu ajutorul aplicației

HyperChem

- 3▶ Aplică un model semiempiric (pentru calculul distribuției de sarcini parțiale pe atomi) și (unde este cazul) cuantic mergând până la cele mai avansate teorii cuantice (Ab-inițio și Time-Dependent Density Functional Theory), folosind module specifice ale aplicației HyperChem (ex: HyperNewton, HyperGauss, HyperNDO) pentru obținerea unui model geometric in vitro sau in vivo optimizat.

- 4▶ Generează familia MDF

- 5▶ Aplică procedura de bias

- 6▶ Obține relații structură-activitate/proprietate cu 1,2, sau mai mulți descriptori de structură

Metodologia MDF-SAR

Date de intrare:

- ▶ Model topologic (2D) și geometric (3D) pentru fiecare moleculă din set (format HyperChem)
- ▶ Valorile proprietății/activității (sau mai multe) măsurate pe setul de compuși
- ▶ Ecuație (sau ecuații) MDF cu 1 sau mai mulți descriptori

- ▶ Valori prezise ale proprietății/activității obținute cu alte modele (din literatura de specialitate)

Proceduri (disponibile online):

- ▶ Statistica descriptivă și inferențială
- ▶ Statistica predictivă "LOO" (leave-one-out)
- ▶ Experiment de inteligență artificială "TvT-E" (training versus test experiment)
- ▶ Analiza corelațiilor corelate (testul Steiger) [Procedura offline].

MDF-SAR Drug Design

Această facilitate a MDF-SAR permite ca având:

- ▶ Un set de compuși de interes cu proprietate/activitate cunoscută și model MDF-SAR obținut, validat și stocat în baza de date
 - ▶ Unul sau mai mulți compuși similari/înrușiți cu setul ales sa se obțină folosind
 - ▶ Ecuația MDF-SAR
 - ▶ Modelul topologic (2D) și geometric (3D) obținute folosind aceleași modele semiempirice/cuantice ca și la moleculele setului
- predicția proprietății/activității compusului (compușilor) noi, chiar dacă aceștia nu au fost încă sintetizați, pentru a vedea dacă noua structură (virtuală în acest moment) aduce sau nu îmbunătățiri în proprietatea/activitatea dorită.

Baza de date MDF

Sunt de fapt 2 baze de date tip MySQL, ambele stocate pe un server FreeBSD în IntraNet [IP:172.27.211.5].

- ▶ `MDFSARs` (la data de 26 Sept 2007 conține 246 tabele, 3.5Gb) - conține seturile pe care s-au finalizat studiile SAR și s-au obținut ecuațiile MDF-SAR.

- ▶ `MDFSARtmp` (la data de 26 Sept 2007 conține 64 tabele, 0.6 Gb) - conține seturile în lucru.

Pentru fiecare set se generează:

- ▶ tabela ``"NumeSet"_tmpx` (câmpuri:molecule, înregistrări:descriptori); are totdeauna 131328 înregistrări (787968/6 - fără linearizare)
- ▶ tabela ``"NumeSet"_data` (câmp:proprietate/activitate, înregistrări:molecule); are totdeauna un număr de înregistrări egal cu numărul de molecule incluse în studiu
- ▶ tabela ``"NumeSet"_valx` (câmpuri:molecule, înregistrări:descriptori); rezultă în urma procedurii de bias; numărul de înregistrări variază și se situează în jurul valorii de 100000
- ▶ tabela ``"NumeSet"_valy` (câmpuri:M(X);M(X*X);M(X*Y);r2(X,Y), înregistrări:descriptori); are același număr de înregistrări cu "NumeSet"_valx; M - operatorul de medie; Y - proprietatea/activitatea măsurată; X - descriptor MDF; r2 - coeficient de determinare; este utilă și folosită în regresiiile multivariate (scade dramatic timpul de execuție)

► tabela `0_MDFSARRes` (una singura per baza de date) conține toate modelele MDF-SAR obținute pe seturile din baza de date; are structura:[name(ume set);eq(ecuație MDF-SAR);r2(coeficient de determinare);m(număr de molecule în set);n(număr de descriptori MDF în ecuația de regresie)

Remarca: Conform celor arătate până aici, n de mai sus joacă rolul de "număr de variabile independente în ecuația de regresie", proprietatea măsurată jucând rolul de "variabila dependentă în ecuația de regresie".

► tabela `ready` (una singura per baza de date) conține de obicei o singură înregistrare, numele setului care este supus investigației curente și este folosită pentru interacțiuni server-client, de către aplicațiile dedicate pentru regresii multiple, care se execută simultan pe mai multe stații de lucru în intranet.

Aplicații dedicate pentru generarea MDF

► au fost realizate folosind limbajul PHP

► se execută pe un server din IntraNet [IP:172.27.211.4] respectând o anumită succesiune

Sucesiunea este următoarea:

► 0_mdf_prepare.php - creează structura tabelor ce vor stoca informațiile; folosește denumirile de director (pt NumeSet) și fișiere (pt NumeMolecula)

► 1_mdf_generate.php - calculează descriptorii pentru fiecare moleculă (poate fi executat pentru fiecare moleculă în parte sau simultan sau pentru tot setul odată); scrie în tabela ``NumeSet``_tmpx`

► 2_mdf_linearize.php - aplica operatorul de linearizare, scrie doar descriptori valizi în tabelele ``NumeSet``_xval` și ``NumeSet``_yval` (procedura monotasking)

► 3_mdf_bias.php - sortează în memorie după r2, șterge simultan degenerările din ambele tabele ``NumeSet``_xval` și ``NumeSet``_yval`

► 4_mdf_order.php - sortează în memorie după r2, creează 2 tabele temporare similare cu ``NumeSet``_xval` și ``NumeSet``_yval`, copiază în ordinea descrescătoare a lui r2 din ``NumeSet``_xval` și ``NumeSet``_yval` în noile tabele, și la sfârșit șterge tabelele vechi, și le redenumeste pe cele noi

Aplicații client-server pentru MDF-SAR

Acestea referă aplicațiile dedicate pentru obținerea de ecuații de regresie multivariate (cea mai bună ecuație de regresie cu o singură variabilă independentă fiind deja obținută la pasul anterior).

► Căutare sistematică în 2 variabile independente

► Căutare sistematică în 3 variabile independente (una fixată și dată ca Dată de Intrare)

► Căutare sistematică în 4 variabile independente (2 fixate și date ca Date de Intrare)

► Căutare sistematică în N variabile independente (N-2 fixate și date ca Date de Intrare)

► Căutare aleatorie evolutivă în N (N>2) variabile independente (se selectează perechea de 2 variabile ce urmează a fi înlocuite pe baza analizei de regresie cu N-2 variabile

► Căutare aleatorie (random) în N variabile independente

- Căutare sistematică în 3 variabile independente (consumatoare de memorie ~2Gb și timp ~120zile)
- aplicație nefinalizată (nepractică)

Rezultate obținute folosind MDF și publicate în extenso

| Nr | Publicație |
|----|--|
| 1 | Lorentz JÄNTSCHI, Molecular Descriptors Family on Structure Activity Relationships 1. Review of the Methodology, Leonardo Electronic Journal of Practices and Technologies, AcademicDirect, ISSN 1583-1078, www, Internet, 4(6), p. 76-98, 2005. |
| 2 | Sorana Daniela BOLBOACA, Lorentz JÄNTSCHI, Molecular Descriptors Family on Structure Activity Relationships 2. Insecticidal Activity of Neonicotinoid Compounds, Leonardo Journal of Sciences, AcademicDirect, ISSN 1583-0233, www, Internet, 4(6), p. 78-85, 2005. |
| 3 | Sorana Daniela BOLBOACA, Lorentz JÄNTSCHI, Molecular Descriptors Family on Structure Activity Relationships 3. Antituberculosic Activity of some Polyhydroxyxanthenes, Leonardo Journal of Sciences, AcademicDirect, ISSN 1583-0233, www, Internet, 4(7), p. 58-64, 2005. |
| 4 | Lorentz JÄNTSCHI, Sorana Daniela BOLBOACA, Molecular Descriptors Family on Structure Activity Relationships 4. Molar Refraction of Cyclic Organophosphorus Compounds, Leonardo Electronic Journal of Practices and Technologies, AcademicDirect, ISSN 1583-1078, www, Internet, 4(7), p. 55-102, 2005. |
| 5 | Lorentz JÄNTSCHI, Mihaela Ligia UNGURESAN, Sorana Daniela BOLBOACA, Complex Structural Information Integration: Inhibitor Activity on Carbonic Anhydrase II of Substituted Disulfonamides, Applied Medical Informatics, SRIMA, ISSN 1224-5593, Cluj-Napoca, Romania, 17(3-4), p. 12-21, 2005. |
| 6 | Lorentz JÄNTSCHI, Sorana Daniela BOLBOACA, Molecular Descriptors Family on QSAR Modeling of Quinoline-based Compounds Biological Activities, The 10th Electronic Computational Chemistry Conference, Monmouth University, New Jersey, USA, N/A, www, Internet, Paper #4, April 1-30, 2005. |
| 7 | Lorentz JÄNTSCHI, Microbiology and Toxicology and Phytochemistry Studies (in Romanian), AcademicDirect, , ISBN 973-86211-8-6, www, Internet, 75 p., 2005. |
| 8 | Monica STEFU, Daniela BUTYKA, Mircea V. DIUDEA, Lorentz JÄNTSCHI, Bazil PÂRV, Nanostructures: Novel Architecture - Mircea V. DIUDEA (ed.), Chapter: Algorithms for Basic Operations on Maps, Nova Science, , ISBN 1-59454-499-9, Huntington, New York, USA, 420 p., 98\$, 2005. |

| | |
|----|--|
| 9 | Lorentz JÄNTSCHI, Sorana BOLBOACA, Molecular Descriptors Family on Structure Activity Relationships 5. Antimalarial Activity of 2,4-Diamino-6-Quinazoline Sulfonamide Derivates, Leonardo Journal of Sciences, AcademicDirect, ISSN 1583-0233, www, Internet, 5(8), p. 77-88, 2006. |
| 10 | Lorentz JÄNTSCHI, Sorana BOLBOACA, Molecular Descriptors Family on Structure Activity Relationships 6. Octanol-Water Partition Coefficient of Polychlorinated Biphenyls, Leonardo Electronic Journal of Practices and Technologies, AcademicDirect, ISSN 1583-1078, www, Internet, 5(8), p. 71-86, 2006. |
| 11 | Sorana Daniela BOLBOACA, Lorentz JÄNTSCHI, Modeling of Structure-Toxicity Relationship of Alkyl Metal Compounds by Integration of Complex Structural Information, Therapeutics: Pharmacology and Clinical Toxicology, RP Press, ISSN 1583-0012, Bucuresti, Romania, X(1), p. 110-114, 2006. |
| 12 | Sorana BOLBOACA, Claudia FILIP, Stefan TIGAN, Lorentz JÄNTSCHI, Antioxidant Efficacy of 3-Indolyl Derivates by Complex Information Integration, Clujul Medical, Editura Iuliu Hatieganu, ISSN 1222-2119, Cluj-Napoca, Romania, LXXIX(2), p. 204-209, 2006. |
| 13 | Sorana BOLBOACA, Lorentz JÄNTSCHI, Pearson Versus Spearman, Kendall's Tau Correlation Analysis on Structure-Activity Relationships of Biologic Active Compounds, Leonardo Journal of Sciences, AcademicDirect, ISSN 1583-0233, www, Internet, 5(9), p. 179-200, 2006. |
| 14 | Lorentz JÄNTSCHI, Sorana BOLBOACA, Modelling the Inhibitory Activity on Carbonic Anhydrase IV of Substituted Thiadiazole- and Thiadiazoline- Disulfonamides: Integration of Structure Information, Electronic Journal of Biomedicine, Red UniNet Spain, ISSN 1697-090X, www, Internet, 2006(2), p. 22-33, 2006. |
| 15 | Sorana D. BOLBOACA, Lorentz JÄNTSCHI, Molecular Descriptors Family on Structure-Activity Relationships: Modeling Herbicidal Activity of Substituted Triazines Class, Bulletin of University of Agricultural Sciences and Veterinary Medicine - Agriculture, AcademicPres, ISSN 1454-2382, Cluj-Napoca, Romania, 62, p. 35-40, 2006. |
| 16 | Sorana Daniela BOLBOACA, Stefan TIGAN, Lorentz JÄNTSCHI, Molecular Descriptors Family on Structure-Activity Relationships on anti-HIV-1 Potencies of HEPTA and TIBO Derivatives, Integrating Biomedical Information: From eCell to ePatient, European Federation for Medical Informatics, ISBN 3-89838-0722-6 (Aka, Germany), ISBN 1-58603-614-9 (IOS Press, Holland), ISBN 973-625-303-1 (Politehnica Timisoara, Romania), Timisoara, Romania, p. 110-114, April 6-8, 2006. |

| | |
|----|--|
| 17 | Lorentz JÄNTSCHI, Sorana-Daniela BOLBOACA, Modeling the Octanol-Water Partition Coefficient of Substituted Phenols: the Use of Structure Information, Third Humboldt Conference on Computational Chemistry, InnosLab Ltd. , ISBN 954-323-199-0 & 978-954-323-199-7, Varna, Bulgaria, p. 65, June 24-28, 2006. |
| 18 | Stefan TIGAN, Lorentz JÄNTSCHI, Sorana-Daniela BOLBOACA, Modeling Herbicidal Activity of a Substituted Triazines Class by Integration of Compounds Complex Structural Information, XXIII International Biometric Conference, International Biometric Society, N/A, Montreal, Quebec, Canada, TP1.219 (509.pdf on CD), July 16-21, 2006. |
| 19 | Lorentz JÄNTSCHI, Sorana-Daniela BOLBOACA, Online System for Molecular Descriptors Family on Structure-Activity Relationships: Assessment and Characterization of Biologic Active Compounds, 6th European Conference on Computational Chemistry, European Association for Chemical and Molecular Sciences, N/A, Tale, Slovakia, September 3-7, 2006. |
| 20 | Sorana Daniela BOLBOACA, Lorentz JÄNTSCHI, Activity Characterization of Triazines Analogues: Statistical Parameters for Models Assessment, International Symposium on Organic Chemistry, December 9-12, 2006, Avangard Prisma, ISBN 954-323-243-1 & 978-954-323-243-7, Sofia, Bulgaria, p. 48-49 [Plenary oral presentation], December 9-12, 2006. |
| 21 | Lorentz JÄNTSCHI, Sorana Daniela BOLBOACA, Molecular Descriptors Family on Structure-Activity and Structure-Property Relationships: Results, SizeMat: Workshop on Size-Dependent Effects in Materials for Environmental Protection and Energy Application, Institute of General and Inorganic Chemistry, Bulgarian Academy of Sciences [http://sizemat.igic.bas.bg], FP6: EC-INCO-CT-2005-016414 Specific Support Action, Varna, Bulgaria, p. 14-15 [Plenary oral presentation], May 25-27, 2006. |
| 22 | Lorentz JÄNTSCHI, Mircea V. DIUDEA, Subgraphs by Pairs of Vertices, TOPMOL2006 - 20 Years Anniversary of Molecular Topology at Cluj, Babes-Bolyai University, CEEEx M3 PR-D11-PT00-41, Cluj-Napoca, Romania, p. 16 [Plenary oral presentation], September 25-30, 2006. |
| 23 | Lorentz JÄNTSCHI, Carmen Elena STOENOIU, Sorana Daniela BOLBOACA, Use of Graph Polynomials for Characterization of Alkanes Properties, International Symposium on Organic Chemistry, December 9-12, 2006, Avangard Prisma, ISBN 954-323-243-1 & 978-954-323-243-7, Sofia, Bulgaria, p. 87-88, December 9-12, 2006. |
| 24 | Lorentz JÄNTSCHI, Madalina Ana VALEANU, Sorana Daniela BOLBOACA, Rapid Programming of Relational Databases Applications (in Romanian), AcademicDirect & |

| | |
|----|--|
| | AcademicPres, , ISBN: 973-86211-5-1 & 978-973-86211-5-2 (AcademicDirect) && ISBN 973-744-044-7 & 978-973-744-044-0 (AcademicPres), www & Cluj-Napoca, Internet & Romania, 233 p., 2006. |
| 25 | Lorentz JÄNTSCHI, Sorana BOLBOACA, Results from the Use of Molecular Descriptors Family on Structure Property/Activity Relationships, International Journal of Molecular Sciences, Molecular Diversity Preservation International, ISSN 1422-0067, www & Basel, Internet & Switzerland, 8(3), p. 189-203, 2007. |
| 26 | Carmen Elena STOENOIU, Sorana Daniela BOLBOACA, Lorentz JÄNTSCHI, Characterization of Marine Sediments Toxicity based on Structural Information, Institute of General and Inorganic Chemistry, Bulgarian Academy of Sciences [http://sizemat.igic.bas.bg], FP6: EC-INCO-CT-2005-016414 Specific Support Action, Plovdiv, Bulgaria, p. 54, April 19-21, 2007. |
| 27 | Sorana Daniela BOLBOACA, Lorentz JÄNTSCHI, Modeling the Property of Compounds from Structure: Statistical Methods for Models Validation, Institute of General and Inorganic Chemistry, Bulgarian Academy of Sciences [http://sizemat.igic.bas.bg], FP6: EC-INCO-CT-2005-016414 Specific Support Action, Plovdiv, Bulgaria, p. 71, April 19-21, 2007. |
| 28 | Sorana D. BOLBOACA, Lorentz JÄNTSCHI, Data Mining on Structure-Activity/Property Relationships Models, ECC11- The 11th Electronic Computational Chemistry Conference, online, Monmouth University, New Jersey, USA, N/A, www, Internet, paper #29, April 2-30, 2007. |
| 29 | Lorentz JÄNTSCHI, Sorana-Daniela BOLBOACA, Modeling the Octanol-Water Partition Coefficient of Substituted Phenols by the Use of Structure Information, International Journal of Quantum Chemistry, Wiley InterScience, ISSN 0020-7608, eISSN 1097-461X, New York, NY, USA, Volume 107, Issue 8 (Special Issue: Proceedings from the 3rd Humboldt Conference on Computational Chemistry . Issue Edited by Georgi N. Vayssilov, Tzonka Mineva), p. 1736-1744, 2007. |
| 30 | Sorana D. BOLBOACA, Lorentz JÄNTSCHI, How Good the Characteristic Polynomial Can Be for Correlations?, International Journal of Molecular Sciences, Molecular Diversity Preservation International, ISSN 1422-0067, www & Basel, Internet & Switzerland, 8(4), p. 335-345, 2007. |
| 31 | Lorentz JÄNTSCHI, Sorana D. BOLBOACA, The Jungle of Linear Regression Revisited, Leonardo Electronic Journal of Practices and Technologies, AcademicDirect, ISSN 1583-1078, www, Internet, 6(10), p. 169-187, 2007. |

| | |
|----|---|
| 32 | Lorentz JÄNTSCHI, Characteristic and Counting Polynomials of Nonane Isomers, AcademicDirect, , ISBN 973-86211-3-5 & ISBN 980-973-86211-3-8, www, Internet, p. 101, 2007. |
| 33 | Sorana D. BOLBOACA, Lorentz JÄNTSCHI, Modelling the Inhibitory Activity on Carbonic Anhydrase I of Some Substituted Thiadiazoleand Thiadiazoline-Disulfonamides: Integration of Structure Information, Computer-Aided Chemical Engineering, Elsevier Netherlands & UK, eISSN 1570-7946, Amsterdam, Netherlands, 24(2007), p. 965-970, 2007. |
| 34 | Lorentz JÄNTSCHI, Sorana D. BOLBOACA, Antiallergic Activity of Substituted Benzamides: Characterization, Estimation and Prediction, Clujul Medical, Editura Iuliu Hatieganu, ISSN 1222-2119, Cluj-Napoca, Romania, LXXX(1), p. 125-132, 2007. |
| 35 | Lorentz JÄNTSCHI, Carmen Elena STOENOIU, Sorana-Daniela BOLBOACA, A Formula for Vertex Cuts in b-Trees, Fourth International Conference of Applied Mathematics and Computing, August 12-18, 2007, University of Chemical Technology and Metallurgy Sofia & Technical University of Plovdiv, Short communication, presented on August 15, from 12.10 to 12.20, Plovdiv, Bulgaria, p. 233, 2007. |
| 36 | Lorentz JÄNTSCHI, Sorana-Daniela BOLBOACA, Cristina Maria FURDUI, Characteristic and Counting Polynomials on Modeling Nonane Isomers Properties, Fourth International Conference of Applied Mathematics and Computing, August 12-18, 2007, University of Chemical Technology and Metallurgy Sofia & Technical University of Plovdiv, Invited lecture, presented on August 15, from 12.20 to 13.00, Plovdiv, Bulgaria, p. 234, 2007. |
| 37 | Sorana-Daniela BOLBOACA, Lorentz JÄNTSCHI, Structure-Property Based Model Estimation of Alkanes Boiling Points, Fourth International Conference of Applied Mathematics and Computing, August 12-18, 2007, University of Chemical Technology and Metallurgy Sofia & Technical University of Plovdiv, Invited lecture, presented on August 13, from 18.30 to 19.00, Plovdiv, Bulgaria, p. 48, 2007. |
| 38 | Sorana D. BOLBOACA, Lorentz JÄNTSCHI, Data Mining on Structure-Activity/Property Relationships Models, World Applied Sciences Journal, IDOSI Publications, ISSN 1818-4952, www, Internet & Faisalabad, Pakistan, 2(4), p. 323-332, 2007. |

Metoda MDF și aplicația sa MDF-SAR s-a dovedit un foarte bun instrument pentru designul compușilor chimici. Câteva lucrări din secțiunea de rezultate expun abilitățile sale pe seturile investigate (peste 50).

Activitatea 2. Dobândire competențe complementare prin participări la manifestări științifice/stagii de documentare-cercetare

S-a participat pentru dobândire de competențe complementare la Accelrys Science Forum 2007 desfășurat la Cambridge în perioada 12-13 noiembrie 2007.

Forumul a cuprins două secțiuni de prezentări științifice ale celor mai recente realizări ale companiei americane Accelrys și colaboratorilor acesteia.

Firma multinațională Accelrys este profilată pe producerea de soft specializat și dedicat pentru aplicații biomedicale și farmaceutice (acesta fiind unul din publicurile țintă ale sale) și aplicații chimice și de știința materialelor (acesta fiind cel de-al doilea public țintă al său).

Prima secțiune de prezentări științifice a cuprins cele mai recente realizări și provocări în domeniul aplicațiilor biomedicale și farmaceutice.

A doua secțiune de prezentări științifice a cuprins cele mai recente realizări și provocări în domeniul aplicațiilor chimice și de știința materialelor.

În cadrul forumului au fost cuprinse și două secțiuni *hands-on*, aceasta însemnând instruire pe care reprezentanții companiei au oferit-o participanților cu aplicațiile soft dezvoltate de aceștia.

În continuare sunt discutate prezentările susținute.

Secțiunea de aplicații biomedicale și farmaceutice

■ Gareth WILDEN (Department of medicinal chemistry, AstraZeneca R & D Charnwood, Loughborough, Leicestershire, UK) a susținut prezentarea "GA: PP Using Pipeline Pilot as a Genetic Algorithm" în care după o scurtă prezentare a ceea ce a fost dezvoltat anterior, și anume:

- ÷ Procesul de design al librăriei;
- ÷ Enumerarea completă a librăriei;
- ÷ Obținerea scorului compușilor bazat pe violațiile de tip Lipinski (numai compușii cu cel mai mare scor - cea mai mare abatere - sunt eliminați în acest stadiu
- ÷ Analiza vecinătății de vârfuri versus banca de compuși
- ÷ Filtrarea chimică
- ÷ hERG, AMES, Solubilitate, Metabolit reactiv (filtrare și sistem de atenționare)
- ÷ Gruparea pe categorii bazată pe amprente (alegerea iterativă din categorii până când numărul necesar de compuși este selectat; compușii cu cel mai mic scor sunt selectați primii)

a prezentat problematica actuală și soluțiile care în perspectiva autorului se profilează și anume:

- ÷ Problema ridicată de chimie: *Cum se poate include acoperirea farmacoforă în analiza diversității*
- ÷ Existența și accesibilitatea locală a amprentelor structurale (metoda de clasificare prin excluziune sferică pentru a obține avantajul amprentării structurale de acest tip, cu validare internă corectă
- ÷ Dezvoltarea procedeelelor de amprentare internă a farmacoforilor (în relație directă cu utilizarea aceleiași metode de clasificare ca mai sus)
- ÷ Cum adăugarea a încă unei amprente de farmacofor aduce o acoperire suplimentară

■ Darren FAYNE, Tim JAMES, Andrew KNOX, Giorgio CARTA, David G. LLOYD (Molecular Design Group, School of Biochemistry and Immunology, Trinity College Dublin, Irlanda) au susținut prezentarea *High-throughput virtual screening drug discovery*, în care după o scurtă prezentare a stadiului cunoașterii, și anume:

- ÷ Diagrama de lucru în Drug Design - Faza I - Crearea modelului de receptor
- ÷ Diagrama de lucru în Drug Design - Faza II - Evaluarea protocolului de examinare și evaluare sistematică pentru identificare substanțelor și atributelor nedorite
- ÷ Diagrama de lucru în Drug Design - Faza III - Aplicații ale protocolului de examinare și evaluare sistematică pentru identificare substanțelor și atributelor nedorite
- ÷ Paradigma complexității în Drug Design (spațiul director - dimensiune de ordinul zecilor; spațiul optimizat/virtual - dimensiune de ordinul zecilor de mii; spațiul protocolului de examinare și evaluare sistematică - dimensiune de ordinul milioanele; spațiul detectării chimice - dimensiune de ordinul zecilor de miliarde; spațiul chimiei medicinale - dimensiune de ordinul 10^{60})

prezentarea și-a concentrat atenția asupra problematicii actuale și a soluțiilor propuse:

- ÷ Încrederea în bazele de date pentru ER (adică setul de date Bissantz)
- ÷ Rezultatele obținute din andocare
- ÷ Construcția unui bun set de învățare
- ÷ Efectul SMILES: Conformerii
- ÷ Funcțiile de scor și ce sunt acestea
- ÷ Studiu de caz la îmbogățirea valorilor pentru funcții de scor
- ÷ Soluții ale cercetării și producerea unei noi funcții de scor
- ÷ Crearea unor noi instrumente pentru design molecular - fFLASH (al firmei IBM)
- ÷ Protocoale de fixare rigidă
- ÷ Protocoale în linie de așteptare/procesare - librării și pre-procesare
- ÷ Conformerii, andocare și atribuire de scor

Discuția a continuat cu prezentarea componentelor sistemului în detaliu și anume:

- ÷ Corina (modul ce poate fi executat pe o mașină independentă)
- ÷ Diagrama de rețea și curgerea proceselor
- ÷ Omega (manipulator configurabil și noduri de execuție)
- ÷ Andocare și atribuire de scor - modulul FRED (și legătura acestuia cu date de analiză de raze X stocate în fișiere Brookhaven PDB; andocare ghidată)
- ÷ Analiza și interpretarea rezultatelor
- ÷ Modulul HITS (modulul predictiv)
- ÷ Distribuția lucrului în sistem
- ÷ Concluzii obținute din derularea experimentului prezentat

■ Tim CLARK (Centre of Computational Chemistry, Friedrich-Alexander University, Erlangen, Nürnberg, Germania, Centre for Molecular Design, University of Portsmouth, Portsmouth, UK, și Editor-in-Chief, Journal of Molecular Modeling, Springer, Germania) a susținut prezentarea *Surface-Based in silico Techniques*, în care după o scurtă prezentare a participanților la proiectul prezentat:

- ÷ Centre of Computational Chemistry, Friedrich-Alexander University, Erlangen, Nürnberg, Germania
- ÷ Centre for Molecular Design, University of Portsmouth, Portsmouth, UK
- ÷ Department of Computer Science, University of Aberdeen, Aberdeen, UK
- ÷ Cepos InSilico Ltd, Erlangen, Germania

a metodelor folosite:

- ÷ AM1 (Austin Model 1)
- ÷ Molecular Electrostatics (în legătură directă cu metoda MNDO)
- ÷ B3LYP/6-31G(d)

÷ MP2/6-31G(d)

a trecut la prezentarea de ansamblu:

- ÷ Proprietăți locale (potențial electrostatic molecular, interacții cu probe, formă, proprietăți derivate din orbitali)
- ÷ Descriptori pentru QSAR (eliminarea tuturor descriptorilor specifici pentru atom/substructură; pot descriptorii statistici bazați pe proprietăți locale să performeze la fel de bine?)
- ÷ QSAR convențional (puncte de fierbere pentru N = 5453 molecule - !)
- ÷ Modele integrale pe suprafețe
- ÷ Energii libere de hidratare (în mediu neutru)
- ÷ CypScore
- ÷ Modele bazate pe ParaSurf - proprietăți atomice de suprafață
- ÷ Setul de validare public
- ÷ Setul de validare internă
- ÷ Extensii armonice sferice
- ÷ Descrierea suprafeței
- ÷ Potențialul electrostatic molecular
- ÷ Aliniamentul rapid ParaFit

■ Peter HAYNES (Materials and Physics Department, Imperial College London) a susținut prezentarea *Linear-scaling Density-functional Theory for Life Sciences* al cărui subiect principal a fost realizarea și utilizarea aplicației ONETEP. Codul ONETEP a fost construit pentru a efectua calcule de scară largă înalt eficiente cu aceeași acuratețe ca metodele tradiționale bazate pe primul principiu. Prezentarea a fost orientată către ilustrarea facilităților codului ONETEP care îl fac în special foarte potrivit pentru studiul sistemelor biologice, ilustrate de studiul inhibiției kinazelor dependente de ciclină.

■ C. M. Venkatchalam (Cercetător, Accelrys) a susținut prezentarea *A Rational Approach to Receptor-flexible Docking: Method and Validation Results*. În prezentare s-a susținut că evidența experimentală arată că structurile proteice adoptă o conformație variată când liganzi diferiți sunt legați de acestea. În opinia prezentatorului, metoda de investigare de înaltă rezoluție în general implică o singură structură de receptor cu o largă varietate de liganzi. De asemenea, oricum, flexibilitatea proteinelor poate juca un rol vital în mecanismul de andocare al ligandului. Echipa de cercetare coordonată de prezentator a observat utilizând Discovery Studio (produs soft comercializat de Accelrys) o metodă automată pentru andocarea liganzilor când flexibilitatea ligandului și a lanțului proteic sunt ambele luate în considerare. Rezultatele andocării liganzilor pentru câteva sisteme proteice au fost obținute și investigate. Metoda prezentată are avantajul că poate fi ușor extinsă pentru a considera schimbările ciclice de conformație.

Secțiunea de aplicații de chimie și știința materialelor

■ James ELLIOTT (Materials Modelling Group, Department of Engineering, University of Cambridge) a susținut prezentarea *A Semi-Empirical Molecular Orbital Study of Fullerene-Encapsulated Mo-S/Mo-O Nanoclusters* în care problematica încapsulării de atomi și ioni metalici în structura nanofulerenelor a fost investigată cu ajutorul modelării moleculare. Aglomerările de metale tranziționale nanoscopice sunt folosite pe scară largă în creșterea catalitică a nanotuburilor de carbon (CNTs) sintetizate utilizând procesul de depunere chimică din vapori (CVD). Este cunoscut că aditia oxigenului la calcogenii grei cum este sulfurul deopotrivă produce creșterea ratei de depunere a CNT și promovează depunerea selectivă a straturilor simple și duble de CNTs. Grupul coordonat de prezentator au raportat rezultate de calcul semi-empiric de orbitali moleculari (utilizând Materials Studio și VAMP - produse de firma Accelrys) ale aglomerărilor de Mo-S și Mo-O, care demonstrează schimbările care au loc în structura electronică și sunt cauzate de aditivii folosiți.

■ Andrei GUSEV (Researcher, MatSim GmbH Zürich și Professor, Department of Materials, Swiss Federal Institute of Technology Zürich) a susținut prezentarea *Finite Element Approaches to Mesoscopic Materials Modelling* al cărui subiect l-a constituit utilizarea metodei elementului finit în modelarea nanomaterialelor. Astfel, simulările de dinamică moleculară prezic că transportul gazelor în interiorul nanotuburilor de carbon (CNT) este cu câteva ordine de mărime mai rapidă decât în orice alt material cunoscut. Grupul de cercetare coordonat de prezentator a utilizat tehnologia elementului finit oferită de platforma PLAMYRA/GRIDDER pentru a estima performanța de permeabilitate a membranelor matriciale mixte CNT/polimer. Colectivul de cercetare a demonstrat că astfel de membrane pot într-un mod favorabil să combine performanța de flux ridicat a nanotuburilor cu selectivitatea intrinsecă a matricilor polimerice.

■ Sam FRENCH (Senior Scientist, Johnson Matthey) a susținut prezentarea *Modelling Reactions at the Active Sites of Chiral Ruthenium Catalysts using Density Functional Theory*. Scurta secțiune introductivă în subiectul prezentării a cuprins subiecte de actualitate despre cataliză și tehnologii chirale. Astfel, selectivitatea este un factor cheie de succes pe piața tehnologiilor de cataliști chirali. Înțelegerea proceselor fundamentale care apar când un reactiv interacționează cu un sit simplu și omogen de catalizator, în situl activ și în absența acestuia, este deci esențial pentru designul rațional al noilor cataliști. Catalizatorii de hidrogenare bazați pe rutheniu asimetrici au făcut parte dintr-un proiect de cercetare colaborativă al grupului reprezentat de prezentator [(S)-XyIBINAP-RuH₂-(S,S)-DPEN], primul descoperit de Noyori [Ryoji NOYORI, *Asymmetric catalysis: Science and opportunities* (nobel lecture), *Angewandte Chemie - International Edition*, 41(12), p. 2008-2022, 2002], [Ryoji NOYORI, T. Ohkuma, *Asymmetric catalysis by architectural and functional molecular engineering: Practical chemo- and stereoselective hydrogenation of ketones*, *Angewandte Chemie - International Edition*, 40(1), p. 40-73, 2001] și [Masato KITAMURA, Masaki TSUKAMOTO, Yuhki BESSHO,

Masahiro YOSHIMURA, Uwe KOBBS, Michael WIDHALM, Ryoji NOYORY, Mechanism of asymmetric hydrogenation of α -(acylamino)acrylic esters catalyzed by BINAP-ruthenium(II) diacetate, Journal of the American Chemical Society, 124(23), p. 6649-6667, 2002] este studiat ca patent sau model prototip pentru o serie de catalizatori de hidrogenare eficienți, printre aceștia numărându-se catalizatori bazați pe familiile de liganzi P-Phos, PhanePhos, și ParaPhos [Antonio ZANOTTI-GEROSA, William HEMS, Michelle GROARKE, Fred HANCOCK, Ruthenium-catalysed asymmetric reduction of ketones, Platinum Metals Review, 49(4), p. 158-165, 2005].

■ Asen ASENOV (Device Modelling Group, Department of Electronics and Electrical Engineering, Glasgow University) a susținut prezentarea (care în opinia noastră a fost dintre cele mai valoroase prezentări) *Simulation of Atomic Scale Effects in Nano-CMOS Devices*. Fluctuațiile parametrilor intrinseci asociate cu discretizarea sarcinii și atomicitatea materiei sunt acum unii dintre factorii majori de limitare a scalei, a integrării și a reducerii tensiunii aplicate și puterii consumate. Profesorul Asenov a prezentat cele mai noi descoperiri în simularea fluctuațiilor de variabilitate la scală atomică în componentele nano CMOS utilizând tehnicile de difuzie de drift (DD), Monte Carlo (MC) și transport cuantic (QT). A discutat de asemenea necesitatea unei legături strânse între aceste componente și modelarea materialelor. Secțiunea introductivă a prezentării a cuprins:

- ÷ Continuarea scenariului More Moore
- ÷ Viitoarele MOSFET-uri sunt componentele la scară atomică
- ÷ Discretizarea sarcinii și materiei introduce variabilitate
- ÷ Variabilitatea devine o durere de cap majoră
- ÷ Variabilitatea deja și-a arătat efectele adverse

Prezentarea a continuat cu expunerea surselor de variabilitate:

- ÷ Dopanți întâmplători discreți
- ÷ Linii de separație iregulate
- ÷ Iregularități ale interfețelor
- ÷ Margini formate din grăunțe de polisilicon
- ÷ Morfologia pentru κ înalt

Secțiunea aplicativă a prezentării a cuprins tehnicile de simulare:

- ÷ Corecții ale soluțiilor cuantice pentru electroni și goluri utilizând DG
- ÷ IPF combinate în MOSFET-uri conglomerate
- ÷ Sarcina suplimentară în conglomeratele MOSFET
- ÷ Investigarea impurităților folosind Ab-initio în ansambluri Monte Carlo
- ÷ Simularea IPF necesită simulări statistice tridimensionale
- ÷ Rugozitatea interfeței în simulările bidimensionale NEGF

- ÷ Rugozitatea interfeței în simulările tridimensionale NEGF
- ÷ Dopanți întâmplători în simulările bidimensionale NEGF
- ÷ Dopanți întâmplători în simulările tridimensionale NEGF

Legătura între designul componentelor nano și modelarea materialelor a cuprins următoarele subiecte de interes:

- ÷ Impactul regiunii tranziționale Si/SiO₂ în calculele de străpungere a porții
- ÷ Incluziunea de stive-poartă cu κ înalt
- ÷ Variabilitatea conductibilității la tunelarea porții
- ÷ Structura de bandă în componentele UTB
- ÷ Analiza rugozității în ansambluri Monte-Carlo folosind metoda Ab-initio

Secțiunea de final a prezentării a cuprins cele mai importante concluzii desprinse din provocările cu care se confruntă acum tehnologia nanocomponentelor semiconductoare:

- ÷ Tranzistorii CMOS se apropie de ordinul de mărime atomic
- ÷ Simularea necesită rezoluție la scară atomică
- ÷ Efectele de la scară atomică implică o variabilitate sporită
- ÷ Legătura cu primele principii ale simulării materialelor devin obligatorii

■ Stephen TODD (Product Manager, Accelrys) a susținut prezentarea *Future Developments in Materials Studio 4.2* în conjuncție cu prezentarea de ansamblu a produsului Materials Studio oferit de firma Accelrys, cu marcarea noilor funcționalități introduse cu ajutorul limbajului de scriptare MaterialsScript, prin intermediul noii interfețe Gaussian și Adsorption Locator și în contextul celor mai noi descoperiri ale Nanotechnology Consortium. Prezentarea a început prin expunerea proiectelor de cercetare în care în prezent firma Accelrys este implicată:

- ÷ Materials Visualizer (segregarea volumelor atomice, performanță în utilizarea memoriei)
- ÷ Simulări (editarea câmpurilor de forțe, dinamica moleculară fin divizată, proiectul GULP)
- ÷ Mezoscala (dezvoltarea instrumentelor suport pentru îmbunătățirea instrumentelor la mezoscală, analiză, utilitate)
- ÷ Mecanica cuantică și cataliza: CASTEP
- ÷ Mecanica cuantică și cataliza: DMol
- ÷ Mecanica cuantică și cataliza: VAMP
- ÷ Mecanica cuantică și cataliza: Gaussian® UI
- ÷ Mecanica cuantică și cataliza: adsorbția
- ÷ Mecanica cuantică și cataliza: ONETEP
- ÷ Mecanica cuantică și cataliza: QMERA
- ÷ Nanotechnology Consortium

÷ Discovery Studio

■ Patricia GESTOSO-SOUTO (Senior Scientist, Accelrys) a susținut prezentarea *New Science in Contact Research Services: Case Studies, SIG's and Recent Initiatives* cu accent pe cercetări pentru surse de energie alternativă și modelarea țesutului pielii umane. Prezentarea s-a concentrat pe următoarele subiecte cheie:

÷ Activități specifice ale serviciilor de contractare a cercetării

÷ Studiu de caz: Îmbunătățirea formulărilor pentru industria cosmetică

÷ Actualități pentru grupuri speciale de interes (polimeri pentru celule de energie, polimeri industriali, (noi) polimeri compoziți avansați)

÷ Alte inițiative recente (cercetări de energie alternativă, modelarea țesutului pielii umane)

■ Richard GILBERT (Principal Scientist, e2v Biosensors), autor a 15 patente și 25 lucrări evaluate prin procedura peer-review (incluzând câteva articole în *Nature* și în *Analytical Chemistry and Biochemistry*) a susținut prezentarea *Versatile Biosensor Labels* (în opinia noastră fiind una dintre cele mai valoroase prezentări). În ultimele două decenii biosenzorii și-au mărit impactul în științele vieții și piața de echipamente medicale, și încep a fi utilizați ca instrumente de cercetare primare în multe laboratoare. Biosenzorii sunt componente care utilizează reacțiile chimice și biologice pentru a detecta analiți în probă, în mod tipic monitorizând nivelul unor proteine sau acizi nucleici specifici în lichide biologice cum sunt sângele sau extractele celulare. Sunt o multitudine de tehnologii de detecție biosenzorială, fiecare dintre ele aducând după sine propriile sale beneficii și restricții, și care sunt uzual cerute de dezvoltarea unor sisteme specifice de identificare chimică pentru a face cea mai bună utilizare a tehnologiilor senzoriale. Autorul prezentării a susținut că utilizarea lui Materials Studio (al companiei Accelrys) a permis companiei pe care o reprezintă e2v să dezvolte o chimie identificativă generică care poate fi aplicată la o varietate de tehnici de detecție. De usa singură, interfața de lucru pentru identificare chimică a fost dezvoltată și este compatibilă cu tehnologiile senzoriale dintre cele mai diverse, cum sunt electrochimia, fluorescența, chemiluminescența, electroluminescența și spectroscopia Raman, dar oricum se adresează tuturor principalelor tipuri de detecție din uzul curent.

Activitatea 3. Sumarizare și stocare metode SAR

Următorul tabel prezintă centralizat compușii și activitățile selectate în studiu, așa cum s-au selectat și prevăzut în propunerea de proiect de cercetare:

| Nr | Compuși | Set | Activitate |
|----|----------------------------------|-----|--|
| 1 | derivați carbochinone | 37 | antitumorală |
| 2 | compuși organici | 56 | traversare barieră hemato-encefalică |
| 3 | derivați de sulfonamide | 18 | inhibitori ai anhidrazei carbonice |
| 4 | Taxoizi | 34 | inhibiția dezvoltării celulare |
| 5 | derivați de trifenilacrilonitril | 25 | afinitate relativă de legare receptori de estrogen |

În continuare se redau detaliile seturilor de compuși:

1.1. Setul derivaților de carbochinone

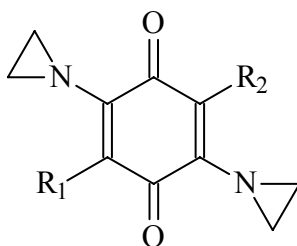
- ÷ Lucrare: “Application of a Self-Organizing Map to Quantitative Structure-Activity Relationship Analysis of Carboquinone and Benzodiazepine”
- ÷ Autori: Junko KAWAKAMI, Kenji HOSHI, Akiko ISHIYAMA, Sayaka MIYAGISHIMA, and Kenichi SATO
- ÷ Referință: Chem. Pharm. Bull., 52(6), p. 751-755, 2004
- ÷ Rezumat: propune utilizarea așa-numitelor hărți auto-organizate pentru analiza relațiilor structură-activitate, incluzând în studiu un set de 37 de derivați de carbochinone
- ÷ Descriere

Hărțile autoorganizate (notate SOM) propuse de Kohonen (Kohonen T., “Self-Organizing Maps,” Springer, Berlin, 2000) par a fi un promițător mod de abordare dincolo de modelul standard al regresiei pentru unele problemele de clasificare care apar în domeniul farmaciei. Autorii lucrării au aplicat acest model la relațiile cantitative structură-activitate ale carbochinonelor și benzodiazepinelor și au arătat utilitatea acestei abordări.

Cele mai multe analize QSAR ce utilizează rețele neuronale au fost obținute prin adoptarea unei rețele neuronale cu învățare supervizată. Din contră, SOM respectă învățarea nesupervizată și inițial nu implică utilizarea datelor țintă dorite. Dacă notăm că o fracție considerabilă a datelor pot să lipsească fără a face comparația similarității imposibilă în SOM dacă numărul de attribute considerate este apreciabil, analiza QSAR utilizând SOM este posibilă ca și o învățare supervizată. Similar cu datele țintă în învățarea supervizată, noi putem să luăm în considerare datele țintă (activitatea observată) ca unul dintre attribute în adiție la alte attribute (descriptorii structurali). Alegerea descriptorilor optimale ca parametri de intrare a fost găsită a fi esențială pentru generarea unui SOM valoros.

Structura și activitatea derivaților carbochinonei

Derivații carbochinonei sunt un grup de compuși având configurația din figura de mai jos, unde R_1 și R_2 sunt radicali. Aceștia au fost sintetizați de Nakao și alții [Nakao H., Arakawa M., Nakamura T., Fukushima M., Chem. Pharm. Bull., 20, 1968-1974 (1972)] cu scopul de a fi folosiți în componența unui medicament pentru tratarea clinică a cancerului.



Următorul tabel conține substituenții R_1 și R_2 , valorile descriptorilor moleculari folosiți și valoarea activității măsurate pentru setul de 37 de carbochinone:

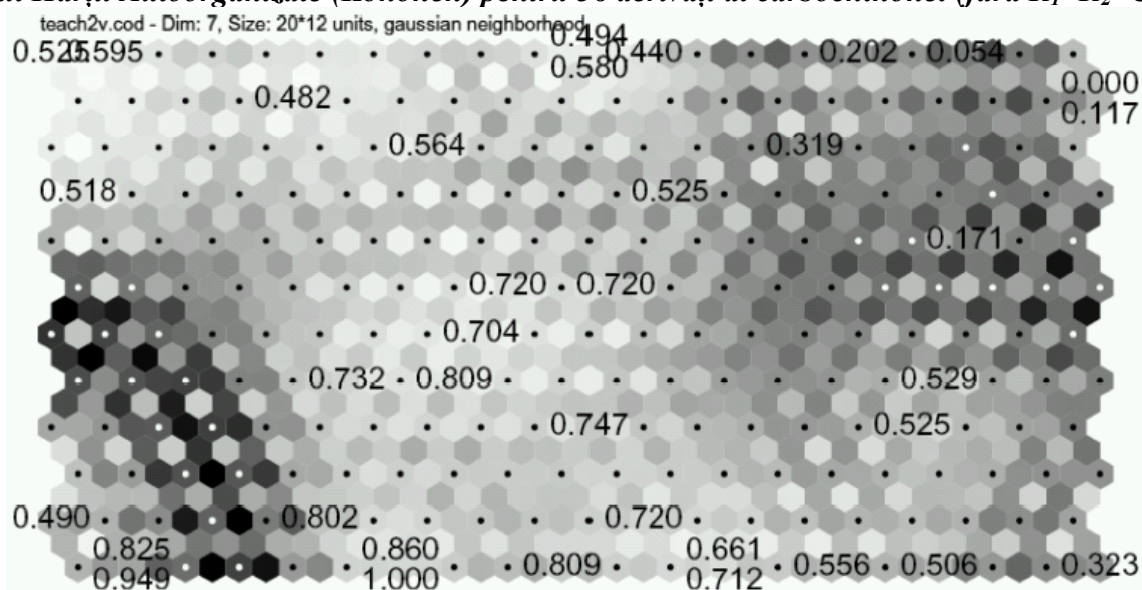
| R_1 | R_2 | $MR_{1,2}$ | $\pi_{1,2}$ | π_2 | MR_1 | F | R | A | NA |
|--------------------------------|--------------------------|------------|-------------|---------|--------|-------|-------|------|-------|
| C_6H_5 | C_6H_5 | 5.08 | 3.92 | 1.96 | 2.54 | 0.16 | -0.16 | 4.33 | 0 |
| CH_3 | $(CH_3)_2C_6H_5$ | 4.5 | 3.66 | 3.16 | 0.57 | -0.08 | -0.26 | 4.47 | 0.054 |
| C_5H_{11} | C_5H_{11} | 4.86 | 5 | 2.5 | 2.43 | -0.08 | -0.26 | 4.63 | 0.117 |
| $CH(CH_3)_2$ | $CH(CH_3)_2$ | 3 | 2.6 | 1.3 | 1.5 | -0.08 | -0.26 | 4.77 | 0.171 |
| CH_3 | $CH_2C_6H_5$ | 3.57 | 2.51 | 2.01 | 0.57 | -0.12 | -0.14 | 4.85 | 0.202 |
| C_3H_7 | C_3H_7 | 3 | 3 | 1.5 | 1.5 | -0.08 | -0.26 | 4.92 | 0.23 |
| CH_3 | $CH_2OC_6H_5$ | 3.79 | 2.16 | 1.66 | 0.57 | -0.04 | -0.13 | 5.15 | 0.319 |
| $R_1=R_2=CH_2CH_2OCON(CH_3)_2$ | | 6.14 | 0.72 | 0.36 | 3.07 | -0.08 | -0.26 | 5.16 | 0.323 |
| C_2H_5 | C_2H_5 | 2.06 | 2 | 1 | 1.03 | -0.08 | -0.26 | 5.46 | 0.44 |
| CH_3 | $CH_2CH_2OCH_3$ | 2.28 | 1.03 | 0.53 | 0.57 | -0.08 | -0.26 | 5.57 | 0.482 |
| OCH_3 | OCH_3 | 1.58 | -0.04 | -0.02 | 0.79 | 0.52 | -1.02 | 5.59 | 0.49 |
| CH_3 | $CH(CH_3)_2$ | 2.07 | 1.8 | 1.3 | 0.57 | -0.08 | -0.26 | 5.6 | 0.494 |
| C_3H_7 | $CH(OCH_3)CH_2OCONH_2$ | 4.24 | 0.98 | -0.52 | 1.5 | -0.04 | -0.13 | 5.63 | 0.506 |
| CH_3 | CH_3 | 1.14 | 1 | 0.5 | 0.57 | -0.08 | -0.26 | 5.66 | 0.518 |
| H | $CH(CH_3)_2$ | 1.6 | 1.3 | 1.3 | 0.1 | -0.04 | -0.13 | 5.68 | 0.525 |
| CH_3 | $CH(OCH_3)C_2H_5$ | 2.75 | 1.53 | 1.03 | 0.57 | -0.04 | -0.13 | 5.68 | 0.525 |
| C_3H_7 | $CH_2CH_2OCONH_2$ | 3.56 | 1.45 | -0.05 | 1.5 | -0.08 | -0.26 | 5.68 | 0.525 |
| $R_1=R_2=CH_2CH_2OCH_3$ | | 3.42 | 1.03 | 0.53 | 1.71 | -0.08 | -0.26 | 5.69 | 0.529 |
| C_2H_5 | $CH(OC_2H_5)CH_2OCONH_2$ | 4.23 | 0.98 | -0.02 | 1.03 | -0.04 | -0.13 | 5.76 | 0.556 |
| CH_3 | $CH_2CH_2OCOCH_3$ | 2.78 | 1.23 | 0.73 | 0.57 | -0.08 | -0.26 | 5.78 | 0.564 |
| CH_3 | $(CH_2)_3$ -dimer | 1.96 | 2 | 1.5 | 0.57 | -0.08 | -0.26 | 5.82 | 0.58 |

| | | | | | | | | | |
|--|---|------|-------|-------|------|-------|-------|------|-------|
| CH ₃ | C ₂ H ₅ | 1.6 | 1.5 | 1 | 0.57 | -0.08 | -0.26 | 5.86 | 0.595 |
| CH ₃ | CH(OCH ₂ CH ₂ OCH ₃) ⁻ | 4.45 | 0.01 | -0.49 | 0.57 | -0.04 | -0.13 | 6.03 | 0.661 |
| CH ₃ | CH ₂ CH(CH ₃)OCONH ₂ | 3.09 | 0.75 | 0.25 | 0.57 | -0.08 | -0.26 | 6.14 | 0.704 |
| C ₂ H ₅ | CH(OCH ₃)CH ₂ OCONH ₂ | 3.77 | 0.48 | -0.52 | 1.03 | -0.04 | -0.13 | 6.16 | 0.712 |
| CH ₃ | CH(C ₂ H ₅)CH ₂ OCONH ₂ | 3.55 | 1.25 | 0.75 | 0.57 | -0.08 | -0.26 | 6.18 | 0.72 |
| CH ₃ | CH(OC ₂ H ₅)CH ₂ OCONH ₂ | 3.77 | 0.48 | -0.02 | 0.57 | -0.04 | -0.13 | 6.18 | 0.72 |
| CH ₃ | (CH ₂) ₃ OCONH ₂ | 3.09 | 0.95 | 0.45 | 0.57 | -0.08 | -0.26 | 6.18 | 0.72 |
| CH ₃ | (CH ₂) ₂ OCONH ₂ | 2.63 | 0.45 | -0.05 | 0.57 | -0.08 | -0.26 | 6.21 | 0.732 |
| C ₂ H ₅ | (CH ₂) ₂ OCONH ₂ | 3.09 | 0.95 | -0.05 | 1.03 | -0.08 | -0.26 | 6.25 | 0.747 |
| CH ₃ | CH ₂ CH ₂ OH | 1.78 | 0.34 | -0.16 | 0.57 | -0.08 | -0.26 | 6.39 | 0.802 |
| CH ₃ | CH(CH ₃)CH ₂ OCONH ₂ | 3.09 | 0.75 | 0.25 | 0.57 | -0.08 | -0.26 | 6.41 | 0.809 |
| CH ₃ | CH(OCH ₃)CH ₂ OCONH ₂ | 3.31 | -0.02 | -0.52 | 0.57 | -0.04 | -0.13 | 6.41 | 0.809 |
| H | N(CH ₂) ₂ | 1.66 | 0.18 | 0.18 | 0.1 | 0.1 | -0.92 | 6.45 | 0.825 |
| R ₁ =R ₂ =CH ₂ CH ₂ OH | | 2.42 | -0.32 | -0.16 | 1.21 | -0.08 | -0.26 | 6.54 | 0.86 |
| CH ₃ | N(CH ₂) ₂ | 2.13 | 0.68 | 0.18 | 0.57 | 0.06 | -1.05 | 6.77 | 0.949 |
| CH ₃ | CH(OCH ₃)CH ₂ OH | 2.47 | -0.13 | -0.63 | 0.57 | -0.04 | -0.13 | 6.9 | 1 |

Parametrii fizicochimici folosiți conform autorilor lucrării au următoarele semnificații:

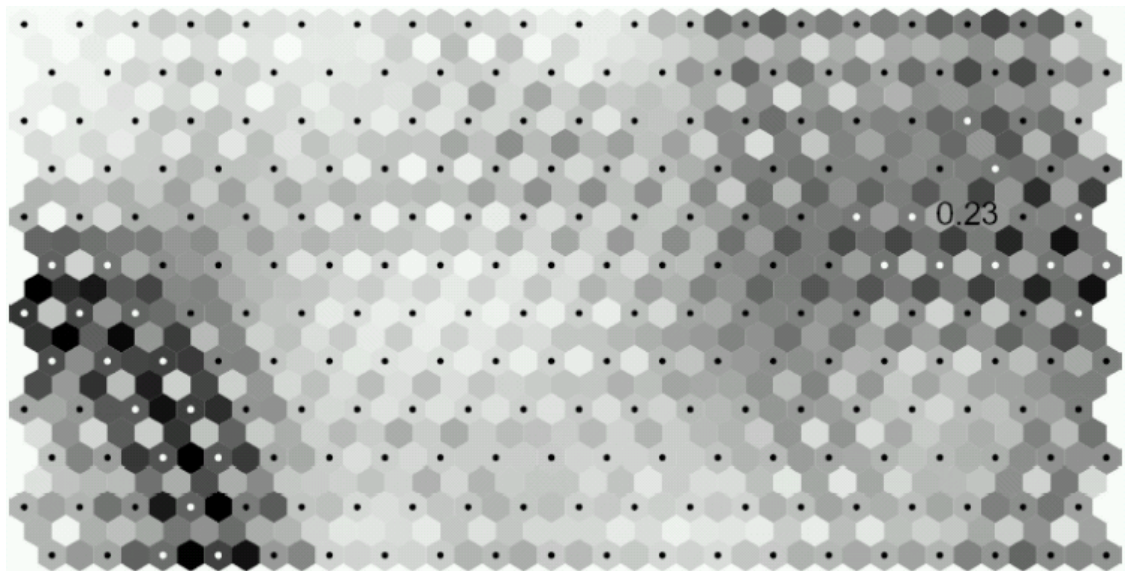
- ÷ MR - refractivitate molară; π - hidrofobicitate; F, R - constante de substituent; MR_{1,2} și $\pi_{1,2}$ - estimatori ai efectului steric al R₁ și R₂ și hidrofobicitatea totală;
- ÷ Utilizând concentrația C care este minimul dozei efective (MED) per kilogram de șoarec, activitatea biologică log(1/C) este codificată cu "A" în tabelul de mai sus.

Modelul Hărții Autoorganizate (Kohonen) pentru 36 derivați ai carbochinonei (fără R₁=R₂=C₃H₇)



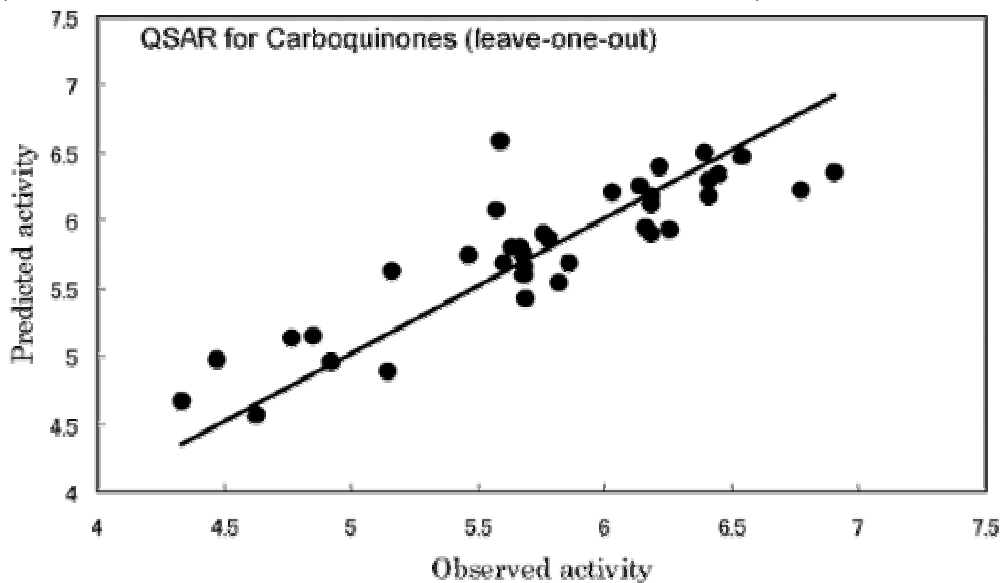
(valorile normalizate ale activității biologice sunt înscrise pentru fiecare neuron câștigător)

Harta Autoorganizată a lui Kohonen derivatul R₁=R₂=C₃H₇ (omis) și neuronul câștigător fixat



(activitatea biologică pentru $R_1=R_2=C_3H_7$ este obținută din valoarea componentei 7-dimensionale a cărții care listează simbolurile vectorului și semnificațiile lor)

Rezultate obținute pentru legătura între structura și activitatea derivaților carbochinonei



(valori prezise din leave-one-out cross-validation - aici cu semnificația că compusul $R_1=R_2=C_3H_7$ nu a făcut parte din setul de învățare al rețelei neuronale - versus valorile observate pentru carbochinone)

Valoarea statisticii r pentru datele prezentate în figura de mai sus este $r = 0.874$, care conform autorilor este aproape egală cu valoarea obținută utilizând BPNN [Ichikawa H., "Kaisougata Nyurarunettowaku," Kyouritsu Co., Tokyo, 1993] sau rețelele neuronale Bayesiene regularizate [Sato K., Nakagawa J., Matuzaki H., J. Tohoku Pharmaceutical University, 44, 187-193 (1997)].

1.2. Setul de compuși organici cu activitate de traversare barieră hemato-encefalică

÷ Lucrare: “Predicting Blood–Brain Barrier Partitioning of Organic Molecules Using Membrane-Interaction QSAR Analysis”

÷ Autori: Manisha IYER, Rama MISHRA, Yi HAN, A. J. HOPFINGER

÷ Referință: *Pharmaceutical Research*, 19(11), p. 1611-1621, 2002

÷ Rezumat: propune utilizarea așa-numitelor relații structură-activitate membrană-interacție pentru partiționarea compușilor organici la bariera sânge-creier, incluzând în studiu un set de 56 de compuși organici pentru care coeficienții de partiție a barierei sânge-creier a fost măsurată. Un set de alți 7 compuși a fost folosit pentru validarea modelelor.

÷ Descriere

Scop. Relații cantitative structură-activitate (QSAR) membrană-interacție (MI-QSAR) au fost folosite pentru a dezvolta modele predictive ale partiționării barierei sânge-creier a compușilor organici prin, în parte, simularea interacției unui compus organic cu regiunile bogate în fosfolipide ale membranelor celulare.

Metodă. Un set de 56 compuși diverși structural ai căror coeficienți de partiție a barierei sânge-creier au fost mășurați au fost folosiți pentru a construi modelele MI-QSAR. Simulări de dinamică moleculară au fost folosite pentru a determina interacțiunea explicită a fiecărui compus testat (solut) cu un model DMPC al modelului de membrană mono-strat. Un set adițional de descriptori de solut intramoleculari au fost calculați și considerați în baza de încercări a descriptorilor pentru construirea modelelor MI-QSAR. Modelele QSAR au fost optimizate utilizând regresia liniară multidimensională și un algoritm genetic. Un set de 7 compuși a fost evaluat folosind modelele MI-QSAR ca parte a procesului de validare.

Rezultate. Modelele MI-QSAR semnificative ($R^2=0.845$, $Q^2=0.795$) ale procesului de partiționare sânge-creier au fost construite. Partiționarea barierei sânge-creier a fost găsită depinzând de suprafața polară, coeficientul de partiție octanol/apă, și flexibilitatea conformațională a compușilor și de asemenea de tăria legăturii lor la modelul membranei biologice. Coeficienții de partiție a barierei sânge-creier au fost preziși cu aceeași acuratețe cu a compușilor din setul de învățare.

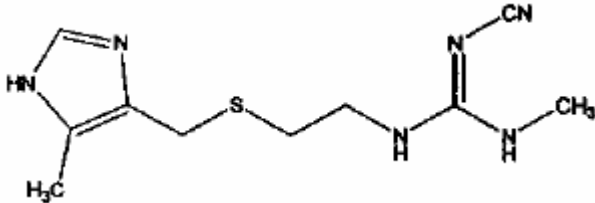
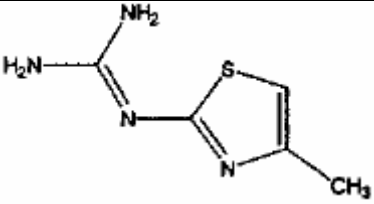
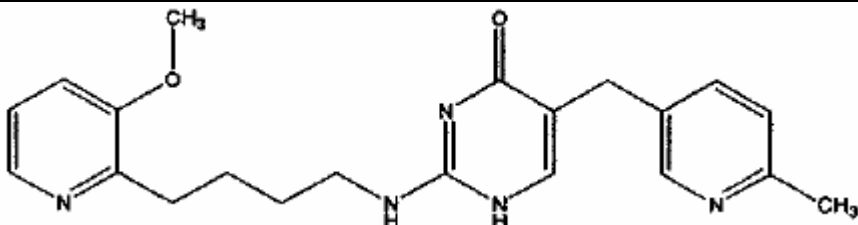
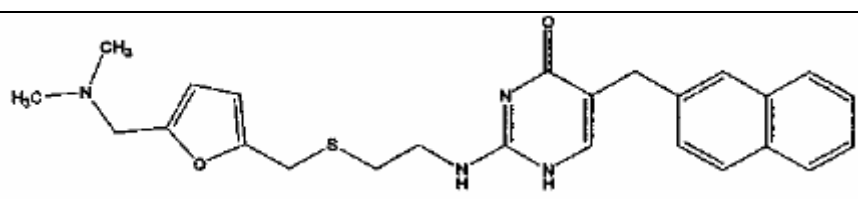
Concluzie. Modelele MI-QSAR indică că procesul de partiționare a barierei sânge-creier poate fi reliabil descris pentru molecule structural diverse furnizând interacțiuni ale moleculei cu regiunile bogate în fosfolipide ale membranei celulare, acestea fiind explicit considerate.

Structura și activitatea derivaților carbochinonei

Abraham și coautorii [M. H. Abraham, H. S. Chadha, and R. C. Mitchell. Hydrogen bonding. 36. Determination of blood-brain barrier distribution using octanol-water partition coefficients. *Drug Des. Discov.* 13:123–131 (1995).] au raportat un studiu BBB (blood-brain barrier - bariera sânge-creier)

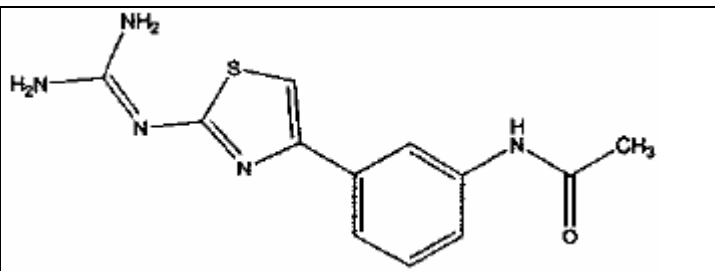
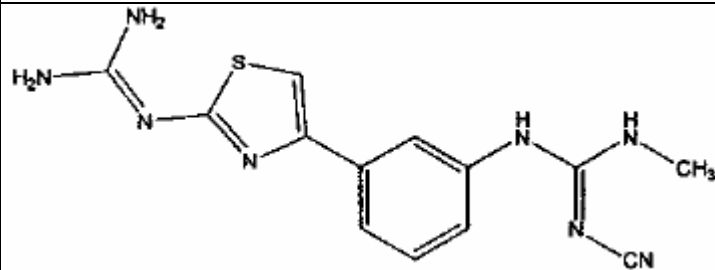
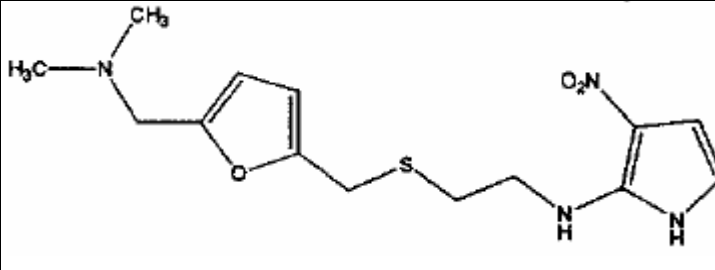
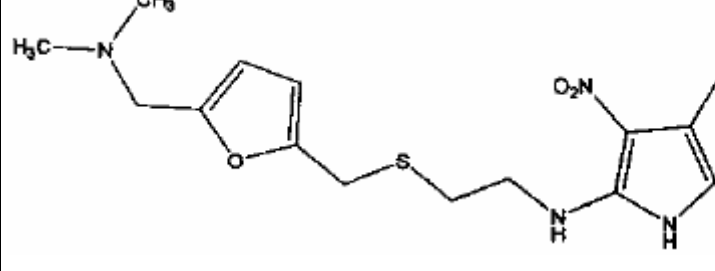
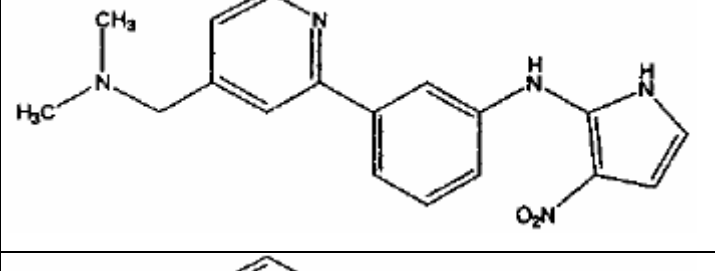
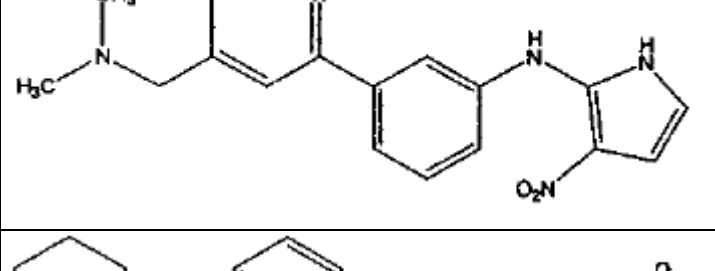
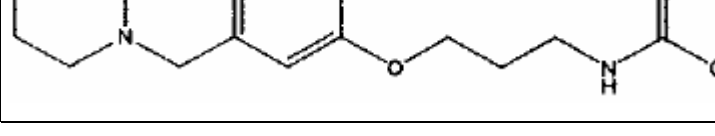
bazat pe un set de învățare de 57 molecule diverse structural și chimic. Acest set popular a fost folosit într-un număr de alte studii raportate de penetrație BBB [R. Liu, H. Sun, and S.-S. So. Development of quantitative structure-property relationship models for early ADME evaluation in drug discovery. 2. Blood-brain barrier penetration. J. Chem. Inf. Comput. Sci. 41:1623–1632 (2001)] și [D. E. Clark. Rapid calculation of polar molecular surface and its application to the prediction of transport phenomena. 2. Prediction of blood-brain barrier penetration. J. Pharm. Sci. 88:815–821 (1999)].

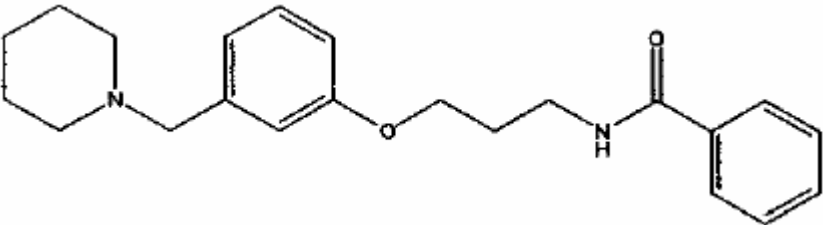
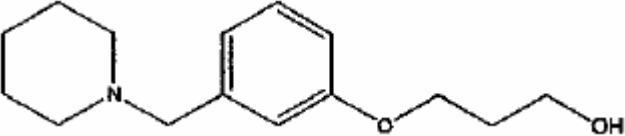
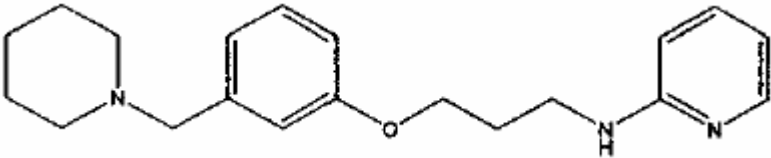
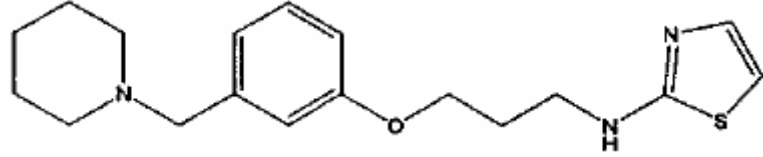
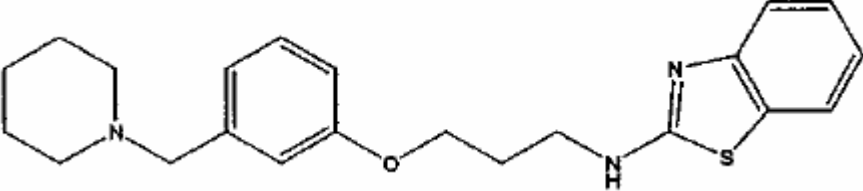
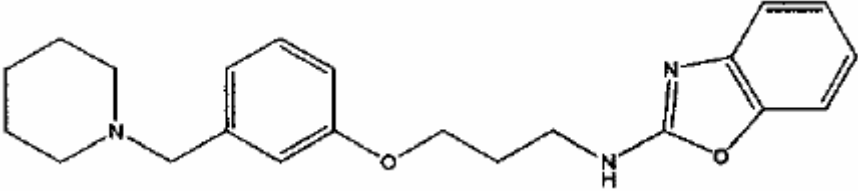
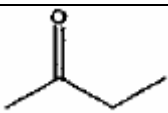
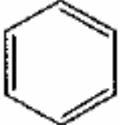
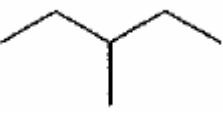
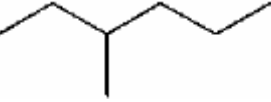
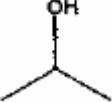
Autorii prezentei lucrări au omis o moleculă (N_2) din setul de date inițial din lipsa unui număr suficient de proprietăți calculate și au utilizat pe cele rămase (56) ca set de învățare pentru studiul BBB MI-QSAR. Acest set de test are domeniul de mase moleculare variind de la 16.03 până la 448.58 unități atomice de masă, iar concentrațiile în sânge și creier au fost măsurate în unități de $\mu\text{M}/\text{mL}$, și acestea fiind variații de încărcare netă la $\text{pH}=7.4$. Prima parte a tabelului de mai jos conține structurile chimice și valorile $\log(\text{BB})$ ale compușilor din setul de învățare. Cei 7 compuși din setul test (în a doua parte a tabelului) au fost selectați din literatură ca aceia la care valorile $\log(\text{B})$ au fost determinate prin același protocol ca la setul de învățare.

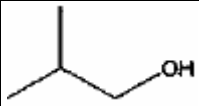


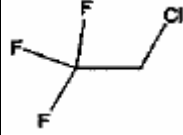
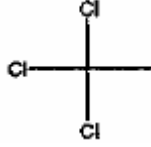
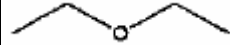
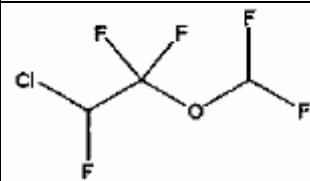

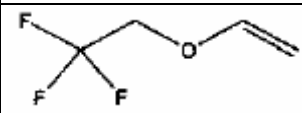
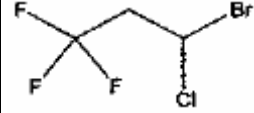


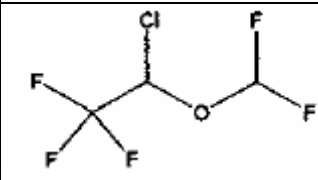
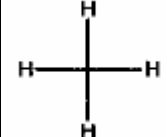
| Nr | Structura | $\log(\text{BB})$ |
|-------------------|--|-------------------|
| Setul de învățare | | |
| 1 |  | -1.42 |
| 2 |  | -0.04 |
| 3 |  | -2.00 |
| 4 |  | -1.30 |

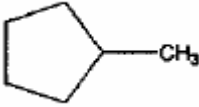


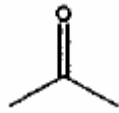
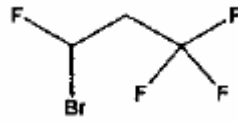
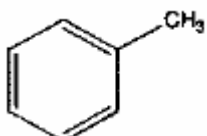
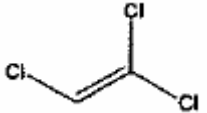
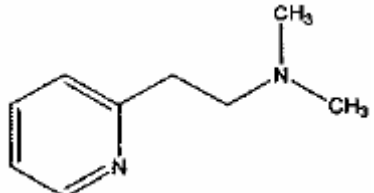
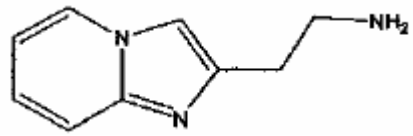
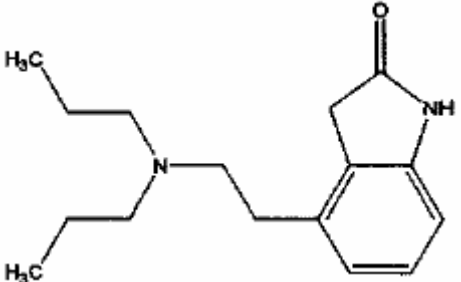
| | | |
|----|--|-------|
| 5 | | -1.06 |
| 6 | | 0.11 |
| 7 | | 0.49 |
| 8 | | 0.83 |
| 9 | | -1.23 |
| 10 | | -0.82 |

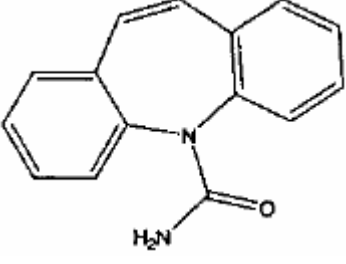
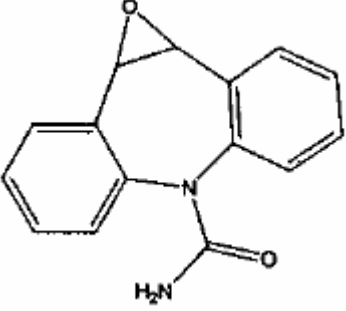
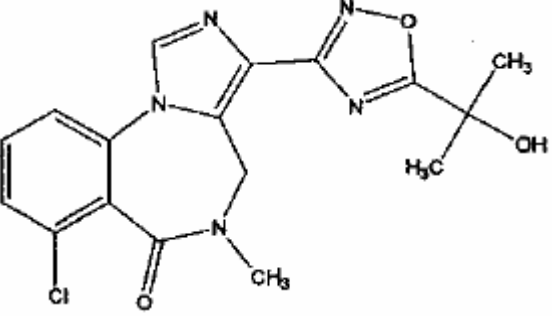
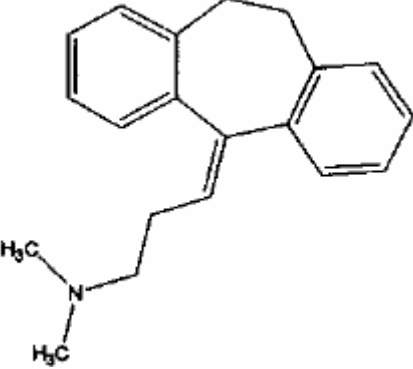
| | | |
|----|--|-------|
| 11 | | -1.17 |
| 12 | | -2.15 |
| 13 | | -0.67 |
| 14 | | -0.66 |
| 15 | | -0.12 |
| 16 | | -0.18 |
| 17 | | -1.15 |

| | | |
|----|--|-------|
| 18 |  | -1.57 |
| 19 |  | -1.54 |
| 20 |  | -1.12 |
| 21 |  | -0.73 |
| 22 |  | -0.27 |
| 23 |  | -0.28 |
| 24 |  | -0.46 |

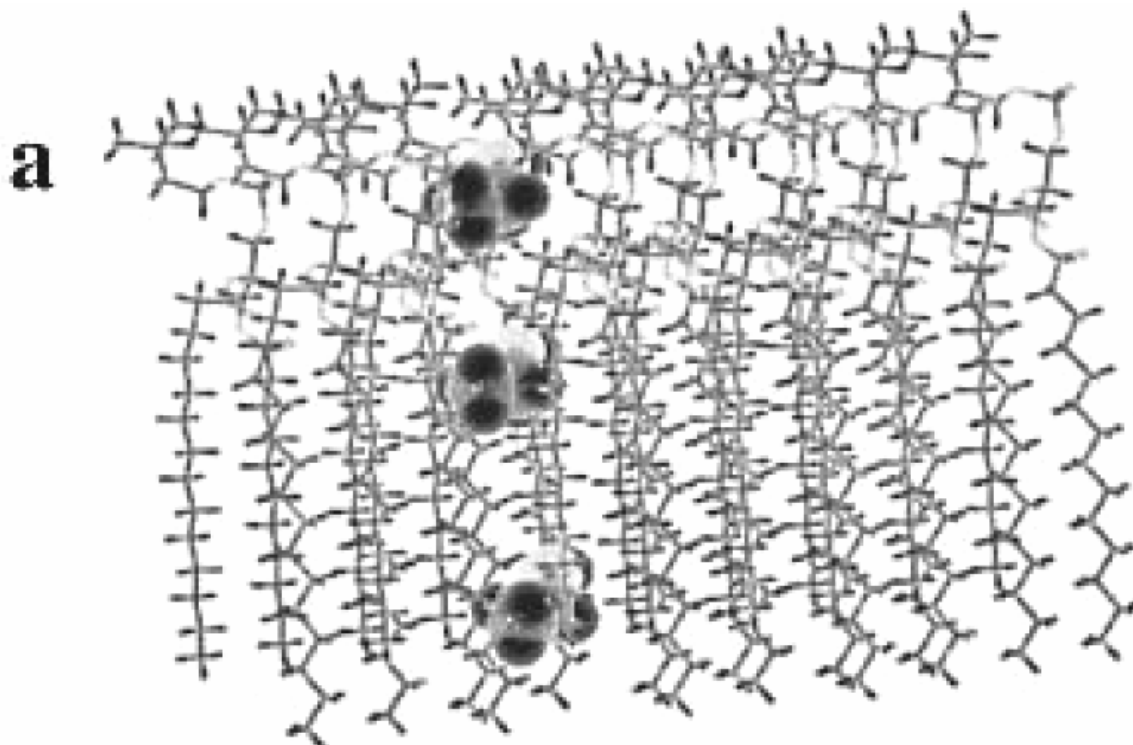
| | | |
|----|--|-------|
| 25 |  | -0.24 |
| 26 |  | -0.02 |
| 27 |  | 0.69 |
| 28 |  | 0.44 |
| 29 |  | 0.14 |
| 30 |  | 0.22 |
| 31 |  | -0.08 |
| 32 |  | 0.37 |
| 33 |  | 1.01 |
| 34 |  | 0.90 |
| 35 |  | -0.15 |

| | | |
|----|---|-------|
| 36 |  | -0.17 |
| 37 |  | 0.97 |
| 38 |  | 1.04 |
| 39 |  | 0.08 |
| 40 |  | 0.40 |
| 41 |  | 0.00 |
| 42 |  | 0.24 |
| 43 |  | -0.16 |
| 44 |  | 0.13 |
| 45 |  | 0.35 |
| 46 |  | 0.81 |
| 47 |  | 0.80 |
| 48 |  | 0.42 |
| 49 |  | 0.04 |

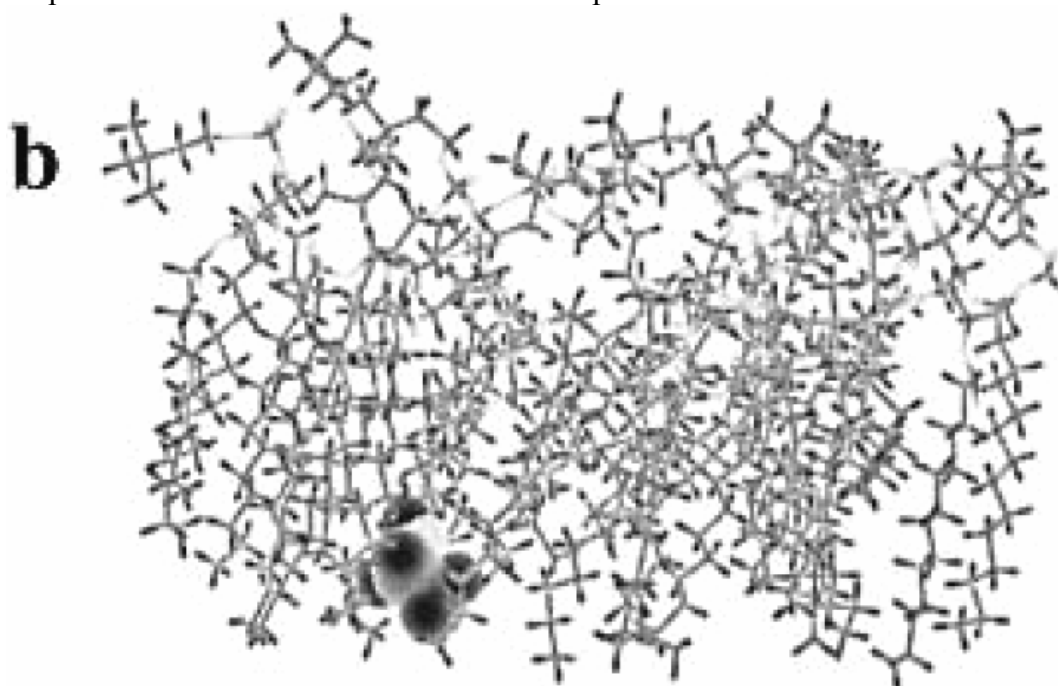
| | | |
|------------------|---|-------|
| 50 |  | 0.93 |
| 51 |  | 0.76 |
| 52 |  | -0.16 |
| 53 |  | -0.15 |
| 54 |  | 0.27 |
| 55 |  | 0.37 |
| 56 |  | 0.34 |
| Setul de testare | | |
| T1 |  | -0.06 |
| T2 |  | -1.40 |
| T3 |  | 0.25 |

| | | |
|----|--|-------|
| T4 |  <p>The structure shows a tricyclic system consisting of two benzene rings fused to a seven-membered ring. A nitrogen atom is part of the seven-membered ring and is also bonded to a carbonyl group (C=O) which has an amino group (H₂N) attached. The nitrogen atom is also bonded to the two benzene rings.</p> | 0.00 |
| T5 |  <p>The structure is similar to T4, but the seven-membered ring contains an epoxide ring (a three-membered ring with one oxygen atom) fused to the benzene rings.</p> | -0.34 |
| T6 |  <p>The structure is a complex heterocyclic molecule. It features a central seven-membered ring with a carbonyl group (C=O) and a methyl group (CH₃) attached to one of the nitrogens. This ring is fused to a benzene ring with a chlorine atom (Cl) at the 3-position. Another nitrogen atom is part of the system, connected to a five-membered ring containing two nitrogens and an oxygen atom. This five-membered ring is further substituted with a methyl group (CH₃), a hydroxyl group (OH), and another methyl group (H₃C).</p> | -1.34 |
| T7 |  <p>The structure shows a tricyclic system with two benzene rings fused to a seven-membered ring. A long chain is attached to the seven-membered ring, ending in a dimethylamino group (N(CH₃)₂).</p> | 0.85 |

Modelul DMPC al modelului de membrană mono-strat



(vedere dintr-o parte a moleculei de etanol inserată în 3 poziții diferite în modelul MDPC mono-strat înainte de pornirea fiecăreia dintre cele 3 DMS corespunzătoare utilizate în modelarea MI-QSAR)



(geometria de cea mai mică energie a complexului DMPC-etanol în DMS)

Descriptori generali intramolecular folosiți ca parte a trialului MI-QSAR

| Descriptor | Semnificație |
|------------|--|
| HOMO | Energia celui mai înalt orbital molecular ocupat |
| LUMO | Energia celui mai jos orbital molecular ocupat |

| | |
|----------------|--|
| Dp | Moment de dipol |
| Vm | Volum molecular |
| SA | Aria suprafeței moleculare |
| Ds | Densitate |
| MW | Masa moleculară |
| MR | Refractivitatea moleculară |
| N(hba) | Numărul de acceptori de legături de hidrogen |
| N(hbd) | Numărul de donori de legături de hidrogen |
| N(B) | Numărul de legături care se pot roti |
| JSSA (X) | Descriptorii de suprafață Jurs-Stanton |
| PSA | Aria suprafeței polare |
| Chi-N, Kappa-M | Descriptori topologici Kier & Hall |
| Rg | Raza de rotație |
| Pm | Momentul principal de inerție |
| Se | Entropia conformațională |
| Q(I) | Densități ale sarcinilor atomice parțiale |

Descriptorii ai interacție intermoleculare folosiți ca parte a trialului MI-QSAR

| Descriptor | Semnificație |
|-----------------------------------|--|
| Descriptorii membrană-solut | |
| $\langle E(\text{total}) \rangle$ | Media energiei potențiale totale a complexului solut-membrană |
| $E_{\text{INTER}}(\text{total})$ | Energia totală a interacției intermoleculare între solut și membrană la energia potențială a sistemului minimă |
| $E_{\text{TT}}(Z)$ | $Z = 1,4$ -nonbonded, general Van der Waal, electrostatic, hydrogen bonding, stretching, bending, torsion and combinations thereof energies of the membrane-solute complex at the total system minimum potential energy. |
| $\Delta E_{\text{TT}}(Z)$ | Change in the $Z = 1,4$ -nonbonded, general Van der Waal, electrostatic, hydrogen bonding, stretching, bending, torsion and combinations thereof at the total [solute and membrane model] intermolecular system minimum potential energy |
| $E_{\text{MS}}(Z)$ | $Z =$ Intermolecular Van der Waal, electrostatic, hydrogen bonding interaction and combinations thereof energies between the solute and the membrane at the total system minimum potential energy |
| $E_{\text{SS}}(Z)$ | $Z =$ Intramolecular $1,4$ -nonbonded, general van der Waal, electrostatic, hydrogen bonding, stretching, bending, |

| | |
|---|--|
| | torsion and combinations thereof energies of the solute within the membrane at total system minimum potential energy |
| $\Delta E_{ss}(Z)$ | Change in the Z = intramolecular 1,4-nonbonded, general van der Waal, electrostatic, hydrogen bonding, stretching, bending, torsion and combinations thereof energies of the solute due to its uptake to the total intermolecular system minimum potential energy. |
| Descriptori ai solvatului pentru dizolvare și solvatare | |
| F(H ₂ O) | Energia liberă la dizolvarea în apă |
| F(OCT) | Energia liberă la dizolvarea în 1-octanol |
| ClogP | Coeficientul de partiție 1-octanol/apă |
| E(coh) | Energia de împachetare coezivă a moleculelor de solut |
| TM | Temperatura ipotetică de tranziție cristal-topitură a solutului |
| TG | Temperatura ipotetică de tranziție la gaz a solutului |

Rezultate obținute pentru legătura între structura și activitatea derivaților carbochinonei

| Nr | Ecuatie (log(BB)=) | Statistică |
|----|---|--|
| 1 | 0.543 - 0.0161PSA | n = 56 R ² = 0.675 Q ² = 0.647 |
| 2 | 0.133 - 0.0153PSA + 0.1522ClogP | n = 56 R ² = 0.744 Q ² = 0.713 |
| 3 | 0.122 - 0.0199PSA + 0.1703ClogP - 0.0049EMS(chg + hbd) | n = 56 R ² = 0.797 Q ² = 0.759 |
| 4 | 0.124 - 0.0226PSA + 0.1613ClogP - 0.0066 EMS(chg + hbd) + 0.0338ESS(tor) | n = 56 R ² = 0.825 Q ² = 0.784 |
| 5 | 0.0156 - 0.0231PSA + 0.1591ClogP - 0.0071 EMS(chg + hbd) + 0.0346ESS(tor) + 0.0075 $\Delta E_{TT}(1-4)$ | n = 56 R ² = 0.845 Q ² = 0.795 |
| 6 | -0.015 - 0.0235PSA + 0.1673ClogP - 0.0076 EMS(chg + hbd) + 0.0388ESS(tor) + 0.01 $\Delta E_{TT}(1-4)$ - 0.0037 $\Delta E_{TT}(\text{stre} + \text{bend})$ | n = 56 R ² = 0.855 Q ² = 0.792 |

Valorile celor 6 descriptori (găsiți a fi semnificativi, tabelul de mai sus) pentru moleculele incluse în set sunt redate în tabelul de mai jos:

| Molecula | PSA (\AA^2) | ClogP | $E_{MS}(\text{chg+hbd})$ | $E_{SS}(\text{tor})$ | $\Delta E_{TT}(1-4)$ | $\Delta E_{TT}(\text{stre+bend})$ |
|-------------------|------------------------|-------|--------------------------|----------------------|----------------------|-----------------------------------|
| Setul de învățare | | | | | | |
| 1 | 92.100 | 0.351 | -52.460 | 3.840 | 20.080 | -5.820 |
| 2 | 78.900 | 0.952 | -206.980 | 3.620 | 2.530 | 10.300 |
| 3 | 94.000 | 2.297 | -65.450 | 4.850 | -4.020 | 37.410 |
| 4 | 73.500 | 4.046 | -10.290 | -3.450 | 20.840 | -19.230 |
| 5 | 87.000 | 1.874 | -36.240 | 6.00 | 16.920 | 0.500 |
| 6 | 39.000 | 0.743 | -7.530 | 5.050 | 56.590 | -6.370 |
| 7 | 26.800 | 2.787 | -7.140 | -1.280 | 12.920 | 16.900 |
| 8 | 6.000 | 4.413 | -16.520 | -2.290 | -4.750 | 26.790 |
| 9 | 84.500 | 1.327 | -10.070 | 5.810 | -18.390 | 32.430 |
| 10 | 139.200 | 0.844 | -164.150 | 6.930 | 6.380 | 6.580 |
| 11 | 88.800 | 0.911 | -99.350 | -1.410 | 39.480 | 40.950 |
| 12 | 73.500 | 2.282 | -4.110 | 1.540 | -0.160 | 34.870 |
| 13 | 83.900 | 2.747 | -6.270 | 9.100 | 26.360 | -31.110 |
| 14 | 84.000 | 1.800 | -62.080 | 10.190 | 12.680 | 16.780 |
| 15 | 78.000 | 3.637 | -79.530 | 2.900 | 15.510 | 18.940 |
| 16 | 76.600 | 2.781 | -188.830 | -5.750 | 19.070 | 37.440 |
| 17 | 104.400 | 1.784 | -166.100 | -1.350 | 2.230 | 19.350 |
| 18 | 108.800 | 1.977 | -139.010 | 2.380 | -1.010 | 6.650 |
| 19 | 135.800 | 1.880 | -126.240 | -2.000 | 38.660 | 0.130 |
| 20 | 85.500 | 2.287 | -74.040 | 11.220 | 1.250 | 66.130 |
| 21 | 79.500 | 4.124 | -18.690 | 6.430 | 17.680 | -34.110 |
| 22 | 82.700 | 3.849 | -80.630 | 7.260 | 46.080 | 29.600 |
| 23 | 85.700 | 3.234 | -13.330 | 6.880 | 36.310 | -42.070 |
| 24 | 47.900 | 2.065 | 2.090 | 4.870 | 4.230 | -16.610 |
| 25 | 45.200 | 4.004 | -9.810 | 4.520 | 23.790 | 35.130 |
| 26 | 38.500 | 2.379 | 0.330 | 5.260 | 15.370 | 32.910 |
| 27 | 39.100 | 4.259 | -7.100 | 6.650 | 14.120 | 21.480 |
| 28 | 40.000 | 4.165 | -27.980 | 7.100 | 21.880 | 9.480 |
| 29 | 39.200 | 5.759 | -10.460 | 4.420 | 39.480 | 23.560 |
| 30 | 54.900 | 5.029 | -33.370 | 1.070 | 17.080 | -18.340 |

| | | | | | | |
|------------|--------|--------|---------|--------|--------|---------|
| 31 | 22.700 | 0.834 | -0.690 | -6.000 | 13.120 | 26.320 |
| 32 | 0.000 | 2.142 | -0.660 | -2.900 | 25.780 | 15.340 |
| 33 | 0.000 | 3.738 | -0.310 | 1.760 | 8.860 | 26.470 |
| 34 | 0.000 | 4.267 | -1.060 | 2.180 | 8.070 | 17.530 |
| 35 | 23.400 | 0.074 | -24.340 | 1.410 | 0.030 | 32.550 |
| 36 | 22.600 | 0.693 | -35.690 | 0.570 | 12.320 | 22.260 |
| 37 | 0.000 | 3.738 | 2.020 | 1.720 | 5.200 | 9.250 |
| 38 | 0.000 | 3.608 | 3.450 | 1.130 | 18.720 | 24.630 |
| 39 | 0.000 | 1.714 | 1.200 | 0.040 | 11.950 | -25.070 |
| 40 | 0.000 | 2.481 | -0.720 | 0.540 | 16.730 | -45.870 |
| 41 | 11.300 | 0.870 | 0.780 | 1.310 | 23.950 | 13.420 |
| 42 | 11.600 | 2.459 | -4.180 | 1.740 | 25.280 | 44.850 |
| 43 | 24.400 | -0.235 | 1.760 | 0.180 | 34.160 | -46.950 |
| 44 | 10.700 | 1.765 | -4.120 | 2.310 | 21.400 | 21.470 |
| 45 | 0.000 | 2.447 | -0.100 | 1.130 | -3.850 | 32.860 |
| 46 | 0.000 | 4.397 | -4.160 | 2.650 | -2.590 | 25.230 |
| 47 | 0.000 | 3.868 | 0.750 | 2.040 | 19.460 | 31.890 |
| 48 | 11.000 | 2.999 | 6.400 | 0.190 | 26.000 | -28.970 |
| 49 | 0.000 | 1.103 | 0.440 | 0.000 | 16.290 | 18.610 |
| 50 | 0.000 | 3.314 | -1.550 | 6.420 | 24.380 | -1.590 |
| 51 | 0.000 | 3.339 | -3.500 | 0.180 | 17.050 | 6.570 |
| 52 | 24.400 | 0.294 | -27.810 | 0.680 | 13.190 | -18.180 |
| 53 | 22.700 | 0.305 | 1.860 | -0.390 | 9.080 | -6.130 |
| 54 | 0.000 | 2.007 | -4.510 | 0.380 | 22.980 | 10.620 |
| 55 | 0.000 | 2.641 | -4.110 | -4.480 | 4.040 | 16.020 |
| 56 | 0.000 | 2.627 | 1.240 | 0.940 | 24.370 | 29.080 |
| Setul test | | | | | | |
| T1 | 18.8 | -0.338 | -2.91 | -1.5 | 11.4 | 19.51 |
| T2 | 46.7 | 0.11 | -3.63 | 9.13 | 7.37 | 20.07 |
| T3 | 44.1 | 1.999 | 44.1 | 7.27 | -14 | 16.83 |
| T4 | 46.7 | 1.98 | -29.3 | 9.85 | 8.96 | 36.58 |
| T5 | 62.7 | 0.26 | -16.03 | 13.57 | 7.43 | 2.87 |
| T6 | 98.5 | 0.38 | 98.5 | 7.69 | -4.07 | 72.84 |
| T7 | 5.4 | 4.641 | -6.78 | 9.07 | -2.46 | 22.28 |

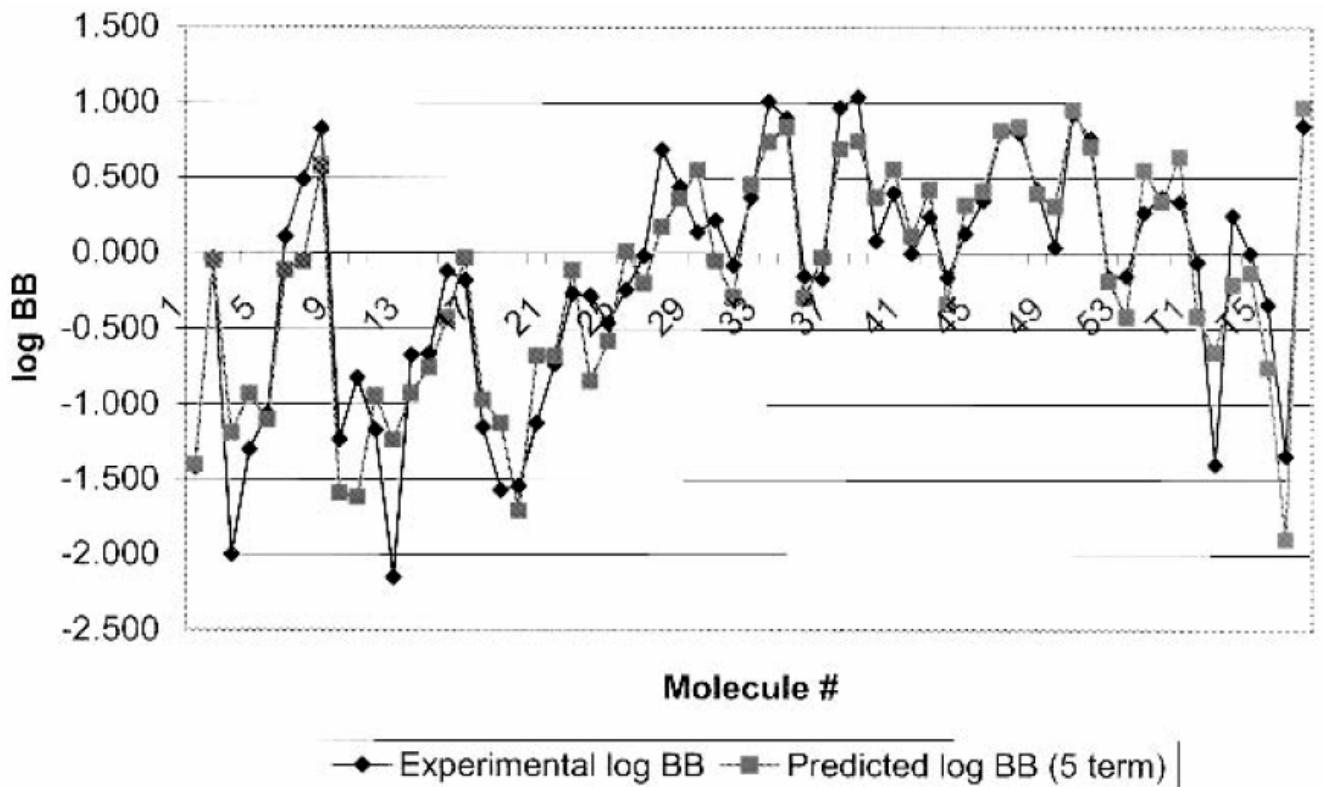
Notă: toate energiile sunt în kcal/mol

Tabelul următor conține valorile prezise de modelele MI-QSAR cu 3-6 termeni.

| Molecula | Log(BB) | model 3 termeni | model 4 termeni | model 5 termeni | model 6 termeni |
|-------------------|---------|-----------------|-----------------|-----------------|-----------------|
| Setul de învățare | | | | | |
| 1 | -1.420 | -1.390 | -1.419 | -1.401 | -1.349 |
| 2 | -0.040 | -0.261 | -0.009 | -0.048 | -0.008 |
| 3 | -2.000 | -1.032 | -1.024 | -1.190 | -1.335 |
| 4 | -1.300 | -0.599 | -0.922 | -0.928 | -0.839 |
| 5 | -1.060 | -1.109 | -1.089 | -1.104 | -1.070 |
| 6 | 0.110 | -0.489 | -0.413 | -0.113 | 0.040 |
| 7 | 0.490 | 0.099 | -0.021 | -0.056 | -0.107 |
| 8 | 0.830 | 0.836 | 0.742 | 0.581 | 0.471 |
| 9 | -1.230 | -1.282 | -1.302 | -1.591 | -1.785 |
| 10 | -0.820 | -1.690 | -1.558 | -1.617 | -1.589 |
| 11 | -1.170 | -0.997 | -1.121 | -0.939 | -1.004 |
| 12 | -2.150 | -0.930 | -1.082 | -1.238 | -1.403 |
| 13 | -0.670 | -1.047 | -0.970 | -0.927 | -0.745 |
| 14 | -0.660 | -0.934 | -0.721 | -0.751 | -0.756 |
| 15 | -0.120 | -0.416 | -0.417 | -0.428 | -0.437 |
| 16 | -0.180 | 0.007 | -0.096 | -0.031 | -0.084 |
| 17 | -1.150 | -0.828 | -0.887 | -0.967 | -1.009 |
| 18 | -1.570 | -1.017 | -1.008 | -1.125 | -1.127 |
| 19 | -1.540 | -1.634 | -1.865 | -1.707 | -1.621 |
| 20 | -1.120 | -0.822 | -0.562 | -0.674 | -0.879 |
| 21 | -0.730 | -0.664 | -0.654 | -0.676 | -0.497 |
| 22 | -0.270 | -0.468 | -0.334 | -0.114 | -0.066 |
| 23 | -0.280 | -0.965 | -0.960 | -0.843 | -0.597 |
| 24 | -0.460 | -0.489 | -0.468 | -0.577 | -0.518 |
| 25 | -0.240 | -0.046 | -0.023 | 0.014 | -0.049 |
| 26 | -0.020 | -0.240 | -0.180 | -0.200 | -0.289 |
| 27 | 0.690 | 0.105 | 0.210 | 0.177 | 0.152 |
| 28 | 0.440 | 0.175 | 0.328 | 0.363 | 0.415 |
| 29 | 0.140 | 0.375 | 0.399 | 0.551 | 0.588 |
| 30 | 0.220 | 0.052 | -0.036 | -0.051 | 0.072 |

| | | | | | |
|------------------|--------|--------|--------|--------|--------|
| 31 | -0.080 | -0.184 | -0.267 | -0.293 | -0.393 |
| 32 | 0.370 | 0.490 | 0.381 | 0.455 | 0.439 |
| 33 | 1.010 | 0.760 | 0.797 | 0.740 | 0.672 |
| 34 | 0.900 | 0.854 | 0.902 | 0.838 | 0.808 |
| 35 | -0.150 | -0.210 | -0.183 | -0.292 | -0.434 |
| 36 | -0.170 | -0.033 | -0.017 | -0.031 | -0.096 |
| 37 | 0.970 | 0.749 | 0.780 | 0.695 | 0.680 |
| 38 | 1.040 | 0.720 | 0.729 | 0.745 | 0.703 |
| 39 | 0.080 | 0.408 | 0.398 | 0.371 | 0.479 |
| 40 | 0.400 | 0.548 | 0.553 | 0.560 | 0.767 |
| 41 | 0.000 | 0.042 | 0.051 | 0.113 | 0.101 |
| 42 | 0.240 | 0.331 | 0.351 | 0.419 | 0.310 |
| 43 | -0.160 | -0.412 | -0.470 | -0.334 | -0.114 |
| 44 | 0.130 | 0.231 | 0.277 | 0.320 | 0.285 |
| 45 | 0.350 | 0.540 | 0.563 | 0.416 | 0.278 |
| 46 | 0.810 | 0.892 | 0.960 | 0.817 | 0.735 |
| 47 | 0.800 | 0.777 | 0.820 | 0.843 | 0.783 |
| 48 | 0.420 | 0.383 | 0.330 | 0.396 | 0.558 |
| 49 | 0.040 | 0.308 | 0.302 | 0.311 | 0.261 |
| 50 | 0.930 | 0.694 | 0.893 | 0.960 | 1.052 |
| 51 | 0.760 | 0.708 | 0.699 | 0.706 | 0.725 |
| 52 | -0.160 | -0.175 | -0.171 | -0.182 | -0.100 |
| 53 | -0.150 | -0.286 | -0.363 | -0.418 | -0.412 |
| 54 | 0.270 | 0.486 | 0.495 | 0.553 | 0.562 |
| 55 | 0.370 | 0.592 | 0.431 | 0.340 | 0.266 |
| 56 | 0.340 | 0.564 | 0.577 | 0.641 | 0.589 |
| Setul de testare | | | | | |
| T1 | -0.06 | -0.295 | -0.386 | -0.418 | -0.508 |
| T2 | -1.4 | -0.77 | -0.578 | -0.648 | -0.714 |
| T3 | 0.25 | -0.181 | 0.01 | -0.206 | -0.281 |
| T4 | 0 | -0.324 | -0.083 | -0.132 | -0.223 |
| T5 | -0.34 | -1.001 | -0.683 | -0.752 | -0.733 |
| T6 | -1.34 | -1.724 | -1.713 | -1.896 | -2.21 |
| T7 | 0.85 | 0.839 | 1.103 | 0.972 | 0.93 |

Figura următoare redă grafic capacitatea predictivă a modelului cu 5 termeni propus de autori:



1.3. Setul de inhibitori ai anhidrazei carbonice (1)

÷ Lucrare: “Carbonic anhydrase inhibitors - Part 57: Quantum chemical QSAR of a group of 1,3,4-thiadiazole- and 1,3,4-thiadiazoline disulfonamides with carbonic anhydrase inhibitory properties”

÷ Autori: Claudiu T. SUPURAN, Brian W. CLARE

÷ Referință: European Journal of Medicinal Chemistry, 19(11), p. 1611-1621, 2002

÷ Rezumat: propune utilizarea de relații structură-activitate în combinație cu energia de solvatare pentru estimarea activității biologice a 40 de inhibitori ai anhidrazei carbonice.

÷ Descriere

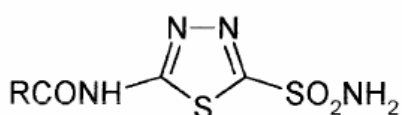
Expresii cuantice QSAR pentru 20 1,3,4-tiadiazol disulfonamide și 20 1,3,4-tiadiazoline disulfonamide, care sunt inhibitori ai anhidrazei carbonice, pentru izomerii CA I, CA II și CA IV au fost dezvoltate pe baza calculelor de tipul AM1 (Austin Model 1). Ca în multe studii anterioare, sarcina pe atomii din gruparea sulfonamică este de o importanță centrală, și de asemenea câmpul electric în punctele centrale ale grupării primare sulfonamidice. De asemenea, ca în studiile anterioare, polarizabilitatea moleculei este implicată într-o manieră anizotropică. O nouă facilitate este corelația cu energia de solvatare a moleculei, calculată prin modelul continuu COSMO.

Structura și activitatea setului de inhibitori ai anhidrazei carbonice

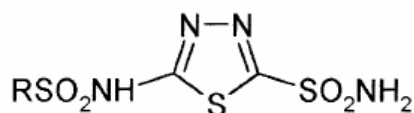
Multe variante structurale au fost derivate folosind acetazolamida A ca moleculă de plecare, cum ar fi aril/alchil sulfonilamido-1,3,4-tiadiazol-2-sulfonamidele B și C [Novel aromatic/heterocyclic sulfonamides and their metal complexes as inhibitors of carbonic anhydrase isozymes I, II and IV, Supuran, C.T., Scozzafava, A., Journal of Enzyme Inhibition, 12 (1), pp. 37-51, 1997] și [Carbonic anhydrase inhibitors - Part 70. Synthesis and ocular pharmacology of a new class of water-soluble, topically effective intraocular pressure lowering agents derived from nicotinic acid and aromatic/heterocyclic sulfonamides, Supuran, C.T., Scozzafava, A., Menabuoni, L., Mincione, F., Briganti, F., Mincione, G., European Journal of Medicinal Chemistry, 34 (10), pp. 799-808, 1999] (din care benzolamida B este cel mai important reprezentant [Maren T.H., In: Karch F.E. (Ed.), Orphan Drugs, Dekker, New York, 1982, pp 89-115], astfel încât alți derivați din această serie, de tipul C, pot fi folosiți pentru dezvoltarea instrumentelor de diagnostic în PET – pozitron emission tomography). Metazolamida D [1,3,4-Thiadiazole- and Thiadiazolinesulfonamides as Carbonic Anhydrase Inhibitors. Synthesis and Structural Studies, Richard W. YOUNG, Kathryn H. WOOD, Joyce A. EICHLER, James R. VAUGHAN, George W. ANDERSON, Journal of the American Chemical Society, 78(18), p. 4649-4654, 1956], a fost în consecință studiat deoarece posedă proprietăți farmacologice complet diferite comparat cu acetazolamida. Metazolamida este mai solubilă în lipide decât sulfonamidele menționate anterior și în consecință dozele sale de administrare sunt diferite; de asemenea penetrează mult mai ușor prin bariera BBB (blood-brain barrier), și este posibil să fie

folosită în tratamentul epilepsiei, de exemplu [Acetazolamide in the treatment of seizures, Reiss, W.G., Oles, K.S., Annals of Pharmacotherapy, 30 (5), pp. 514-518, 1996].

Recent, un număr mare de derivați de sulfonamido-1,3,4-tiazol-2-sulfonamidă de tipul C și tiadiazolii de tipul E corespunzători au fost raportați de autorii prezentei lucrări, în căutarea instrumentelor de diagnostic cu aplicații în PET. Acești compuși s-au dovedit foarte puternici inhibitori ai câtorva izomeri CA de importanță pentru psihologic, cum sunt CA I, II și IV. A rezultat astfel importanța încercării de a raționaliza activitatea biologică în clasa inhibitorilor CA prin modele QSAR, pentru cel puțin 2 motive: (i) pentru a încerca să se explice diferențele importante în afinitate pentru inhibitorii sulfonamidei între izomerii menționați, și de unde posibilitatea de a construi inhibitori izozime-specifici; și (ii) pentru a prezice, bazat pe astfel de calcule QSAR structura pentru inhibitori mult mai selectivi și mai puternici pentru acest tip de aplicații (PET), atâta timp cât compușii investigați anterior, de altfel acționând ca inhibitori puternici, duc la radio-randamente joase datorită problemelor ce apar în inserția izotopului emițător de pozitroni în structura lor moleculară (^{18}F în cazul menționat anterior).

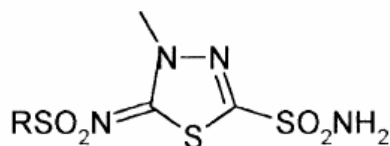


A: R = Me



B: R = Ph

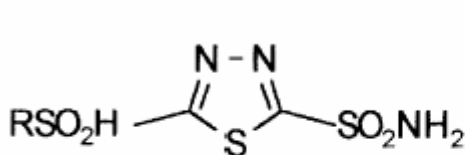
C: R = alkyl/aryl



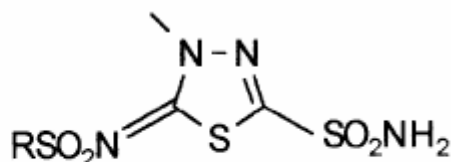
D: R = Me

E: R = alkyl/aryl

Structurile A-E discutate anterior



C



E

Structurile C și E incluse în acest studiu (sus, figura) și activitatea biologică (jos, tabel)

| Nr | Tip | R | IC ₅₀ (nM) | | |
|-----|-----|---|-----------------------|--------|--------|
| | | | HCA I | HCA II | HCA IV |
| Crt | R | R | | | |
| 1 | C | Me | 10 | 6 | 5 |
| 2 | C | PhCH ₂ | 7 | 5 | 6 |
| 3 | C | 4-Me-C ₆ H ₄ | 5 | 4 | 3 |
| 4 | C | 4-F-C ₆ H ₄ | 4 | 4 | 7 |
| 5 | C | 4-Cl-C ₆ H ₄ | 4 | 3 | 5 |
| 6 | C | 4-Br-C ₆ H ₄ | 3 | 2 | 4 |
| 7 | C | 4-MeO-C ₆ H ₄ | 5 | 3 | 4 |
| 8 | C | 4-AcNH-C ₆ H ₄ | 10 | 3 | 8 |
| 9 | C | 4-H ₂ N-C ₆ H ₄ | 6 | 2 | 5 |
| 10 | C | 3-H ₂ N-C ₆ H ₄ | 9 | 1 | 7 |
| 11 | C | 4-O ₂ N-C ₆ H ₄ | 3 | 1 | 2 |
| 12 | C | 3-O ₂ N-C ₆ H ₄ | 2 | 0.9 | 1 |
| 13 | C | 2-O ₂ N-C ₆ H ₄ | 5 | 3 | 4 |
| 14 | C | Me ₂ N | 19 | 8 | 13 |
| 15 | C | 2-HO ₂ CC ₆ H ₄ | 1 | 0.5 | 0.6 |
| 16 | C | 4-(2,4,6-Me ₃ Py ⁺)C ₆ H ₄ | 18 | 4 | 10 |
| 17 | C | 4-(2,4,6-Ph ₃ Py ⁺)C ₆ H ₄ | 360 | 110 | 320 |
| 18 | C | 2,4-(O ₂ N) ₂ C ₆ H ₃ | 12 | 5 | 28 |
| 19 | C | 4-Cl-3-O ₂ N-C ₆ H ₃ | 9 | 3 | 7 |
| 20 | C | 2,4,6-Me ₃ C ₆ H ₄ | 15 | 9 | 12 |
| 21 | E | Me | 17 | 4 | 8 |
| 22 | E | PhCH ₂ | 6 | 8 | 9 |
| 23 | E | 4-Me-C ₆ H ₄ | 5 | 3 | 3 |
| 24 | E | 4-F-C ₆ H ₄ | 8 | 4 | 7 |
| 25 | E | 4-Cl-C ₆ H ₄ | 8 | 3 | 5 |
| 26 | E | 4-Br-C ₆ H ₄ | 5 | 2 | 6 |
| 27 | E | 4-MeO-C ₆ H ₄ | 6 | 3 | 5 |
| 28 | E | 4-AcNH-C ₆ H ₄ | 2 | 0.7 | 2 |
| 29 | E | 4-H ₂ N-C ₆ H ₄ | 1 | 0.6 | 0.8 |
| 30 | E | 3-H ₂ N-C ₆ H ₄ | 1 | 0.5 | 0.8 |
| 31 | E | 4-O ₂ N-C ₆ H ₄ | 8 | 4 | 6 |

| | | | | | |
|----|---|----------------------|-----|-----|-----|
| 32 | E | 3-O2N-C6H4 | 7 | 2 | 5 |
| 33 | E | 2-O2N-C6H4 | 5 | 1 | 3 |
| 34 | E | Me2N | 9 | 5 | 8 |
| 35 | E | 2-HO2CC6H4 | 1 | 0.2 | 0.5 |
| 36 | E | 4-(2,4,6-Me3Py+)C6H4 | 17 | 4 | 12 |
| 37 | E | 4-(2,4,6-Ph3Py+)C6H4 | 455 | 110 | 180 |
| 38 | E | 2,4-(O2N)2C6H3 | 10 | 4 | 8 |
| 39 | E | 4-Cl-3-O2N-C6H3 | 7 | 2 | 5 |
| 40 | E | 2,4,6-Me3C6H4 | 13 | 7 | 9 |

Rezultate pentru legătura între structura și activitatea setului de inhibitori ai anhidrazei carbonice

În continuare, sarcinile atomilor grupării primare sulfonamidice vor fi referite ca QS1, QO1 QN1 și QH, și carbonul atașat prin inel ca QCr1. Valorile corespunzătoare pentru al doilea grup sulfonamidă sunt QS2, QO2, QN2 și QCr2. Sarcina pe atomul de hidrogen al celei de-a doua sulfonamidă nu a fost considerat. Astfel referindu-ne la figura de mai sus, QS1 este sarcina pe S6, QO1 este suma sarcinilor pe O10 și O11, și QH este suma sarcinilor pe H14 și H15, și QCr1 este sarcina pe C4. Sarcinile pe N2 și N3 sunt QNr1 și respectiv QNr2. Sarcina pe S5 nu a fost niciodată semnificativă.

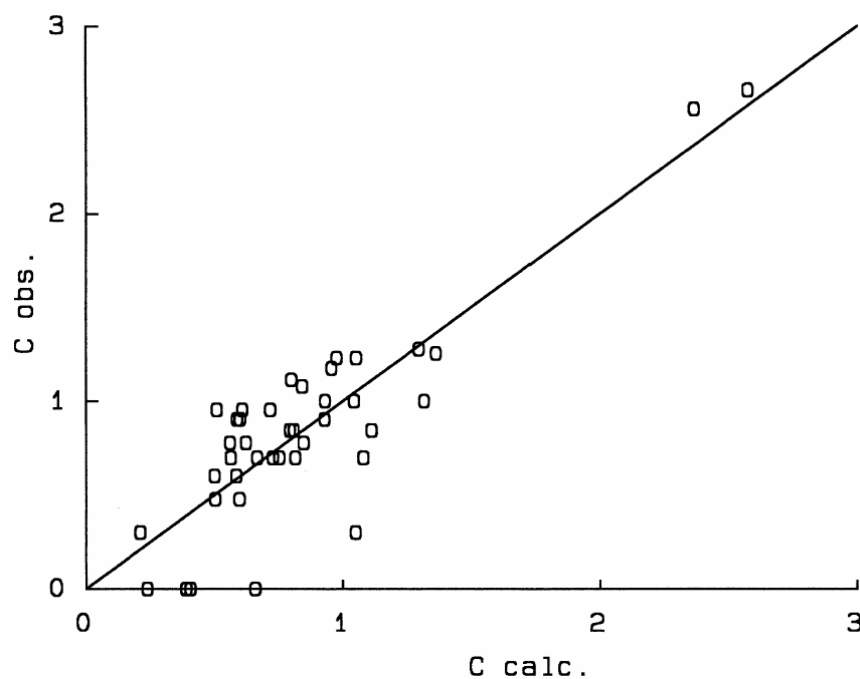
Momentele de dipol μ și componentele lor X, Y și Z au fost calculate ca valori așteptate ale operatorului de moment de dipol. Media sarcinilor absolute Mulliken a tuturor atomilor [Systematic QSAR Procedures with Quantum Chemical Descriptors, Osamu KIKUCHI, Quantitative Structure-Activity Relationships, 6(4), p. 153-210, 1987], indicele de dipol local [idem] și suma superdelocalizabilităților electrofilice au fost de asemenea calculate, utilizând un program scris de autori. Aria suprafeței și volumul fiecărei molecule a fost calculat prin metoda GEOPOL de teselație utilizând programul ARVOMOL. O valoare experimentală a coeficientului de distribuție bufer cloroform-apă a fost determinat pentru a fi folosit ca descriptor.

Următorul tabel redă rezultatele:

| Nr | CA | Compuși | Ecuatie |
|----|------|---------|--|
| 1 | CA I | Toți | $\text{Log IC}_{50} = C_1 \Pi_{xx} + C_2 \Pi_{zz} + C_3 Q_{Nr2} + C_4 Q_{S1} + C_5 Q_{S2} + C_6 \mu_x + C_7$ |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|----------|-----------------------|------------------------|---------|---------|---------|-----------------------|---------|
| C_i | 9.29×10^{-3} | -5.72×10^{-3} | -13.04 | 17.07 | 1.560 | 6.90×10^{-2} | -50.83 |
| σ | 1.38×10^{-3} | 2.31×10^{-3} | 2.87 | 4.59 | 0.790 | 1.82×10^{-2} | 12.29 |
| α | 0.00000 | 0.01863 | 0.00007 | 0.00075 | 0.05677 | 0.00062 | 0.00023 |

$N = 40, R^2 = 0.753, Q^2 = 0.628, s = 0.289, F = 16.78, \Lambda = 2.87$



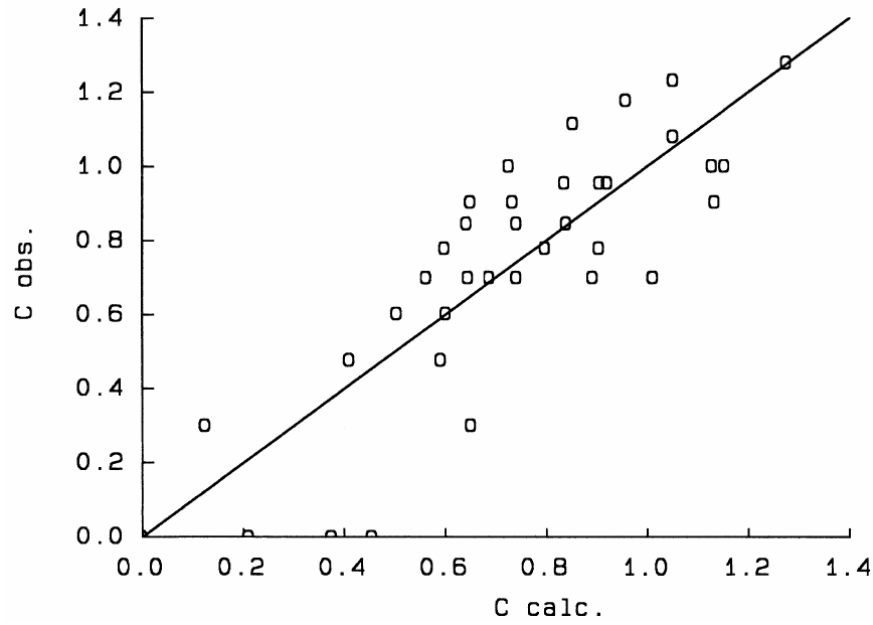
2

Fără
16, 17
36, 37

$$\text{Log IC}_{50} = C_1 \Pi_{zz} + C_2 Q_{Cr2} + C_3 \mu_x + C_4 \text{Log } P + C_5 Q_{O1} + C_6$$

| | 1 | 2 | 3 | 4 | 5 | 6 |
|----------|------------------------|---------|---------|---------|---------|---------|
| C_i | -3.68×10^{-3} | 3.152 | 0.157 | 0.400 | -24.62 | -44.1 |
| σ | 1.44×10^{-3} | 1.074 | 0.023 | 0.168 | 4.01 | 7.4 |
| α | 0.01588 | 0.00632 | 0.00000 | 0.02319 | 0.00000 | 0.00000 |

$N = 36, R^2 = 0.700, Q^2 = 0.570, s = 0.201, F = 13.98, \Lambda = 2.41$



| 3 | CA I | 1-20 | $\text{Log IC}_{50} = C_1 Q_{S1} + C_2 \mu_x + C_3 \mu_z + C_4 \Delta H_S + C_5 Q_{O1} + C_6$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|-----------------------|---------|---|---------|---------|---------|---------|---------|---|---|-------|-------|--------|-----------------------|---------|--------|--------|----------|--------|--------|--------|----------|-----------------------|-------|----------|---------|---------|---------|---------|---------|----------|---------|---------|---------|---------|---------|---------|---------|---------|
| <table border="1"> <thead> <tr> <th></th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> </tr> </thead> <tbody> <tr> <td>C_i</td> <td>59.43</td> <td>0.1359</td> <td>-0.0300</td> <td>-0.0204</td> <td>98.87</td> <td>27.83</td> </tr> <tr> <td>σ</td> <td>6.57</td> <td>0.0325</td> <td>0.0116</td> <td>0.0073</td> <td>10.30</td> <td>10.39</td> </tr> <tr> <td>α</td> <td>0.00000</td> <td>0.00093</td> <td>0.02193</td> <td>0.01451</td> <td>0.00000</td> <td>0.01800</td> </tr> </tbody> </table> <p>$N = 20, R^2 = 0.909, Q^2 = 0.502, s = 0.18, F = 27.94, \Lambda = 4.45$</p> | | | | | 1 | 2 | 3 | 4 | 5 | 6 | C_i | 59.43 | 0.1359 | -0.0300 | -0.0204 | 98.87 | 27.83 | σ | 6.57 | 0.0325 | 0.0116 | 0.0073 | 10.30 | 10.39 | α | 0.00000 | 0.00093 | 0.02193 | 0.01451 | 0.00000 | 0.01800 | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C_i | 59.43 | 0.1359 | -0.0300 | -0.0204 | 98.87 | 27.83 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| σ | 6.57 | 0.0325 | 0.0116 | 0.0073 | 10.30 | 10.39 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| α | 0.00000 | 0.00093 | 0.02193 | 0.01451 | 0.00000 | 0.01800 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4 | CA I | 21-40 | $\text{Log IC}_{50} = C_1 \Pi_{yy} + C_2 Q_{S2} + C_3 E_H + C_4 E_L + C_5 \Delta H_S + C_6 Q_{O1} + C_7 Q_{O2} + C_8$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1"> <thead> <tr> <th></th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> <th>7</th> <th>8</th> </tr> </thead> <tbody> <tr> <td>C_i</td> <td>8.47×10^{-3}</td> <td>-5.871</td> <td>-1.787</td> <td>-1.575</td> <td>0.0501</td> <td>-82.31</td> <td>-16.36</td> <td>-182.6</td> </tr> <tr> <td>σ</td> <td>1.43×10^{-3}</td> <td>1.791</td> <td>0.367</td> <td>0.329</td> <td>0.0100</td> <td>17.76</td> <td>4.16</td> <td>32.8</td> </tr> <tr> <td>α</td> <td>0.00007</td> <td>0.00659</td> <td>0.00038</td> <td>0.00044</td> <td>0.00031</td> <td>0.00058</td> <td>0.00198</td> <td>0.00012</td> </tr> </tbody> </table> <p>$N = 20, R^2 = 0.917, Q^2 = 0.712, s = 0.21, F = 18.92, \Lambda = 4.54$</p> | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | C_i | 8.47×10^{-3} | -5.871 | -1.787 | -1.575 | 0.0501 | -82.31 | -16.36 | -182.6 | σ | 1.43×10^{-3} | 1.791 | 0.367 | 0.329 | 0.0100 | 17.76 | 4.16 | 32.8 | α | 0.00007 | 0.00659 | 0.00038 | 0.00044 | 0.00031 | 0.00058 | 0.00198 | 0.00012 |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C_i | 8.47×10^{-3} | -5.871 | -1.787 | -1.575 | 0.0501 | -82.31 | -16.36 | -182.6 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| σ | 1.43×10^{-3} | 1.791 | 0.367 | 0.329 | 0.0100 | 17.76 | 4.16 | 32.8 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| α | 0.00007 | 0.00659 | 0.00038 | 0.00044 | 0.00031 | 0.00058 | 0.00198 | 0.00012 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 5 | CA II | Tođi | $\text{Log IC}_{50} = C_1 \Pi_{xx} + C_2 Q_{Cr1} + C_3 Q_{S1} + C_4 E_H + C_5 \mu_x + C_6 \mu_z + C_7 \Delta H_S + C_8$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

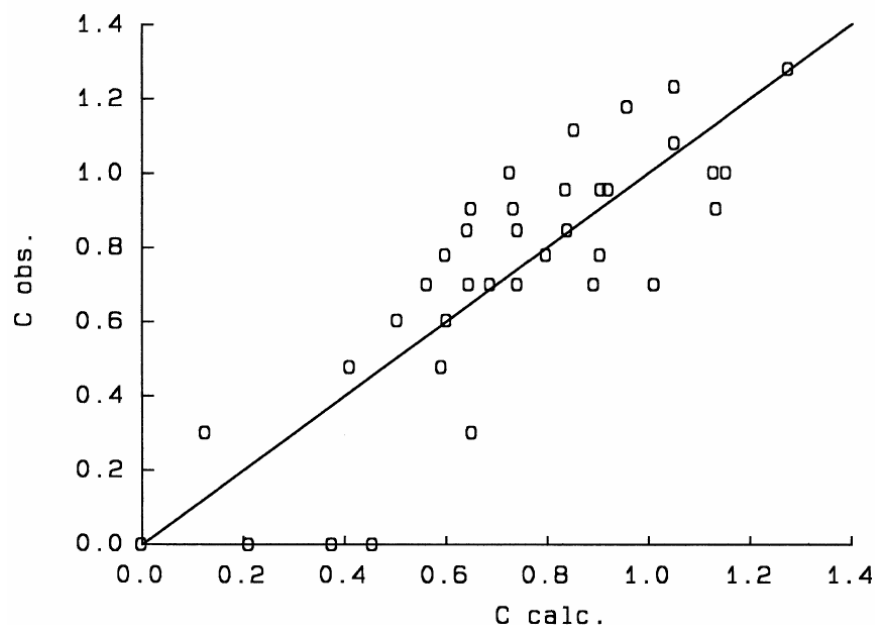
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|----------|-----------------------|---------|---------|---------|---------|---------|---------|---------|
| C_i | 8.92×10^{-3} | -6.68 | 18.97 | -0.736 | 0.0667 | -0.0417 | 0.0275 | -64.15 |
| σ | 1.20×10^{-3} | 1.56 | 4.94 | 0.266 | 0.0211 | 0.0160 | 0.0081 | 14.45 |
| α | 0.00000 | 0.00016 | 0.00056 | 0.00945 | 0.00345 | 0.01339 | 0.00178 | 0.00010 |

$N = 40, R^2 = 0.719, Q^2 = 0.475, s = 0.304, F = 11.70, \Lambda = 2.47$

| | | | |
|---|-------|------------------------|--|
| 6 | CA II | Fără 16,17 37,37 | $\text{Log IC}_{50} = C_1 Q_{Cr1} + C_2 Q_{S1} + C_3 E_H$ $+ C_4 \mu_x + C_5 \mu + C_6 \text{Log } P + C_7 \Delta H_S + C_8$ |
|---|-------|------------------------|--|

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|----------|---------|---------|---------|---------|---------|---------|---------|---------|
| C_i | -7.05 | 13.19 | -0.677 | 0.126 | -0.0421 | 0.298 | 0.0302 | -46.52 |
| σ | 1.07 | 4.35 | 0.140 | 0.016 | 0.0144 | 0.133 | 0.0083 | 11.93 |
| α | 0.00000 | 0.00520 | 0.00004 | 0.00000 | 0.00697 | 0.03277 | 0.00106 | 0.00055 |

$N = 36, R^2 = 0.876, Q^2 = 0.777, s = 0.152, F = 28.66, \Lambda = 2.60$



| | | | |
|---|-------|------|---|
| 7 | CA II | 1-20 | $\text{Log IC}_{50} = C_1 Q_{S1} + C_2 E_H + C_3 \mu_x$ $+ C_4 \mu_z + C_5 \Delta H_S + C_6 Q_{O1} + C_7$ |
|---|-------|------|---|

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|----------|---------|---------|---------|---------|---------|---------|---------|
| C_i | 63.34 | -0.696 | 0.1346 | -0.0504 | 0.0406 | 104.7 | 22.46 |
| σ | 6.97 | 0.204 | 0.0345 | 0.0135 | 0.0077 | 10.6 | 11.49 |
| α | 0.00000 | 0.00474 | 0.00181 | 0.00254 | 0.00016 | 0.00000 | 0.07243 |

$N = 20, R^2 = 0.902, Q^2 = 0.156, s = 0.18, F = 19.93, \Lambda = 4.24$

8 CA II 21-40 Nici o ecuație semnificativă statistic nu a fost obținută

9 CA IV Toți $\text{Log IC}_{50} = C_1 \Pi_{xx} + C_2 Q_{Cr1} + C_3 Q_{S1} + C_4 \mu_x + C_5$

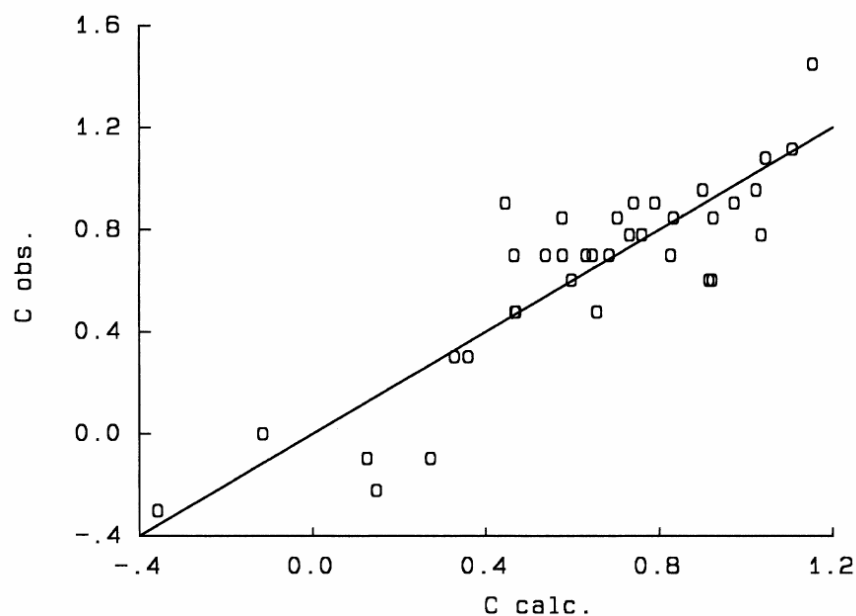
| | 1 | 2 | 3 | 4 | 5 |
|----------|-----------------------|---------|---------|---------|---------|
| C_i | 7.31×10^{-3} | -5.570 | 11.46 | 0.0602 | -37.16 |
| σ | 9.65×10^{-4} | 1.431 | 4.90 | 0.0201 | 13.68 |
| α | 0.00000 | 0.00043 | 0.02523 | 0.00492 | 0.01022 |

$N = 40, R^2 = 0.632, Q^2 = 0.444, s = 0.34, F = 15.07, \Lambda = 1.48$

10 CA IV Fără
16, 17
36, 37 $\text{Log IC}_{50} = C_1 Q_{Nr1} + C_2 Q_{S1} + C_3 \mu_x + C_4 \text{Log } P + C_5 \Delta H_{S2} + C_6$

| | 1 | 2 | 3 | 4 | 5 | 6 |
|----------|---------|---------|---------|---------|---------|---------|
| C_i | -2.798 | 8.447 | 0.1800 | 0.5295 | 0.0298 | -20.11 |
| σ | 0.336 | 3.732 | 0.0230 | 0.1682 | 0.0076 | 10.02 |
| α | 0.00000 | 0.03103 | 0.00000 | 0.00371 | 0.00047 | 0.05375 |

$N = 36, R^2 = 0.769, Q^2 = 0.646, s = 0.20, F = 20.05, \Lambda = 1.99$



| | | | | | | | |
|----|-------------|-------|---|---------|---------|---------|---------|
| 11 | Log IC50 | 1-20 | $\text{Log IC}_{50} = C_1 \Pi_{xx} + C_2 Q_{N1} + C_3 \mu_x + C_4$ | | | | |
| | | | 1 | 2 | 3 | 4 | |
| | C_i | | 7.36×10^{-3} | -18.90 | 0.1228 | -22.29 | |
| | σ | | 1.04×10^{-3} | 6.72 | 0.0452 | 7.78 | |
| | α | | 0.00000 | 0.01260 | 0.01525 | 0.01124 | |
| | | | $N = 20, R^2 = 0.760, Q^2 = 0.340, s = 0.29, F = 16.8, \Lambda = 1.14$ | | | | |
| 12 | Log IC50 | 21-40 | $\text{Log IC}_{50} = C_1 \Pi_{yy} + C_2 E_H + C_3 \mu_x$ $+ C_4 Q_{O1} + C_5$ | | | | |
| | | | 1 | 2 | 3 | 4 | 5 |
| | C_i | | 7.29×10^{-3} | -1.628 | 0.0977 | -105.4 | -208.9 |
| | σ | | 1.22×10^{-3} | 0.384 | 0.0226 | 21.5 | 40.1 |
| | α | | 0.00003 | 0.00072 | 0.00062 | 0.00019 | 0.00011 |
| | | | $N = 20, R^2 = 0.822, Q^2 = 0.358, s = 0.25, F = 17.23, \Lambda = 1.96$ | | | | |

Tabelul următor prezintă estimatorii semnificațiilor statistice ale modelelor de regresie obținute:

| | CA I | CA II | CA IV |
|--------------------|----------------------|----------------------|----------------------|
| All compounds | 9.7×10^{-4} | 1.3×10^{-3} | 2.6×10^{-3} |
| First 20 compounds | 2.3×10^{-5} | 7.5×10^{-4} | 5.0×10^{-4} |
| Last 20 compounds | 0.036 | 0.236 | 0.063 |

1.4. Setul de inhibitori ai anhidrazei carbonice (2)

÷ Lucrare: “Modelling the Inhibitory Activity on Carbonic Anhydrase IV of Substituted Thiadiazole- and Thiadiazoline- Disulfonamides: Integration of Structure Information”

÷ Autori: Lorentz JÄNTSCHI, Sorana-Daniela BOLBOACĂ

÷ Referință: Electronic Journal of Biomedicine, 2006(2), p. 22-33, 2006

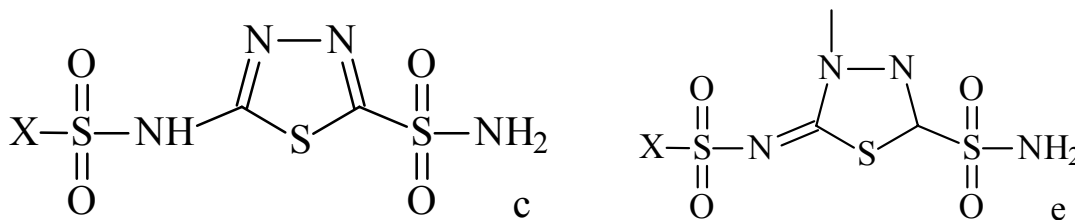
÷ Rezumat: propune relațiilor structură-activitate în convergență cu familia descriptorilor moleculari (MDF) pentru modelarea activității biologice a 40 de inhibitori ai anhidrazei carbonice IV (CA IV). Autorii obțin că activitatea substituenților 1,3,4-tiadiazol- și 1,3,4-tiadiazolin- disulfonamidelor este manifestată ca efect deopotrivă al topologiei moleculare și geometriei moleculare a acestor compuși și este datorată sarcinii electrice parțiale pe atomii acestora.

÷ Descriere

Pornind de la rezultatele de succes obținute de metoda originală a familiei de descriptori moleculari (MDF) aplicată pe relațiile structură-activitate (SAR), scopul cercetării autorilor a fost de a investiga abilitățile estimative și predictive ale metodologiei MDF-SAR pe seria de compuși substituiți de tiazol- și tiazolin- disulfonamide. Rezultatele obținute s-au dovedit superioare celor raportate cu alte metode de investigare (mai puțin elaborate).

Structura și activitatea setului de inhibitori ai anhidrazei carbonice IV

Tabelul următor redă compușii incluși în studiu și activitățile acestora (exprimate în logIC₅₀):



| Abb. | X | log IC ₅₀ (nM) | Abb. | X | log IC ₅₀ (nM) |
|------|--|---------------------------|------|--|---------------------------|
| c_01 | Me | 0.6990 | e_01 | Me | 0.9031 |
| c_02 | PhCH ₂ | 0.7782 | e_02 | PhCH ₂ | 0.9542 |
| c_03 | 4-Me-C ₆ H ₄ | 0.4771 | e_03 | 4-Me-C ₆ H ₄ | 0.4771 |
| c_04 | 4-F-C ₆ H ₄ | 0.8451 | e_04 | 4-F-C ₆ H ₄ | 0.8451 |
| c_05 | 4-Cl-C ₆ H ₄ | 0.6990 | e_05 | 4-Cl-C ₆ H ₄ | 0.6990 |
| c_06 | 4-Br-C ₆ H ₄ | 0.6021 | e_06 | 4-Br-C ₆ H ₄ | 0.7782 |
| c_07 | 4-MeO-C ₆ H ₄ | 0.6021 | e_07 | 4-MeO-C ₆ H ₄ | 0.6990 |
| c_08 | 4-AcNH-C ₆ H ₄ | 0.9031 | e_08 | 4-AcNH-C ₆ H ₄ | 0.3010 |
| c_09 | 4-H ₂ N-C ₆ H ₄ | 0.6990 | e_09 | 4-H ₂ N-C ₆ H | -0.0969 |
| c_10 | 3-H ₂ N-C ₆ H ₄ | 0.8451 | e_10 | 3-H ₂ N-C ₆ H ₄ | -0.0969 |
| c_11 | 4-O ₂ N-C ₆ H ₄ | 0.3010 | e_11 | 4-O ₂ N-C ₆ H ₄ | 0.7782 |

| | | | | | |
|------|---|---------|------|---|---------|
| c_12 | 3-O ₂ N-C ₆ H ₄ | 0.0000 | e_12 | 3-O ₂ N-C ₆ H ₄ | 0.6990 |
| c_13 | 2-O ₂ N-C ₆ H ₄ | 0.6021 | e_13 | 2-O ₂ N-C ₆ H ₄ | 0.4771 |
| c_14 | Me ₂ N | 1.1139 | e_14 | Me ₂ N | 0.9031 |
| c_15 | 2-HO ₂ CC ₆ H ₄ | -0.2218 | e_15 | 2-HO ₂ CC ₆ H ₄ | -0.3010 |
| c_16 | 4-(2,4,6-Me ₃ Py ⁺)C ₆ H ₄ | 1.0000 | e_16 | 4-(2,4,6-Me ₃ Py ⁺)C ₆ H ₄ | 1.0792 |
| c_17 | 4-(2,4,6-Ph ₃ Py ⁺)C ₆ H ₄ | 2.5051 | e_17 | 4-(2,4,6-Ph ₃ Py ⁺)C ₆ H ₄ | 2.2553 |
| c_18 | 2,4-(O ₂ N) ₂ C ₆ H ₃ | 1.4472 | e_18 | 2,4-(O ₂ N) ₂ C ₆ H ₃ | 0.9031 |
| c_19 | 4-Cl-3-O ₂ N-C ₆ H ₃ | 0.8451 | e_19 | 4-Cl-3-O ₂ N-C ₆ H ₃ | 0.6990 |
| c_20 | 2,4,6-Me ₃ C ₆ H ₄ | 1.0792 | e_20 | 2,4,6-Me ₃ C ₆ H ₄ | 0.9542 |

Următoarele două ecuații (una cu 2 descriptori și cealaltă cu 4) s-au dovedit cele mai bune modele predictive:

$$\hat{Y}_{2d} = 0.802 + 0.111 \cdot \text{inPRIQg} + 9.980 \cdot 10^{-9} \cdot \text{iHMMTQt}$$

$$\hat{Y}_{4d} = 0.625 + 0.105 \cdot \text{inPRIQg} + 9.919 \cdot 10^{-9} \cdot \text{iHMMTQt} - 9.248 \cdot \text{IHMDTQg} + 1.727 \cdot \text{InPdJQg}$$

Tabelul următor redă valorile calculate ale descriptorilor și valorile prezise ale activității biologice cu fiecare din ecuațiile de mai sus:

| Abb. | inPRIQg | iHMMTQt | IHMDTQg | InPdJQg | \hat{Y}_{2d} | \hat{Y}_{4d} |
|------|---------|-------------------|---------|---------|----------------|----------------|
| c_01 | -0.7353 | $1.18 \cdot 10^5$ | 0.0007 | 0.0956 | 0.7222 | 0.7076 |
| c_02 | -2.2423 | $3.95 \cdot 10^6$ | 0.0003 | 0.2080 | 0.5927 | 0.7850 |
| c_03 | -2.8218 | $1.43 \cdot 10^6$ | 0.0012 | 0.1283 | 0.5029 | 0.5535 |
| c_04 | -0.1521 | $1.11 \cdot 10^5$ | 0.0218 | 0.2435 | 0.7870 | 0.8293 |
| c_05 | -0.0950 | $2.06 \cdot 10^5$ | 0.0105 | 0.1693 | 0.7943 | 0.8126 |
| c_06 | -0.0102 | $1.47 \cdot 10^6$ | 0.0009 | 0.0309 | 0.8164 | 0.6836 |
| c_07 | -0.2617 | $1.91 \cdot 10^5$ | 0.0163 | 0.1145 | 0.7756 | 0.6462 |
| c_08 | -2.0460 | $3.17 \cdot 10^5$ | 0.0296 | 0.3202 | 0.5782 | 0.6923 |
| c_09 | -3.5436 | $1.63 \cdot 10^5$ | 0.0130 | 0.3811 | 0.4100 | 0.7923 |
| c_10 | -3.5385 | $1.51 \cdot 10^5$ | 0.0386 | 0.3287 | 0.4104 | 0.4655 |
| c_11 | -0.5461 | $1.64 \cdot 10^5$ | 0.0405 | 0.1993 | 0.7437 | 0.5392 |
| c_12 | -1.7861 | $1.42 \cdot 10^5$ | 0.1154 | 0.3446 | 0.6054 | -0.0333 |
| c_13 | -1.6233 | $1.22 \cdot 10^5$ | 0.0522 | 0.2935 | 0.6233 | 0.4799 |
| c_14 | -0.8071 | $5.36 \cdot 10^4$ | 0.0045 | 0.2888 | 0.7135 | 0.9980 |
| c_15 | -3.6509 | $1.34 \cdot 10^5$ | 0.0399 | 0.0129 | 0.3977 | -0.1043 |
| c_16 | -1.0801 | $4.10 \cdot 10^7$ | 0.0009 | 0.0611 | 1.0918 | 1.0155 |
| c_17 | -0.0312 | $1.67 \cdot 10^8$ | 0.0016 | 0.1169 | 2.4673 | 2.4664 |

| | | | | | | |
|------|---------|-------------------|--------|--------|---------|---------|
| c_18 | -0.4637 | $4.29 \cdot 10^5$ | 0.0313 | 0.7259 | 0.7555 | 1.5449 |
| c_19 | -0.4667 | $3.24 \cdot 10^5$ | 0.0067 | 0.1848 | 0.7541 | 0.8363 |
| c_20 | -0.0543 | $1.54 \cdot 10^6$ | 0.0011 | 0.1266 | 0.8122 | 0.8433 |
| e_01 | -0.8572 | $5.09 \cdot 10^5$ | 0.0003 | 0.0546 | 0.7125 | 0.6317 |
| e_02 | -0.9993 | $4.97 \cdot 10^6$ | 0.0004 | 0.1940 | 0.7412 | 0.9006 |
| e_03 | -0.0549 | $1.78 \cdot 10^6$ | 0.0007 | 0.0203 | 0.8145 | 0.6655 |
| e_04 | -0.6339 | $3.50 \cdot 10^5$ | 0.0098 | 0.0531 | 0.7358 | 0.5631 |
| e_05 | -0.0465 | $4.11 \cdot 10^5$ | 0.0065 | 0.1281 | 0.8018 | 0.7856 |
| e_06 | -0.3026 | $4.00 \cdot 10^6$ | 0.0002 | 0.0613 | 0.8091 | 0.7366 |
| e_07 | -0.8485 | $5.46 \cdot 10^5$ | 0.0076 | 0.1111 | 0.7138 | 0.6627 |
| e_08 | -0.6516 | $1.01 \cdot 10^6$ | 0.0121 | 0.1075 | 0.7404 | 0.6402 |
| e_09 | -9.0604 | $4.27 \cdot 10^5$ | 0.0056 | 0.1805 | -0.2016 | -0.0627 |
| e_10 | -8.7988 | $4.10 \cdot 10^5$ | 0.0087 | 0.1296 | -0.1726 | -0.1519 |
| e_11 | -0.3044 | $5.87 \cdot 10^5$ | 0.0139 | 0.1104 | 0.7748 | 0.6609 |
| e_12 | -0.5566 | $5.19 \cdot 10^5$ | 0.0200 | 0.2415 | 0.7461 | 0.8040 |
| e_13 | -0.7842 | $4.47 \cdot 10^5$ | 0.0211 | 0.2449 | 0.7200 | 0.7744 |
| e_14 | -0.2030 | $1.27 \cdot 10^5$ | 0.0043 | 0.2208 | 0.7815 | 0.9468 |
| e_15 | -7.5361 | $4.67 \cdot 10^5$ | 0.0195 | 0.1050 | -0.0315 | -0.1610 |
| e_16 | -0.2658 | $2.75 \cdot 10^7$ | 0.0009 | 0.0621 | 1.0475 | 0.9687 |
| e_17 | -0.0708 | $1.50 \cdot 10^8$ | 0.0010 | 0.1308 | 2.2873 | 2.3170 |
| e_18 | -0.0811 | $9.32 \cdot 10^5$ | 0.0202 | 0.2532 | 0.8031 | 0.8759 |
| e_19 | -0.8712 | $6.99 \cdot 10^5$ | 0.0128 | 0.2206 | 0.7128 | 0.8033 |
| e_20 | -0.0068 | $3.85 \cdot 10^6$ | 0.0009 | 0.0937 | 0.8405 | 0.8158 |

Parametrii statistici pentru modelele obținute sunt redade în tabelul următor:

| Parametru | Valoare | |
|-------------|------------------------|------------------------|
| | Model cu 2 descriptori | Model cu 4 descriptori |
| n | 40 | 40 |
| v | 2 | 4 |
| r | 0.8673 | 0.9593 |
| r^2 | 0.7521 | 0.9202 |
| r^2_{adj} | 0.7387 | 0.9111 |
| S_{est} | 0.2741 | 0.1599 |
| F_{est} | 56 | 101 |

| | | |
|-------------------------|---------|---------|
| p_{est} | < 0.001 | < 0.001 |
| r^2_{cv-loo} | 0.7279 | 0.9034 |
| S_{loo} | 0.2873 | 0.1762 |
| F_{pred} | 49 | 82 |
| p_{pred} | < 0.001 | < 0.001 |
| $r^2 - r^2_{cv-loo}$ | 0.0242 | 0.0168 |
| $r^2(inPRIQg, iHMMTQt)$ | 0.0271 | 0.0271 |
| $r^2(inPRIQg, IHMDTQg)$ | n.a. | 0.0122 |
| $r^2(inPRIQg, InPdJQg)$ | n.a. | 0.0010 |
| $r^2(iHMMTQt, IHMDTQg)$ | n.a. | 0.0365 |
| $r^2(iHMMTQt, InPdJQg)$ | n.a. | 0.0212 |
| $r^2(IHMDTQg, InPdJQg)$ | n.a. | 0.2085 |

n = number of compounds, v = number of descriptors, n.a. = not applicable

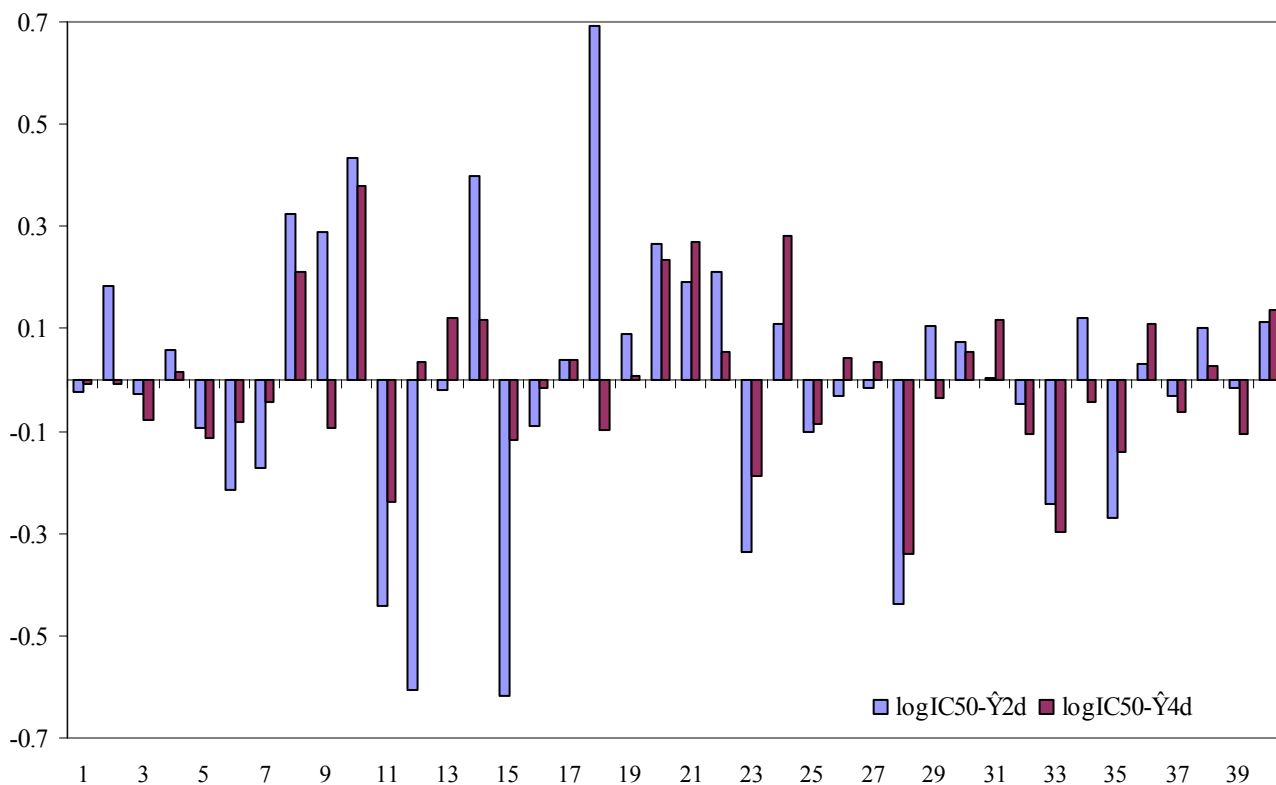
Semnificația statistică a parametrilor de regresie ai modelelor obținute sunt redate în tabelul următor:

| | StdError | $r^2(Y, desc)$ | t | 95%CI _c |
|------------------------|----------------------|----------------|----------|--|
| Model cu 2 descriptori | | | | |
| <i>Intercept</i> | 0.0549 | | 14.616* | [0.6915, 0.9141] |
| <i>inPRIQg</i> | 0.0197 | 0.3335 | 5.6589* | [0.0715, 0.1512] |
| <i>iHMMTQt</i> | $1.26 \cdot 10^{-9}$ | 0.5376 | 7.9051* | $[7.4 \cdot 10^{-9}, 1.2 \cdot 10^{-8}]$ |
| Model cu 4 descriptori | | | | |
| <i>Intercept</i> | 0.0485 | | 12.891* | [0.5265, 0.7233] |
| <i>inPRIQg</i> | 0.0115 | 0.3335 | 9.1175* | [0.0817, 0.1284] |
| <i>iHMMTQt</i> | 0.0000 | 0.5376 | 13.223* | $[8.4 \cdot 10^{-9}, 1.1 \cdot 10^{-8}]$ |
| <i>IHMDTQg</i> | 1.3826 | 0.1198 | -6.6894* | [-12.055, -6.4418] |
| <i>InPdJQg</i> | 0.2218 | 0.0208 | 7.7875* | [1.2771, 2.1778] |

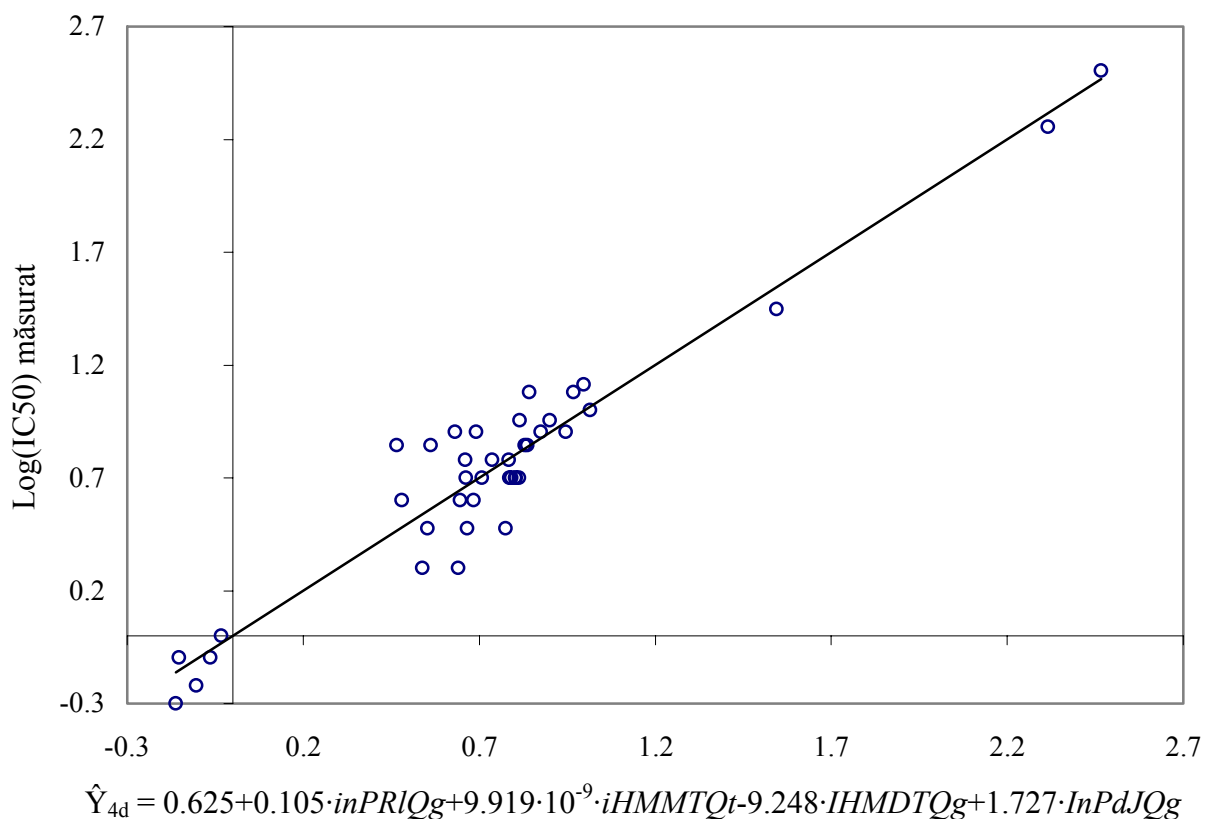
StdError = standard error, Y = log IC₅₀, desc = molecular descriptor, t = parameter of the Student test; * p < 0.001

Rezultate pentru legătura între structura și activitatea setului de inhibitori ai CA IV:

Figura următoare reprezintă diferențele reziduurilor obținute între model și experiment pentru cele 2 modele selectate:



Legătura între structură prin intermediul descriptorilor MDF și activitate prin intermediul lui logIC50 este redată în figura de mai jos:



Analiza corelațiilor corelate a fost aplicată pentru a se vedea dacă există diferență semnificativă între modelele propuse. Rezultatele testului (numit Steiger, după numele inventatorului) sunt redată mai jos:

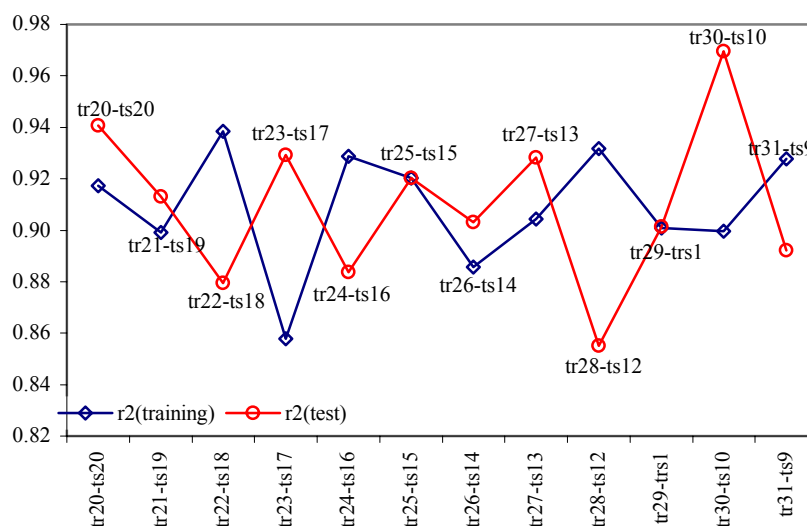
| Parametru | Valoare |
|-------------------------------|----------------------|
| $r_{(\log IC50 - \hat{Y}4d)}$ | 0.9593 |
| $r_{(\log IC50 - \hat{Y}2d)}$ | 0.8670 |
| $r_{(\hat{Y}4d - \hat{Y}2d)}$ | 0.9041 |
| Steiger's Z | 3.9769 |
| $p_{Steiger's Z}$ | $3.49 \cdot 10^{-5}$ |

Testul *Training versus Test* a fost aplicat pentru a verifica capacitatea predictivă a modelelor propuse. Rezultatele acestui test (în care s-au inclus întâmplător moleculele în cele 2 seturi, de învățare și de test) sunt în tabelul de mai jos:

| No. | a_0 | a_1 | a_2 | a_3 | a_4 | Training | | | | Test | | | | $r_{tr} vs. r_{ts}$ |
|-----|-------|-------|----------------------|--------|-------|-----------|----------|----------------|----------|-----------|----------|----------------|----------|---------------------|
| | | | | | | No_{tr} | r_{tr} | 95%CI $_{tr}$ | F_{tr} | No_{ts} | r_{ts} | 95% CI $_{ts}$ | F_{ts} | |
| 1 | 0.307 | 0.086 | $1.01 \cdot 10^{-8}$ | -8.543 | 1.603 | 20 | 0.958 | [0.894, 0.983] | 42*** | 20 | 0.970 | [0.924, 0.988] | 30*** | 0.51† |
| 2 | 0.411 | 0.102 | $9.79 \cdot 10^{-9}$ | -10.04 | 2.263 | 21 | 0.948 | [0.874, 0.979] | 36*** | 19 | 0.956 | [0.886, 0.983] | 35*** | 0.23† |
| 3 | 0.153 | 0.094 | $1.02 \cdot 10^{-8}$ | -8.355 | 1.726 | 22 | 0.969 | [0.925, 0.987] | 65*** | 18 | 0.938 | [0.838, 0.977] | 20*** | 1.02† |
| 4 | 0.145 | 0.070 | $1.04 \cdot 10^{-8}$ | -8.587 | 1.563 | 23 | 0.926 | [0.831, 0.968] | 27*** | 17 | 0.964 | [0.901, 0.987] | 32*** | 1.06† |
| 5 | 0.232 | 0.115 | $9.94 \cdot 10^{-9}$ | -9.229 | 1.756 | 24 | 0.964 | [0.917, 0.984] | 62*** | 16 | 0.940 | [0.832, 0.979] | 19*** | 0.73† |
| 6 | 0.499 | 0.122 | $1.02 \cdot 10^{-8}$ | -9.243 | 1.563 | 25 | 0.959 | [0.908, 0.982] | 58*** | 15 | 0.959 | [0.879, 0.987] | 22*** | 0.00† |
| 7 | 0.406 | 0.057 | $1.03 \cdot 10^{-8}$ | -9.432 | 1.572 | 26 | 0.941 | [0.871, 0.973] | 4*** | 14 | 0.950 | [0.846, 0.984] | 13*** | 0.24† |
| 8 | 0.406 | 0.104 | $9.59 \cdot 10^{-9}$ | -9.076 | 1.414 | 27 | 0.951 | [0.894, 0.978] | 52*** | 13 | 0.964 | [0.879, 0.989] | 19*** | 0.24† |
| 9 | 0.154 | 0.101 | $1.01 \cdot 10^{-8}$ | -9.838 | 1.823 | 28 | 0.965 | [0.926, 0.984] | 78*** | 12 | 0.925 | [0.747, 0.979] | 5* | 1.02† |
| 10 | 0.320 | 0.111 | $9.93 \cdot 10^{-9}$ | -9.160 | 1.642 | 29 | 0.927 | [0.849, 0.965] | 76*** | 11 | 0.949 | [0.812, 0.987] | 10** | 0.47† |
| 11 | 0.073 | 0.104 | $9.44 \cdot 10^{-9}$ | -8.660 | 1.622 | 30 | 0.949 | [0.893, 0.975] | 56*** | 10 | 0.985 | [0.934, 0.996] | 20** | 1.45† |
| 12 | 0.320 | 0.102 | $1.03 \cdot 10^{-8}$ | -9.795 | 1.838 | 31 | 0.963 | [0.924, 0.982] | 84*** | 9 | 0.944 | [0.752, 0.988] | 7* | 0.47† |

* $0.01 < p < 0.05$, ** $0.001 < p < 0.01$, *** $p < 0.001$, † $p > 0.05$

Variația coeficientului de corelație obținută pentru cele 12 experimente TvT executate este redată în figura de mai jos:



1.5. Setul de inhibitori ai anhidrazei carbonice (2)

÷ Lucrare: “ Complex Structural Information Integration: Inhibitor Activity on Carbonic Anhydrase II of Substituted Disulfonamides”

÷ Autori: Lorentz JÄNTSCHI, Mihaela Ligia UNGUREȘAN, Sorana Daniela BOLBOACĂ

÷ Referință: Applied Medical Informatics, 17(3-4), p. 12-21, 2005

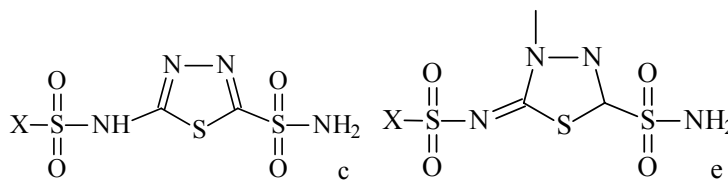
÷ Rezumat: propune relațiilor structură-activitate în convergență cu familia descriptorilor moleculari (MDF) pentru modelarea activității biologice a 40 de inhibitori ai anhidrazei carbonice II (CA II).

÷ Descriere

În acord cu abilitățile estimative și predictive obținute prin aplicarea metodei familiei de descriptori moleculari pe relații structură-activitate (MDF-SAR) [Jäntschi L. Delphi Client - Server Implementation of Multiple Linear Regression Findings: a QSAR/QSPR Application. Applied Medical Informatics 2004;15(3-4):48-55] și [Bolboacă S, Jäntschi L. Molecular Descriptors Family on Structure Activity Relationships 3. Antituberculosic Activity of some Polyhydroxyxanthenes, Leonardo Journal of Sciences 2005;5(7):58-64], această abordare a fost aplicată pe un set de 40 substituenți de 1,3,4-tiazol- și 1,3,4-tiazolin- disulfonamide cu activitate inhibitoare pe CAII și abilitățile în estimare și predicție ale modelelor obținute au fost investigate.

Structura și activitatea setului de inhibitori ai anhidrazei carbonice

Tabelul următor redă compuşii incluși în studiu și activitățile acestora (exprimate în logIC₅₀):



| Abb. | X | log IC ₅₀ (nM) |
|------|--|---------------------------|
| c_01 | Me | 0.7782 |
| c_02 | PhCH ₂ | 0.6990 |
| c_03 | 4-Me-C ₆ H ₄ | 0.6021 |
| c_04 | 4-F-C ₆ H ₄ | 0.6021 |
| c_05 | 4-Cl-C ₆ H ₄ | 0.4771 |
| c_06 | 4-Br-C ₆ H ₄ | 0.3010 |
| c_07 | 4-MeO-C ₆ H ₄ | 0.4771 |
| c_08 | 4-AcNH-C ₆ H ₄ | 0.4771 |
| c_09 | 4-H ₂ N-C ₆ H ₄ | 0.3010 |
| c_10 | 3-H ₂ N-C ₆ H ₄ | 0.0000 |
| c_11 | 4-O ₂ N-C ₆ H ₄ | 0.0000 |

| | | |
|------|---|---------|
| c_12 | 3-O ₂ N-C ₆ H ₄ | -0.0458 |
| c_13 | 2-O ₂ N-C ₆ H ₄ | 0.4771 |
| c_14 | Me ₂ N | 0.9031 |
| c_15 | 2-HO ₂ CC ₆ H ₄ | -0.3010 |
| c_16 | 4-(2,4,6-Me ₃ Py ⁺)C ₆ H ₄ | 0.6021 |
| c_17 | 4-(2,4,6-Ph ₃ Py ⁺)C ₆ H ₄ | 2.0414 |
| c_18 | 2,4-(O ₂ N) ₂ C ₆ H ₃ | 0.6990 |
| c_19 | 4-Cl-3-O ₂ N-C ₆ H ₃ | 0.4771 |
| c_20 | 2,4,6-Me ₃ C ₆ H ₄ | 0.9542 |
| e_01 | Me | 0.6021 |
| e_02 | PhCH ₂ | 0.9031 |
| e_03 | 4-Me-C ₆ H ₄ | 0.4771 |
| e_04 | 4-F-C ₆ H ₄ | 0.6021 |
| e_05 | 4-Cl-C ₆ H ₄ | 0.4771 |
| e_06 | 4-Br-C ₆ H ₄ | 0.3010 |
| e_07 | 4-MeO-C ₆ H ₄ | 0.4771 |
| e_08 | 4-AcNH-C ₆ H ₄ | -0.1549 |
| e_09 | 4-H ₂ N-C ₆ H | -0.2218 |
| e_10 | 3-H ₂ N-C ₆ H ₄ | -0.3010 |
| e_11 | 4-O ₂ N-C ₆ H ₄ | 0.6021 |
| e_12 | 3-O ₂ N-C ₆ H ₄ | 0.3010 |
| e_13 | 2-O ₂ N-C ₆ H ₄ | 0.0000 |
| e_14 | Me ₂ N | 0.6990 |
| e_15 | 2-HO ₂ CC ₆ H ₄ | -0.6990 |
| e_16 | 4-(2,4,6-Me ₃ Py ⁺)C ₆ H ₄ | 0.6021 |
| e_17 | 4-(2,4,6-Ph ₃ Py ⁺)C ₆ H ₄ | 2.0414 |
| e_18 | 2,4-(O ₂ N) ₂ C ₆ H ₃ | 0.6021 |
| e_19 | 4-Cl-3-O ₂ N-C ₆ H ₃ | 0.3010 |
| e_20 | 2,4,6-Me ₃ C ₆ H ₄ | 0.8451 |

Me = methyl; Ph = phenyl; Ac = acetyl; Py⁺ = pyridine

Rezultate pentru legătura între structura și activitatea setului de inhibitori ai CA II

Tabelul de mai jos prezintă ecuațiile MDF-SAR obținute:

| N | Ecuatie |
|---|--|
| 1 | $\hat{Y}_{2-v} = -4.4479 + 2.4352 \cdot imDdSCg + 9.4635 \cdot 10^{-2} \cdot iiMrqQg$ |
| 2 | $\hat{Y}_{4-v} = -9.9859 + 4.5643 \cdot imDdSCg + 2.945 \cdot 10^{-3} \cdot isDrqQg + 5.2036 \cdot IIMDQQg + 1.4832 \cdot lmMrsGg$ |

Valorile descriptorilor moleculari sunt redate în tabelul de mai jos, împreună cu valorile prezise de modelul cu 2 și cu 4 descriptori:

| No. | Abb. | Bi-varied model | | | Tetra-varied model | | | \hat{Y}_{4-v} |
|-----|------|-----------------|---------|---------|--------------------|---------|---------|-----------------|
| | | \hat{Y}_{2-v} | iiMrqQg | imDdSCg | isDrqQg | IIMDQQg | lmMrsGg | |
| 1 | c_01 | 0.7022 | 8.0614 | 1.8016 | 105.65 | 0.1203 | 1.0679 | 0.7582 |
| 2 | c_02 | 0.6276 | 8.6874 | 1.7466 | 162.91 | 0.0405 | 1.3894 | 0.7374 |
| 3 | c_03 | 0.5761 | 9.1547 | 1.7073 | 172.34 | 0.0951 | 1.0866 | 0.4209 |
| 4 | c_04 | 0.2722 | 9.3193 | 1.5761 | 177.76 | 0.1023 | 1.5409 | 0.5489 |
| 5 | c_05 | 0.5774 | 9.1583 | 1.7077 | 170.76 | 0.1152 | 1.0090 | 0.4074 |
| 6 | c_06 | 0.5427 | 8.7839 | 1.7080 | 161.88 | 0.0973 | 1.0086 | 0.2890 |
| 7 | c_07 | 0.5824 | 10.395 | 1.6617 | 200.39 | 0.0914 | 1.1896 | 0.4290 |
| 8 | c_08 | 0.1158 | 14.978 | 1.2920 | 355.18 | 0.1477 | 1.7094 | 0.2610 |
| 9 | c_09 | 0.3207 | 10.553 | 1.5481 | 219.09 | 0.1204 | 1.1992 | 0.1305 |
| 10 | c_10 | 0.3314 | 10.676 | 1.5477 | 220.87 | 0.1374 | 1.2246 | 0.2599 |
| 11 | c_11 | 0.3700 | 11.074 | 1.5481 | 242.14 | 0.1060 | 1.2809 | 0.2446 |
| 12 | c_12 | 0.3575 | 10.952 | 1.5477 | 237.39 | 0.1252 | 1.2809 | 0.3284 |
| 13 | c_13 | 0.2797 | 11.136 | 1.5086 | 244.70 | 0.1483 | 1.4664 | 0.5670 |
| 14 | c_14 | 0.6869 | 9.2454 | 1.7493 | 132.88 | 0.1450 | 1.0871 | 0.7565 |
| 15 | c_15 | -0.4185 | 10.835 | 1.2336 | 261.86 | 0.1013 | 1.7094 | -0.5220 |
| 16 | c_16 | 0.4546 | 11.921 | 1.5499 | 373.43 | 0.0746 | 1.4068 | 0.6628 |
| 17 | c_17 | 1.9327 | 28.797 | 1.5011 | 961.96 | 0.0452 | 1.3894 | 1.9945 |
| 18 | c_18 | 0.3124 | 12.385 | 1.4735 | 337.21 | 0.2088 | 1.2809 | 0.7190 |
| 19 | c_19 | 0.2432 | 10.529 | 1.5172 | 252.22 | 0.1182 | 1.2809 | 0.1965 |
| 20 | c_20 | 0.7904 | 13.082 | 1.6427 | 262.52 | 0.1051 | 1.3656 | 0.8571 |
| 21 | e_01 | 0.5344 | 6.7974 | 1.7818 | 92.229 | 0.0676 | 1.2671 | 0.6493 |
| 22 | e_02 | 0.5952 | 7.4401 | 1.7818 | 149.79 | 0.0402 | 1.3894 | 0.8579 |
| 23 | e_03 | 0.6359 | 7.8693 | 1.7818 | 142.78 | 0.0780 | 1.0866 | 0.5845 |

| | | | | | | | | |
|----|------|---------|--------|--------|--------|--------|--------|---------|
| 24 | e_04 | 0.1903 | 8.4772 | 1.5752 | 167.93 | 0.0932 | 1.5409 | 0.4687 |
| 25 | e_05 | 0.6545 | 8.0666 | 1.7818 | 155.24 | 0.0878 | 1.0085 | 0.5568 |
| 26 | e_06 | 0.6370 | 7.8816 | 1.7818 | 151.04 | 0.0931 | 1.0093 | 0.5732 |
| 27 | e_07 | 0.4772 | 9.3113 | 1.6606 | 182.88 | 0.0951 | 1.1896 | 0.3912 |
| 28 | e_08 | 0.3150 | 17.054 | 1.2931 | 430.95 | 0.1147 | 1.7094 | 0.3176 |
| 29 | e_09 | 0.1396 | 8.6623 | 1.5472 | 190.86 | 0.0675 | 1.1992 | -0.2318 |
| 30 | e_10 | 0.0954 | 8.4144 | 1.5387 | 189.87 | 0.0766 | 1.2246 | -0.1887 |
| 31 | e_11 | 0.4220 | 11.646 | 1.5472 | 335.20 | 0.1100 | 1.2809 | 0.5353 |
| 32 | e_12 | 0.2555 | 10.106 | 1.5387 | 231.36 | 0.1101 | 1.2809 | 0.1914 |
| 33 | e_13 | 0.0889 | 9.1145 | 1.5088 | 216.24 | 0.0941 | 1.4664 | 0.2020 |
| 34 | e_14 | 0.6320 | 7.8282 | 1.7818 | 113.66 | 0.0735 | 1.1504 | 0.5702 |
| 35 | e_15 | -0.6248 | 8.8112 | 1.2275 | 238.79 | 0.0819 | 1.7094 | -0.7184 |
| 36 | e_16 | 0.8854 | 16.757 | 1.5389 | 392.97 | 0.0651 | 1.4068 | 0.6209 |
| 37 | e_17 | 2.0926 | 30.486 | 1.5011 | 949.42 | 0.0543 | 1.3894 | 2.0050 |
| 38 | e_18 | 0.4339 | 13.653 | 1.4741 | 442.43 | 0.1121 | 1.2809 | 0.5283 |
| 39 | e_19 | 0.2318 | 10.676 | 1.5068 | 258.51 | 0.1175 | 1.2809 | 0.1644 |
| 40 | e_20 | 0.6189 | 9.0003 | 1.7309 | 199.79 | 0.0597 | 1.3656 | 0.8391 |

Evaluarea statistică a modelelor este redată în tabelul următor:

| | StdError | $r^2(Y, desc)$ | T | 95%CI _{lower} | 95%CI _{upper} | p (%) |
|-------------------|----------|----------------|---------|------------------------|------------------------|-----------------------|
| Bi-varied model | | | | | | |
| <i>Intercept</i> | 0.5000 | | -8.8959 | -5.4610 | -3.4348 | $1.01 \cdot 10^{-8}$ |
| <i>imDdSCg</i> | 0.2836 | 0.1265 | 8.5882 | 1.8607 | 3.0099 | $2.45 \cdot 10^{-8}$ |
| <i>iiMrqQg</i> | 0.0089 | 0.3573 | 10.654 | 0.0766 | 0.1126 | $7.92 \cdot 10^{-11}$ |
| Four-varied model | | | | | | |
| <i>Intercept</i> | 0.9280 | | -10.757 | -11.869 | -8.1013 | $1.21 \cdot 10^{-10}$ |
| <i>imDdSCg</i> | 0.3613 | 0.1265 | 12.634 | 3.8308 | 5.2977 | $1.34 \cdot 10^{-12}$ |
| <i>isDrqQg</i> | 0.0002 | 0.3231 | 16.314 | 0.0026 | 0.0033 | $6.23 \cdot 10^{-16}$ |
| <i>IIMDQQg</i> | 1.0205 | 0.0568 | 5.0990 | 3.1319 | 7.2754 | $1.19 \cdot 10^{-3}$ |
| <i>lmMrsGg</i> | 0.2430 | 0.0210 | 6.1028 | 0.9897 | 1.9763 | $5.65 \cdot 10^{-5}$ |

StdError = standard error, Y = log IC⁵⁰, desc = molecular descriptor, t = parameter of the Student test

Statistica descriptivă a modelelor este redată mai jos:

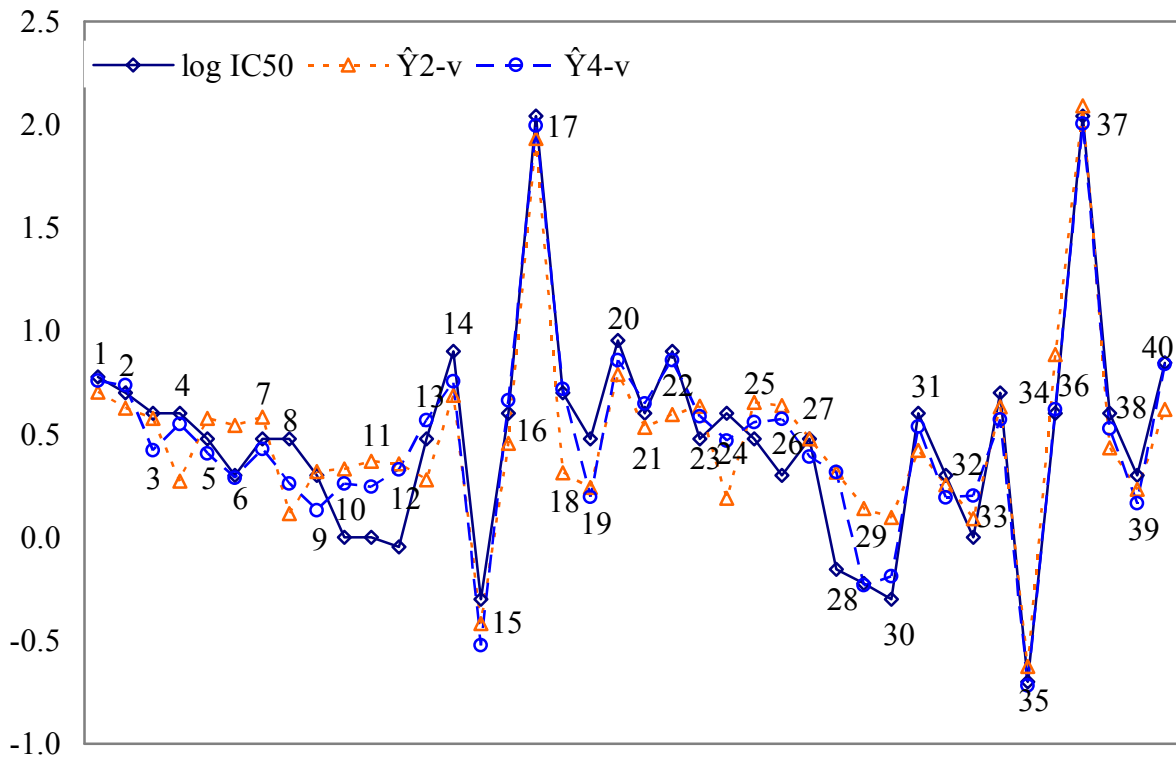
| Parameter | Value | |
|---|-----------------------|-----------------------|
| n | 40 | 40 |
| v | 2 | 4 |
| r | 0.8862 | 0.9506 |
| 95%CI _{r_lower} | 0.7937 | 0.9079 |
| 95%CI _{r_upper} | 0.9385 | 0.9737 |
| r ² | 0.7853 | 0.9037 |
| r ² _{adj} | 0.7737 | 0.8927 |
| S _{est} | 0.2477 | 0.1706 |
| F _{est} | 68 | 82 |
| p _{est} (%) | 4.4·10 ⁻¹¹ | 2.7·10 ⁻¹⁵ |
| r ² _{cv-loo} | 0.7564 | 0.8804 |
| S _{loo} | 0.2640 | 0.1902 |
| F _{pred} | 57 | 64 |
| p _{pred} (%) | 4.6·10 ⁻¹⁰ | 1.2·10 ⁻¹³ |
| r ² - r ² _{cv-loo} | 0.0289 | 0.0234 |
| r ² (<i>imDdSCg, iiMrqQg</i>) | 0.1643 | n.a. |
| r ² (<i>imDdSCg, isDrqQg</i>) | n.a. | 0.1960 |
| r ² (<i>imDdSCg, IIMDQQg</i>) | n.a. | 0.0836 |
| r ² (<i>imDdSCg, lmMrsGg</i>) | n.a. | 0.5933 |
| r ² (<i>isDrqQg, IIMDQQg</i>) | n.a. | 0.0259 |
| r ² (<i>isDrqQg, lmMrsGg</i>) | n.a. | 0.1062 |
| r ² (<i>IIMDQQg, lmMrsGg</i>) | n.a. | 0.1062 |

n = number of compounds, v = number of descriptor, n.a. = not applicable

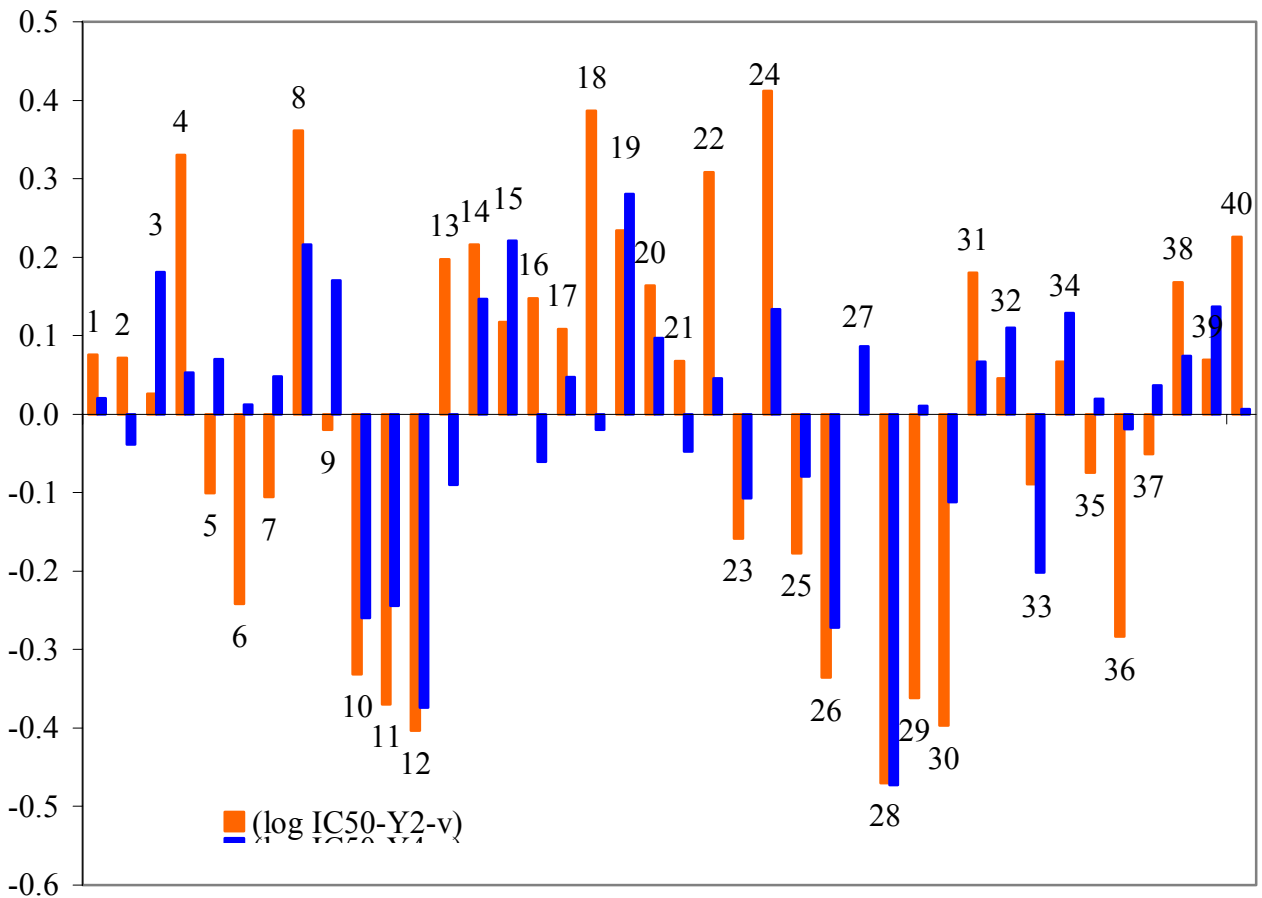
Rezultatul analizei corelațiilor corelate este redat în tabelul următor:

| Parameter | Value |
|--|---------|
| r(log IC ₅₀ - \hat{Y}_{4-v}) | 0.95064 |
| r(log IC ₅₀ - \hat{Y}_{2-v}) | 0.88617 |
| r(\hat{Y}_{four-v} - \hat{Y}_{2-v}) | 0.93439 |
| Steiger's Z | 3.17474 |
| p (%) | 0.075 |

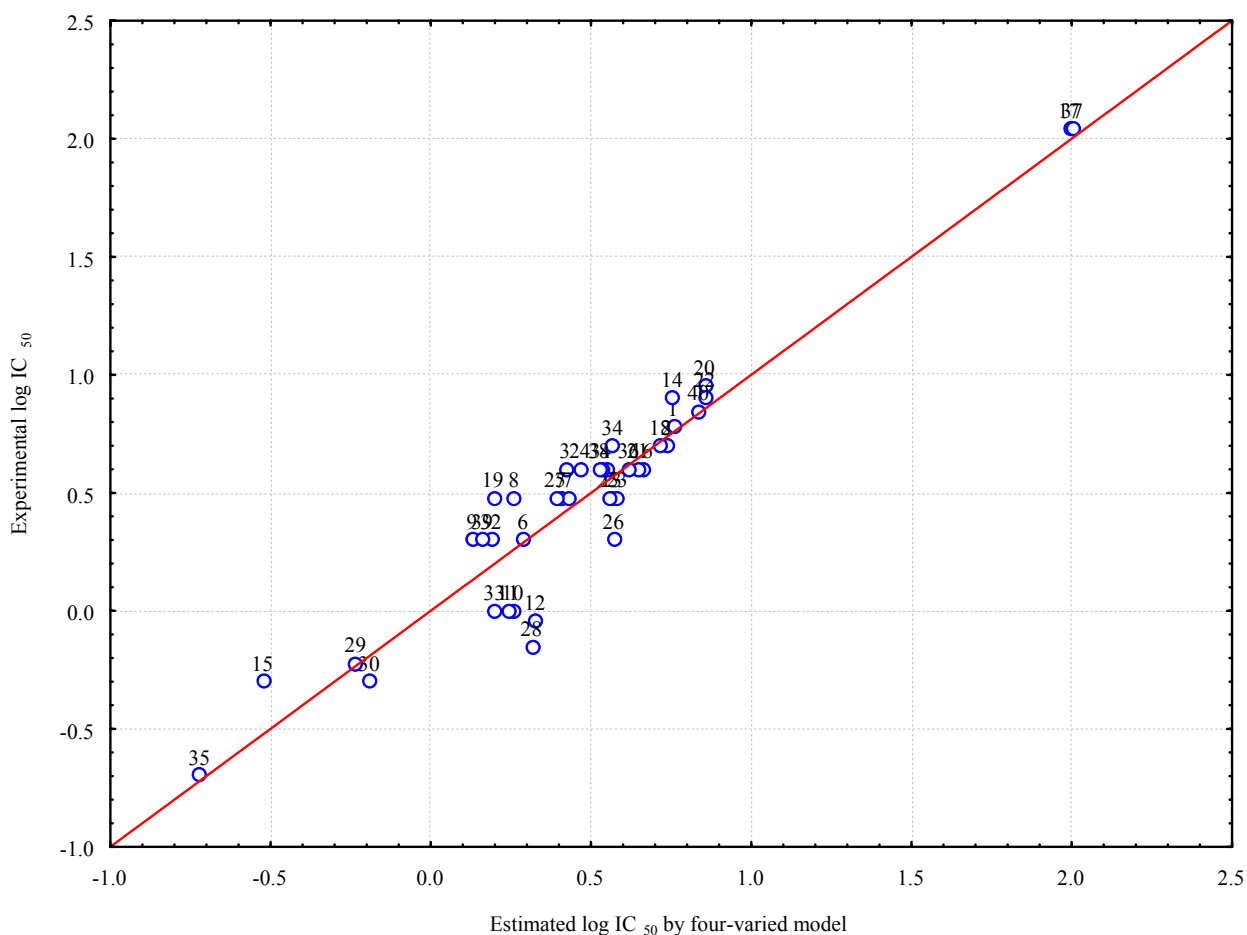
Valorile măsurate și estimate de modele MDF-SAR ale activității inhibitoare sunt redade în următorul grafic:



Distribuția reziduurilor (diferențele între măsurat și estimat) sunt redată în figura următoare:



Regresia dintre valorile măsurate și valorile prezise de modelul linear cu 4 descriptori este redată în figura următoare:



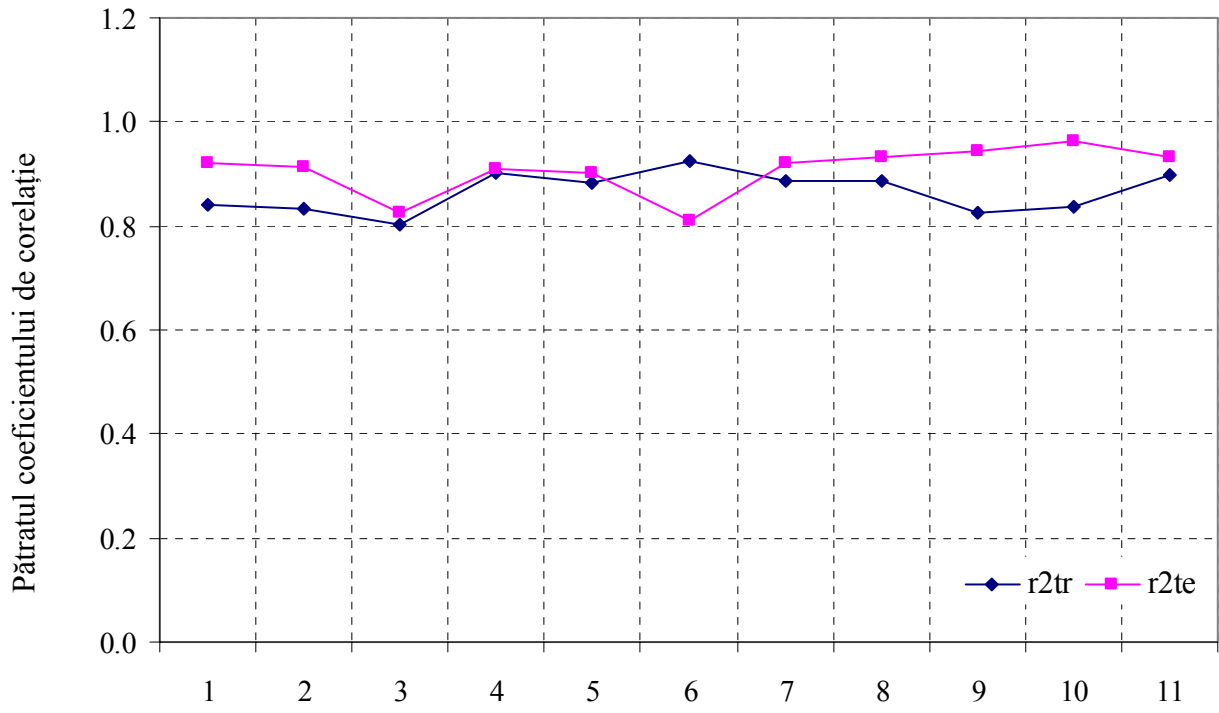
Calitatea modelelor redată prin testul TvT este sintetizată din acest experiment în tabelul de mai jos:

| No. | MDF-SAR equation | | | | | Training | | | | Test | | | | r _{tr} vs r _{ts} |
|-----|------------------|---------|-----------------------|---------|---------|------------------|-----------------|---------------------|-----------------|------------------|-----------------|----------------------|-----------------|------------------------------------|
| | Intercept | imDdSCg | isDrqQg | IIMDQQg | ImMrsGg | No _{tr} | r _{tr} | 95%CI _{tr} | F _{tr} | No _{ts} | r _{ts} | 95% CI _{ts} | F _{ts} | |
| 1 | -10.808 | 4.8727 | 3.03·10 ⁻³ | 5.6355 | 1.7079 | 20 | 0.9179 | [0.801, 0.967] | 20** | 20 | 0.9604 | [0.901, 0.984] | 40** | 1.10 [†] |
| 2 | -9.7091 | 4.4868 | 2.55·10 ⁻³ | 5.5492 | 1.3963 | 21 | 0.9137 | [0.796, 0.965] | 20** | 19 | 0.9559 | [0.887, 0.983] | 28** | 1.01 [†] |
| 3 | -8.5266 | 4.1108 | 1.87·10 ⁻³ | 3.6226 | 1.2308 | 22 | 0.8950 | [0.760, 0.956] | 17** | 18 | 0.9076 | [0.765, 0.965] | 11** | 0.20 [†] |
| 4 | -10.714 | 4.9179 | 3.03·10 ⁻³ | 5.3688 | 1.5626 | 23 | 0.9498 | [0.883, 0.979] | 41** | 17 | 0.9547 | [0.876, 0.984] | 24** | 0.15 [†] |
| 5 | -10.659 | 4.6905 | 2.97·10 ⁻³ | 5.6210 | 1.8122 | 24 | 0.9399 | [0.864, 0.974] | 36** | 16 | 0.9497 | [0.858, 0.983] | 25** | 0.26 [†] |
| 6 | -8.9966 | 4.3022 | 2.89·10 ⁻³ | 5.0608 | 1.0878 | 25 | 0.9624 | [0.915, 0.984] | 63** | 15 | 0.9011 | [0.722, 0.967] | 9* | 1.39 [†] |
| 7 | -9.2808 | 4.3025 | 2.86·10 ⁻³ | 4.1786 | 1.3324 | 26 | 0.9416 | [0.872, 0.974] | 41** | 14 | 0.9589 | [0.872, 0.987] | 17** | 0.49 [†] |
| 8 | -9.2919 | 4.2780 | 2.89·10 ⁻³ | 5.3242 | 1.3077 | 27 | 0.9406 | [0.872, 0.973] | 42** | 13 | 0.9649 | [0.884, 0.990] | 24** | 0.72 [†] |
| 9 | -9.8900 | 4.4827 | 2.42·10 ⁻³ | 5.5150 | 1.5653 | 28 | 0.9090 | [0.811, 0.957] | 27** | 12 | 0.9710 | [0.897, 0.992] | 12* | 1.51 [†] |
| 10 | -9.4715 | 4.3030 | 2.35·10 ⁻³ | 5.8316 | 1.4684 | 29 | 0.9142 | [0.824, 0.959] | 31** | 11 | 0.9820 | [0.929, 0.995] | 10* | 1.97 [†] |
| 11 | -10.194 | 4.6598 | 2.97·10 ⁻³ | 5.1893 | 1.5175 | 30 | 0.9483 | [0.893, 0.975] | 56** | 10 | 0.9662 | [0.859, 0.992] | 15* | 0.51 [†] |

* 0.001 < p < 0.01, ** p < 0.001, † p > 0.05

95% CI = 95% confidence intervals; r_{tr} = correlation coefficient – training set; r_{ts} = correlation coefficient – test set; F_{Z-test} = Fisher's Z test; No_{tr} = number of compounds in training sets; No_{ts} = number of compounds in test sets

Reprezentarea grafică a rezultatelor din experimental TvT este redată în figura de mai jos:



No. din tabelul training vs. test de mai sus (no=1 => 20 training; no=11 => 30 training)

1.6. Setul de inhibitori ai anhidrazei carbonice (2)

- ÷ Lucrare: “Structure versus Biological Role of Substituted Thiadiazole- and Thiadiazoline- Disulfonamides Modeling the Inhibition Activity on Carbonic Anhydrase I of Substituted Thiadiazole- and Thiadiazoline- Disulfonamides by Integration of Structure Information”
- ÷ Autori: Sorana-Daniela BOLBOACĂ, Lorentz JÄNTSCHI
- ÷ Referință: Computer-Aided Chemical Engineering, 24(2007), p. 965-970, 2007
- ÷ Rezumat: propune relațiilor structură-activitate în convergență cu familia descriptorilor moleculari (MDF) pentru modelarea activității biologice a 40 de inhibitori ai anhidrazei carbonice I (CA I).
- ÷ Descriere

Pornind de la ipoteza că este o relație între structura și activitatea biologică a compușilor chimici, o metodă originală numită familia de descriptori moleculari pe relații structură activitate a fost dezvoltată [Lorentz JÄNTSCHI, Molecular Descriptors Family on Structure Activity Relationships 1. Review of the Methodology, Leonardo Electronic Journal of Practices and Technologies, 4(6), p. 76-98, 2005] și aplicată cu succes pentru predicția activității inhibitoare pe CA IV [Lorentz JÄNTSCHI, Sorana BOLBOACĂ, Modelling the Inhibitory Activity on Carbonic Anhydrase IV of Substituted Thiadiazole- and Thiadiazoline- Disulfonamides: Integration of Structure Information, Electronic Journal of Biomedicine, 2006(2), p. 22-33, 2006] și CA II [Lorentz JÄNTSCHI, Mihaela Ligia UNGUREȘAN, Sorana Daniela BOLBOACĂ, Complex Structural Information Integration: Inhibitor Activity on Carbonic Anhydrase II of Substituted Disulfonamides, Applied Medical Informatics, 17(3-4), p. 12-21, 2005]. Scopul cercetării prezente a fost studiul abilităților estimative și predictive ale metodologiei MDF SAR în modelarea activității inhibitoare a anhidrazei carbonice I a unui eșantion de 40 de substituenți de 1,3,4-tiazol- și 1,3,4-tiazolin- disulfonamide.

Rezultate pentru legătura între structura și activitatea setului de inhibitori ai CA II

Modelul SAR folosind descriptori MDF obținut prin selecție genetică din familia de descriptori este:

$$\hat{Y}_{4d} = 1.14 + 8.79 \cdot 10^{-2} \cdot \text{inPRIQg} + 3.52 \cdot 10^{-3} \cdot \text{IPDMoMg} + 2.43 \cdot \text{iAMRqQg} + 1.04 \cdot \text{inMRkQt}$$

Rezultatele analizei statistice a modelului este redată în tabelul de mai jos:

| Parameter (abbreviation) | Value |
|---|-----------------|
| | (n = 40, v = 4) |
| Correlation coefficient (r) | 0.9579 |
| 95% CI for correlation coefficient (95%CI _r) | [0.9212-0.9776] |
| Squared correlation coefficient (r ²) | 0.9175 |
| Adjusted squared correlation coefficient (r ² _{adj}) | 0.9081 |
| Standard error of estimation (S _{est}) | 0.1624 |

| | |
|---|-----------------|
| Fisher parameter (F_{est}) | 97 [†] |
| Cross-validation leave-one-out score (r^2_{cv-loo}) | 0.8911 |
| Standard error of leave-one-out analysis (s_{loo}) | 0.1869 |
| Fisher parameter of loo analysis (F_{pred}) | 71 [†] |
| $r^2 - r^2_{cv-loo}$ | 0.0264 |

n = the number of compound;

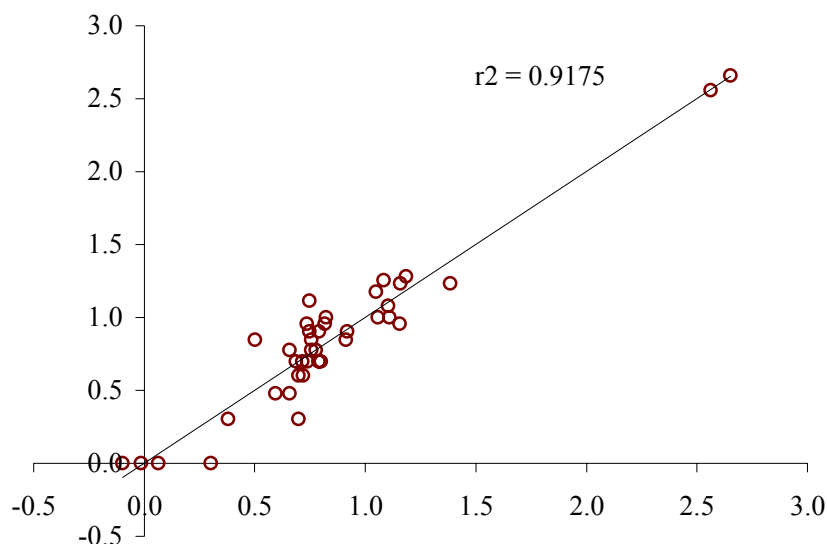
v = the number of descriptors used by the model; [†] p < 0.001

Rezultatele experimentului TvT (training versus test) aplicat modelului obținut sunt redată în tabelul de mai jos:

| n_{tr} | a_0 | a_1 | a_2 | a_3 | a_4 | r_{tr} | 95% CI _{tr} | F_{tr} | N_{ts} | r_{ts} | 95% CI _{ts} | F_{ts} | Z_{trr-ts} |
|----------|-------|----------------------|----------------------|-------|-------|----------|----------------------|------------------|----------|----------|----------------------|-----------------|-------------------|
| 20 | 1.26 | $9.92 \cdot 10^{-2}$ | $3.65 \cdot 10^{-3}$ | 2.21 | 1.15 | 0.94 | [0.842-0.975] | 27 [‡] | 20 | 0.97 | [0.929, 0.989] | 60 [‡] | 1.23 |
| 21 | 1.08 | $9.01 \cdot 10^{-2}$ | $3.58 \cdot 10^{-3}$ | 2.72 | 1.15 | 0.96 | [0.905-0.984] | 49 [‡] | 19 | 0.95 | [0.881, 0.982] | 34 [‡] | 0.27 |
| 22 | 0.90 | $8.01 \cdot 10^{-2}$ | $2.91 \cdot 10^{-3}$ | 2.53 | 0.61 | 0.98 | [0.961-0.993] | 129 [‡] | 18 | 0.90 | [0.750, 0.963] | 14 [‡] | 2.69 [†] |
| 23 | 1.07 | $8.63 \cdot 10^{-2}$ | $3.41 \cdot 10^{-3}$ | 2.46 | 0.83 | 0.95 | [0.873-0.977] | 38 [‡] | 17 | 0.97 | [0.902, 0.987] | 38 [‡] | 0.65 |
| 24 | 0.69 | $7.51 \cdot 10^{-2}$ | $2.87 \cdot 10^{-3}$ | 3.32 | 1.26 | 0.93 | [0.851-0.971] | 33 [‡] | 16 | 0.94 | [0.837, 0.980] | 18 [‡] | 0.18 |
| 25 | 1.46 | $9.16 \cdot 10^{-2}$ | $4.46 \cdot 10^{-3}$ | 2.10 | 0.45 | 0.89 | [0.766-0.951] | 19 [‡] | 15 | 0.96 | [0.881, 0.988] | 6 [†] | 1.53 |
| 26 | 0.91 | $1.06 \cdot 10^{-1}$ | $3.09 \cdot 10^{-3}$ | 2.84 | 0.85 | 0.95 | [0.880-0.975] | 44 [‡] | 14 | 0.95 | [0.848, 0.985] | 18 [‡] | 0.15 |
| 27 | 1.19 | $8.45 \cdot 10^{-2}$ | $3.52 \cdot 10^{-3}$ | 2.21 | 0.98 | 0.95 | [0.882-0.975] | 46 [‡] | 13 | 0.97 | [0.905, 0.992] | 32 [‡] | 0.88 |
| 28 | 1.17 | $8.88 \cdot 10^{-2}$ | $3.50 \cdot 10^{-3}$ | 2.33 | 1.02 | 0.93 | [0.860-0.969] | 39 [‡] | 12 | 0.99 | [0.957, 0.997] | 70 [‡] | 2.24 [†] |
| 29 | 1.09 | $9.77 \cdot 10^{-2}$ | $3.61 \cdot 10^{-3}$ | 2.69 | 1.16 | 0.96 | [0.911-0.982] | 67 [‡] | 11 | 0.97 | [0.872, 0.991] | 13 [†] | 0.28 |
| 30 | 1.09 | $9.07 \cdot 10^{-2}$ | $3.05 \cdot 10^{-3}$ | 2.17 | 0.98 | 0.92 | [0.830-0.960] | 33 [‡] | 10 | 0.98 | [0.897, 0.994] | 15 [†] | 1.49 |
| 31 | 1.21 | $8.66 \cdot 10^{-2}$ | $3.59 \cdot 10^{-3}$ | 2.22 | 1.01 | 0.95 | [0.887-0.973] | 55 [‡] | 9 | 0.98 | [0.908, 0.996] | 21 [†] | 1.18 |

a_0 =intercept; a_1 =inPRlQg; a_2 =lPDMoMg; a_3 =iAMRqQg; a_4 =inMRkQt; [‡]p ≤ 0.001; [†]0.001 < p < 0.05

Reprezentarea grafică a modelului liniar obținut este redată în figura de mai jos:



1.7. Setul de taxoizi (1)

- ÷ Lucrare: “3D QSAR Analysis of Taxoids from Taxus Cuspidata var. Nana by Comparative Molecular Field Approach”
- ÷ Autori: Hiroshi MORITA, Akira GONDA, Lan WEI, Koichi TAKEYA, Hideji ITOKAWA
- ÷ Referință: Bioorganic & Medicinal Chemistry Letters, 7(18), p. 2387-2392, 1997
- ÷ Rezumat: propune relațiilor structură-activitate în convergență cu metoda analizei comparative a câmpului molecular (CoMFA) pentru modelarea activității biologice a 35 taxoizi citotoxici.
- ÷ Descriere

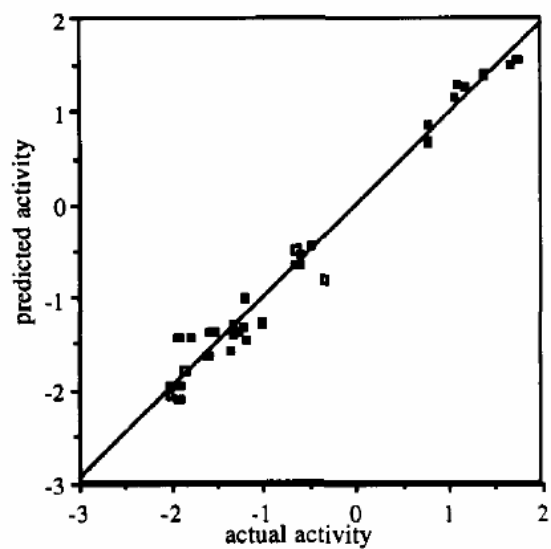
CoMFA a fost inițiat utilizând conformațiile de energie minimă. Configurațiile implicite SYBYL au fost utilizate în acest sens dacă nu a fost specificat explicit altceva. Compușii au fost aliniați după fitarea radical din suma pătratelor distanțelor între atomii taxolului utilizat ca moleculă template. Sarcinile atomice parțiale au fost obținute utilizând metoda PM3 [Optimization of parameters for semiempirical methods II. Applications, Stewart, J.J.P., Journal of Computational Chemistry, 1989, 10(2), p. 221-264, 1989]. Modelul 3D-QSAR de calibrare astfel obținut a fost folosit pentru predicția activității inhibitoare a unor taxoizi grupați în setul test după ce analiza conformațională și aliniamentele au fost efectuate prin aceeași metodă ca pentru setul de învățare.

Rezultate pentru legătura între structura și activitatea setului de inhibitori ai CA II

Rezultatul analizei statistice aplicate modelului obținut este redată în tabelul de mai jos:

| | |
|--------------------------------|--------------------------|
| opt. no. of components | 5 |
| probe atom | c (sp ³ , +1) |
| cross-validated r ² | 0.818 |
| std error of estimate | 0.196 |
| r ² | 0.979 |
| F values (n1=5, n2=29) | 267.621 |
| contributions | |
| steric | 0.520 |
| electrostatic | 0.480 |

Regresia liniară între activitatea măsurată și cea estimată de modelul CoMFA este ilustrată mai jos:

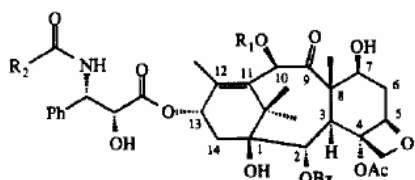


Valorile măsurate și estimate cu modelul CoMFA pentru compușii incluși în setul de învățare sunt redată în tabelul de mai jos:

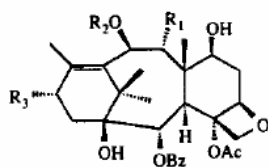
| compounds | log 1/IC ₅₀ | | |
|-----------|------------------------|-----------|----------|
| | actual | predicted | residual |
| 1 | 1.66 | 1.48 | 0.18 |
| 2 | 1.37 | 1.41 | -0.04 |
| 3 | 0.77 | 0.84 | -0.09 |
| 4 | 1.18 | 1.26 | -0.08 |
| 5 | 1.09 | 1.27 | -0.19 |
| 6 | 1.05 | 1.14 | -0.09 |
| 7 | 1.39 | 1.38 | 0.01 |
| 8 | 1.74 | 1.54 | 0.20 |
| 9 | 0.77 | 0.66 | 0.11 |
| 10 | -1.20 | -1.02 | -0.18 |
| 11 | -1.28 | -1.38 | 0.10 |
| 12 | -1.00 | -1.27 | 0.27 |
| 13 | -1.54 | -1.39 | -0.15 |
| 14 | -1.32 | -1.29 | -0.04 |
| 15 | -1.60 | -1.63 | 0.03 |
| 16 | -0.34 | -0.80 | 0.46 |
| 17 | -0.64 | -0.49 | -0.15 |
| 18 | -2.00 | -2.05 | 0.05 |
| 19 | -1.78 | -1.42 | -0.36 |
| 20 | -0.62 | -0.64 | 0.01 |
| 21 | -1.20 | -1.31 | 0.11 |
| 22 | -0.48 | -0.46 | -0.02 |
| 23 | -1.36 | -1.56 | 0.20 |
| 24 | -2.00 | -1.96 | -0.04 |
| 25 | -1.90 | -2.09 | 0.19 |
| 26 | -1.91 | -1.42 | -0.49 |
| 27 | -1.18 | -1.46 | 0.28 |
| 28 | -0.59 | -0.53 | -0.06 |
| 29 | -1.85 | -1.79 | -0.06 |
| 30 | -1.91 | -1.94 | 0.03 |
| 31 | -1.57 | -1.39 | -0.18 |
| 32 | -2.00 | -1.95 | -0.05 |
| 33 | -0.64 | -0.65 | 0.01 |
| 34 | -2.00 | -1.97 | -0.03 |
| 35 | -1.32 | -1.40 | 0.08 |

Structura taxoizilor

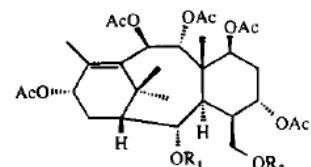
Formulele structurale ale taxoizilor incluși în studiu sunt redată în figura următoare:



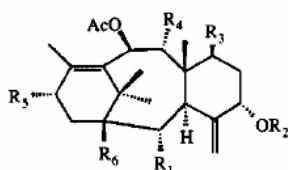
- 1 $R_1=Ac, R_2=Ph$ (taxol)
- 2 $R_1=H, R_2=Ph$
- 3 $R_1=Ac, R_2=C_4H_7$
- 4 $R_1=H, R_2=C_4H_7$
- 5 $R_1=Ac, R_2=n-C_5H_{11}$
- 6 $R_1=H, R_2=n-C_5H_{11}$ C-7 epimer
- 7 $R_1=H, R_2=n-C_5H_{11}$
- 8 $R_1=Ac, R_2=C_6H_{13}$
- 9 $R_1=Ac, R_2=n-C_3H_7$



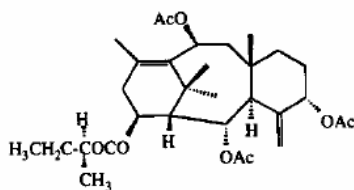
- 10 $R_1=O, R_2=Ac, R_3=OH$
- 11 $R_1=OH, R_2=Ac, R_3=OAc$
- 12 $R_1=OH, R_2=H, R_3=O$
- 13 $R_1=OH, R_2=H, R_3=OAc$
- 14 $R_1=OH, R_2=Ac, R_3=O$



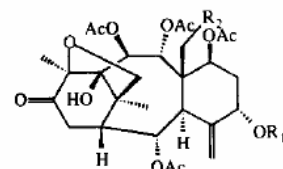
- 29 $R_1=Ac, R_2=H$
- 30 $R_1=H, R_2=Ac$



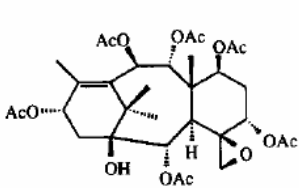
- 15 $R_1=H, R_2=H, R_3=OBz, R_4=OAc, R_5=OH, R_6=OH$
- 16 $R_1=H, R_2=Ac, R_3=H, R_4=OAc, R_5=OAc, R_6=H$
- 17 $R_1=H, R_2=COCH=CHPh, R_3=OAc, R_4=OAc, R_5=OAc, R_6=H$
- 18 $R_1=OAc, R_2=COCH=CHPh, R_3=H, R_4=OAc, R_5=O, R_6=H$
- 19 $R_1=OAc, R_2=Ac, R_3=H, R_4=OAc, R_5=OAc, R_6=H$
- 20 $R_1=H, R_2=Ac, R_3=OAc, R_4=OAc, R_5=OAc, R_6=H$
- 21 $R_1=H, R_2=H, R_3=OAc, R_4=OAc, R_5=OH, R_6=H$
- 22 $R_1=OCOC_4H_9, R_2=Ac, R_3=OAc, R_4=H, R_5=H, R_6=H$
- 23 $R_1=H, R_2=H, R_3=OAc, R_4=OAc, R_5=OAc, R_6=H$
- 24 $R_1=OAc, R_2=COCH=CHPh, R_3=OAc, R_4=OAc, R_5=O, R_6=H$
- 25 $R_1=OAc, R_2=H, R_3=OAc, R_4=OAc, R_5=OAc, R_6=H$
- 26 $R_1=H, R_2=H, R_3=OAc, R_4=OAc, R_5=O, R_6=H$
- 27 $R_1=OAc, R_2=H, R_3=OAc, R_4=OAc, R_5=O, R_6=H$
- 28 $R_1=OH, R_2=COCH=CHPh, R_3=OAc, R_4=OAc, R_5=OAc, R_6=H$



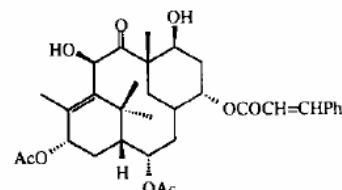
33



- 31 $R_1=H, R_2=OBz$
- 32 $R_1=COCH=CHPh, R_2=H$



34



35

1.8. Setul de taxoizi (2)

- ÷ Lucrare: "Synthesis and structure-activity relationships of taxuyunnanine C derivatives as multidrug resistance modulator in MDR cancer cells"
- ÷ Autori: Toshiaki HASEGAWA, Jiao BAI, Jungui DAI, Liming BAI, Junichi SAKAI, Shigenori NISHIZAWA, Yuhua BAI, Midori KIKUCHI, Mariko ABE, Takao YAMORI, Akihiro TOMIDA, Takashi TSURUO, Katsutoshi HIROSEE, Masayoshi ANDO
- ÷ Referință: Bioorganic & Medicinal Chemistry Letters, 17(13), p. 3722-3728, 2007
- ÷ Rezumat: O serie de taxoizi de nouă generație comportând grupuri voluminoase în poziții diferite cum sunt C-2, C-5, C-7, C-9, C-10 sau C-14 au fost obținuți prin modificări chimice și biotransformări ale taxuiunaninei C și analogilor săi 4, 5 și 10. Parte din seria de noi compuși obținuți (compușii 3, 5 și 9a) au prezentat activitate inhibitoare celulară pe VA-13 și/sau HepG2 și de asemenea activitate de acumulare pe calceină și/sau vistirină în MDR 2780AD și se așteaptă că aceștia vor fi compuși lider pentru un nou tip de agenți împotriva cancerului.

Structura și activitatea taxoizilor

Figura următoare redă structura grupărilor aciloxi- în structurile derivaților taxilunaninei C:

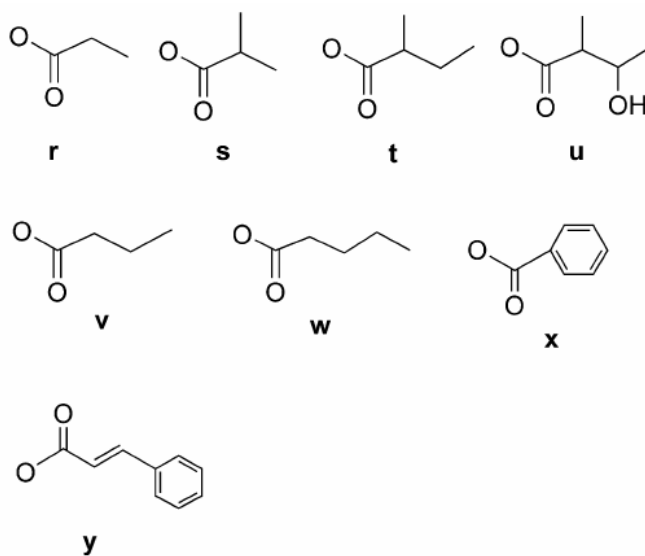


Figura următoare redă taxilunanina C și analogii săi:

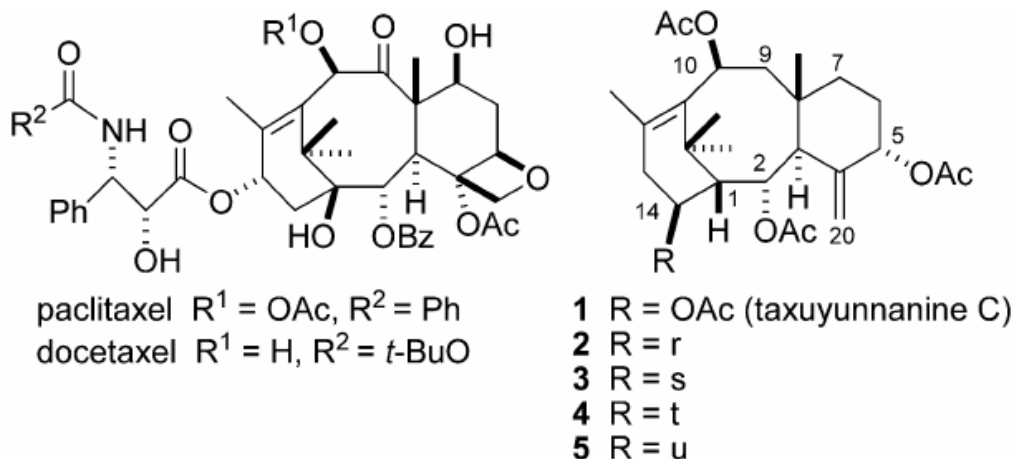
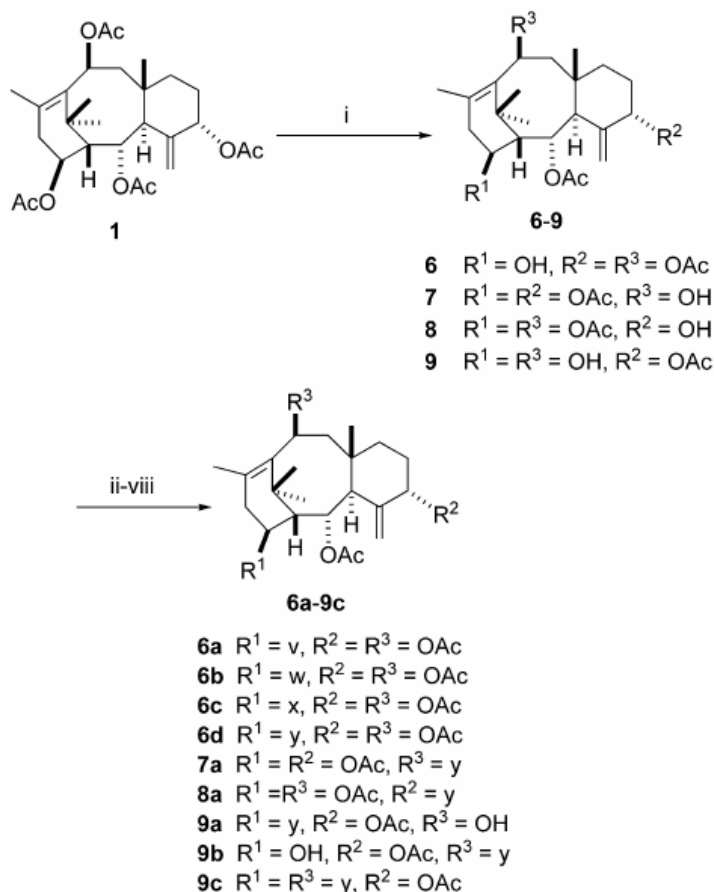


Figura următoare redă prima schemă de sinteză utilizată:



Scheme 1. Reagents and conditions: (i) **1**, 1 M K_2CO_3 (5 equiv), 1:3 THF/MeOH, 45 °C, 7 h (23.9% for **6**, 4.3% for **7**, 0.5% for **8**, 20.7% for **9**); (ii) **6**, $CH_3(CH_2)_2COCl$ (13 equiv), DMAP, dry Py, 80 °C, 12.5 h (87.6% for **6a**); (iii) **6**, $CH_3(CH_2)_3COCl$ (13 equiv), DMAP, dry Py, 80 °C, 13.5 h (88.0% for **6b**); (iv) **6**, $PhCOCl$ (10 equiv), DMAP, dry Py, 80 °C, 15 h (92.7% for **6c**); (v) **6**, $PhCH=CHCOCl$ (8 equiv), DMAP, dry Py, 85 °C, 16 h (92.8% for **6d**); (vi) **7**, $PhCH=CHCOCl$ (8 equiv), DMAP, dry Py, 85 °C, 13 h (91.5% for **7a**); (vii) **8**, $PhCH=CHCOCl$ (20 equiv), DMAP, dry Py, 85 °C, 19h (80% for **8a**); (viii) **9**, $PhCH=CHCOCl$ (2.2 equiv), DMAP, dry Py, 85–90 °C, 23 h (8.4% for **9a**, 29.1% for **9b**, 5.9% for **9c**).

Figura următoare redă a doua schemă de sinteză folosită:

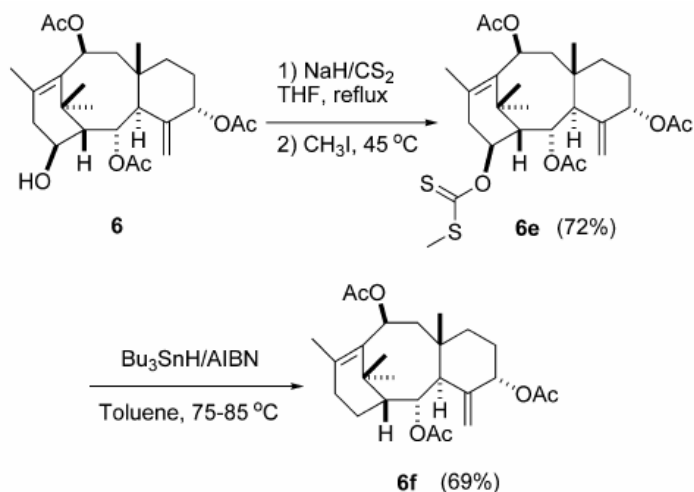


Figura următoare redă a treia schemă de sinteză folosită:

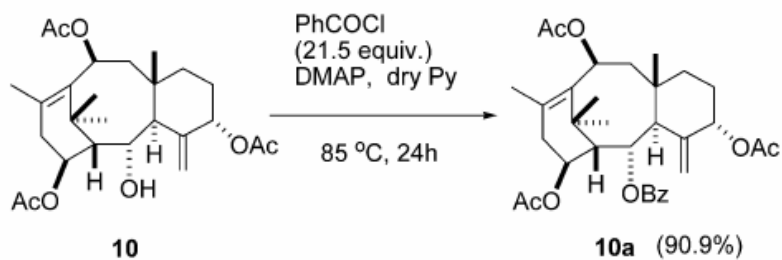
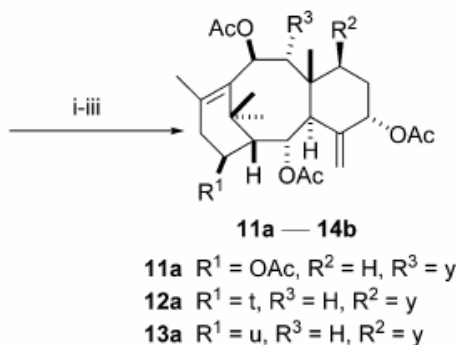
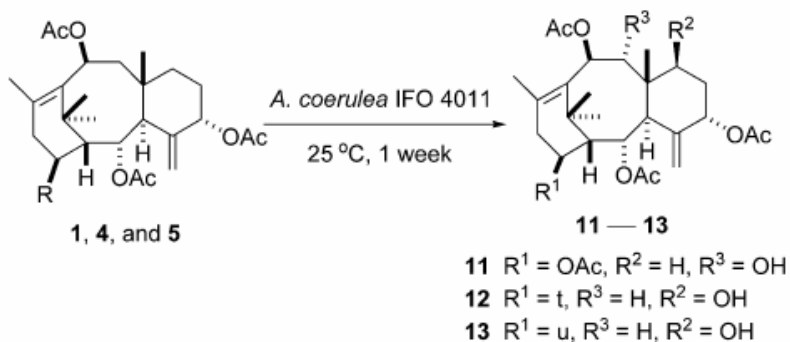
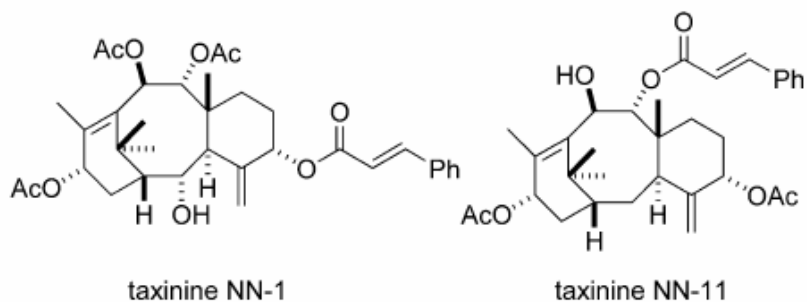


Figura următoare redă a 4-a schemă de sinteză folosită:



Scheme 4. Reagents and conditions: (i) **11**, PhCH=CHCOCl (10 equiv), DMAP, dry Py, 85–90 °C, 9 h (84.6% for **11a**); (ii) **12**, PhCH=CHCOCl (10 equiv), DMAP, dry Py, 85–90 °C, 10 h (88.8% for **12a**); (iii) **13**, PhCH=CHCOCl (6 equiv), DMAP, dry Py, 85–90 °C, 5.5 h (13.9% for **13a**).

Figura următoare redă 2 derivați de taxinină:



Rezultate pentru noii taxoizi obținuți

Tabelul următor conține efectul compușilor în acumularea calceinei în celule rezistente la produși farmaceutici 2780AD și activitatea inhibitoare pe celulele WI-38, VA-13, și HepG2

| Compound | Calcein accumulation (% of control) ^{a,b} | | |
|------------|--|-----------|----------|
| | 0.25 µg/ml | 2.5 µg/ml | 25 µg/ml |
| Paclitaxel | | | |
| Adriamycin | | | |
| 1 | 130 | 141 | 127 |
| 2 | 104 | 98 | 119 |
| 3 | 133 | 139 | 148 |
| 4 | 100 | 86 | 119 |
| 5 | 81 | 94 | 168 |
| 6 | 104 | 86 | 148 |
| 6a | 97 | 98 | 114 |
| 6b | 94 | 88 | 111 |
| 6c | 90 | 107 | 133 |
| 6d | 123 | 126 | 125 |
| 6f | 89 | 105 | 91 |
| 7 | 103 | 98 | 111 |
| 7a | 92 | 123 | 126 |
| 8 | 118 | 118 | 141 |
| 8a | 102 | 95 | 121 |
| 9a | 107 | 132 | 190 |
| 9b | 109 | 113 | 113 |
| 9c | 83 | 80 | 80 |
| 10 | 99 | 99 | 111 |
| 10a | 98 | 96 | 125 |
| 11 | 102 | 101 | 110 |
| 11a | 102 | 103 | 125 |
| 12 | 94 | 98 | 110 |
| 12a | 96 | 98 | 102 |
| 13 | 87 | 104 | 107 |
| 13a | 97 | 102 | 130 |

| Compound | Cell growth inhibitory activity (IC ₅₀ µM) ^c | | |
|------------|--|-------|-------|
| | WI-38 | VA-13 | HepG2 |
| Paclitaxel | 0.04 | 0.005 | 8.1 |
| Adriamycin | 0.66 | 0.38 | 1.2 |
| 1 | 130 | 53.1 | 18.5 |
| 2 | >193 | 93.7 | 86.4 |
| 3 | >188 | 90.9 | 28.0 |
| 4 | >183 | 61.1 | 41.7 |
| 5 | 79.4 | 9.29 | 18.5 |
| 6 | 133 | 80.0 | 216 |
| 6a | 73.4 | 17.2 | 18.5 |
| 6b | >183 | 56.7 | 16.6 |
| 6c | >176 | >176 | 144 |
| 6d | 154 | >169 | >169 |
| 6f | 8.93 | 12.2 | 10.3 |
| 7 | 105 | 12.6 | 108 |
| 7a | >169 | >169 | 106 |
| 8 | 149 | 105 | 183 |
| 8a | 99.0 | 125 | 9.95 |
| 9a | 13.5 | 37.4 | 15.7 |
| 9b | 71.7 | 11.5 | 15.3 |
| 9c | >147 | >147 | >147 |
| 10 | 130.4 | 104 | 122 |
| 10a | 125 | 175 | 51.9 |
| 11 | 129 | >192 | 128 |
| 11a | 151 | >154 | >154 |
| 12 | 81.6 | 79.1 | 75.2 |
| 12a | >144 | >144 | >144 |
| 13 | 140 | >173 | 153 |
| 13a | 62.4 | 49.9 | 80.4 |

Legenda tabelului de mai sus este după cum urmează:

a The amount of calcein accumulated in multidrug-resistant human ovarian cancer 2780AD cells was determined with the control in the presence of 0.25, 2.5 and 25 lg/ml of test compounds.

b The values are the relative amount of calcein accumulated in the cell compared with the control experiment. The values represent the mean of triplicate determination.

c IC₅₀ represents the mean of duplicate determination.

Tabela următoare prezintă efectul compușilor în acumularea vincristinei (VCR) în celulele rezistente la medicamente de tipul 2780AD:

| | Concd (μg/mL) | Average ^b (dpm/well) | % of control ^c | Activities ^d | Verapamil ^e (%) |
|-----------|---------------|---------------------------------|---------------------------|-------------------------|----------------------------|
| 6 | 0.1 | 360 | 102 | ± | 94 |
| | 1 | 517 | 147 | + | 79 |
| | 10 | 853 | 242 | + | 53 |
| 6d | 0.1 | 336 | 95 | ± | 87 |
| | 1 | 494 | 140 | + | 75 |
| | 10 | 985 | 280 | + | 62 |
| 7 | 0.1 | 324 | 92 | ± | 84 |
| | 1 | 420 | 119 | + | 64 |
| | 10 | 805 | 229 | + | 51 |
| 7a | 0.1 | 404 | 132 | + | 130 |
| | 1 | 641 | 209 | + | 118 |
| | 10 | 1385 | 451 | ++ | 81 |
| 9a | 0.1 | 425 | 121 | + | 111 |
| | 1 | 959 | 272 | + | 146 |
| | 10 | 1855 | 527 | +++ | 116 |
| 9b | 0.1 | 402 | 114 | + | 105 |
| | 1 | 669 | 190 | + | 102 |
| | 10 | 1749 | 497 | ++ | 110 |
| 9c | 0.1 | 320 | 91 | ± | 83 |
| | 1 | 325 | 92 | ± | 49 |
| | 10 | 429 | 122 | + | 27 |
| 10 | 0.1 | 306 | 101 | ± | 89 |
| | 1 | 436 | 144 | + | 84 |
| | 10 | 786 | 259 | + | 74 |
| 11 | 0.1 | 303 | 86 | — | 79 |
| | 1 | 379 | 108 | ± | 58 |
| | 10 | 592 | 168 | + | 37 |

Legendă:

^aThe amount of VCR accumulated in MDR2780AD cells was determined with the control in the presence of 0.1, 1, and 10 μg/ml of taxoids.

^bThe values represents triplicated determinations.

^cThe values are the relative amount of VCR accumulated in the cell compared with the control experiment.

^dThe indices are expressed on a scale of seven by the range of the relative amount of VCR accumulation as compared with the control experiment(%): +++, 501–1000%; ++, 301–500%; +, 111–300%; ±, 91–100%; —, <90%.

^eThe values are expressed as the relative amount of vincristine(VCR) accumulation in the cell as compared with that of verapamil.

Tabela următoare prezintă efectul compușilor pe citotoxicitatea agenților împotriva cancerului raportat la celulele AD2780 și MDR 2780AD:

| Cell lines | MDR modulator (μM) or (μM) | Viability of the cell (%) ^a | IC ₅₀ (nM) of anticancer agents ^b | | | | |
|------------|--|--|---|-----|-----|-----|-----|
| | | | Taxol | ADM | VCR | | |
| A2780 | No modulator | 0 | 100 | 2.8 | 5.7 | 3.1 | |
| | Verapamil | 0.2 | 93 | 2.7 | 5.8 | 3.4 | |
| | | 2 | 82 | 2.9 | 6.0 | 2.0 | |
| | | 10 | 53 | 3.5 | 7.7 | 0.5 | |
| | 6 | 0.2 | 97 | 2.6 | 7.0 | 2.6 | |
| | | 2 | 95 | 2.3 | 6.0 | 1.6 | |
| | | 10 | 20 | 6.6 | 7.8 | 4.8 | |
| | 9a | 0.2 | 89 | 2.6 | 7.0 | 2.6 | |
| | | 2 | 74 | 2.3 | 6.0 | 1.6 | |
| | | 10 | 27 | 6.6 | 7.8 | 4.8 | |
| | 2780AD | No modulator | 0 | 100 | 183 | 298 | 311 |
| | | Verapamil | 0.2 | 84 | 36 | 269 | 41 |
| 2 | | | 90 | 5.5 | 201 | 25 | |
| 10 | | | 71 | 4.9 | 164 | 3.9 | |
| 6 | | 0.2 | 101 | 21 | 215 | 55 | |
| | | 2 | 96 | 26 | 256 | 26 | |
| | | 10 | 77 | 5.3 | 181 | 11 | |
| 9a | | 0.2 | 105 | 22 | 284 | 51 | |
| | | 2 | 86 | 4.2 | 138 | 39 | |
| | | 10 | 38 | 3.9 | 73 | 23 | |

Legendă:

^a Cytotoxicity of the compounds was evaluated in the absence of the anticancer agents.

^b Enhancing effects of the compounds on the cytotoxicity of taxol, adriamycin (ADM), and vincristin (VCR) toward A2780 cells and MDR A2780 (2780AD) cells were determined in the presence of 0.2, 2.0, and 10 mM of each compound. The values represent means of triplicate determination.

Sumarul evaluării compusului 7a bazat pe 39 de linii de celule cancerigene umane este redat în tabelul următor:

| | GI ₅₀ | TGI | LC ₅₀ |
|---|------------------|----------------|-----------------------------|
| <i>Parameters of effective concentrations</i> | | | |
| MG-MID | -4.41 | -4.04 | -4.00 |
| Rank | Compound | r ^a | Molecular targets/drug type |
| <i>Results of the COMPARE analysis</i> | | | |
| 1 | ICRF-154 | 0.581 | Topoisomerase |
| 2 | ICRF-193 | 0.580 | Topoisomerase |
| 3 | TNP-470 | 0.552 | |

1.9. Setul de taxoizi (3)

- ÷ Lucrare: "Antitumor agents. 258. Syntheses and evaluation of dietary antioxidant-taxoid conjugates as novel cytotoxic agents"
- ÷ Autori: Kyoko NAKAGAWA-GOTO, Koji YAMADA, Seikou NAKAMURA, Tzu-Hsuan CHEN, Po-Cheng CHIANG, Kenneth F. BASTOW, Shao-Chun WANG, Bill SPOHN, Mien-Chie HUNG, Fang-Yu LEE, Fang-Chen LEE, Kuo-Hsiung LEE
- ÷ Referință: Bioorganic & Medicinal Chemistry Letters, 17(18), p. 5204-5209, 2007
- ÷ Rezumat: O gamă variată de antioxidanți destinați consumului uman, incluzând vitamine, flavonoide, curcumin, și o cumarină au fost conjugați cu paclitaxel printr-o legătură esterică. Nii compuși sintetizați au fost evaluați pentru activitatea citotoxică pe câteva linii de celule tumorale umane și pe liniile corespunzătoare de celule normale. Interesant, cei mai testați conjugați sunt selectivi la inhibiția creșterii la celulele tumorale 1A9 (ovarian) și KB (nasofaringial) și fără a avea activitate pe alte linii celulare. Specific, conjugații 16 și 20 sunt foarte activi pe celulele 1A9 (ED50 = 0.005 μg/mL) și pe KB (ED50 = 0.005 și respectiv 0.14 μg/mL). Compusul 22b, sarea esterului glicinat al vitaminei E conjugat cu 1 pare a fi un lider pentru dezvoltările ulterioare de candidați de trialuri clinice atâta timp cât dovedește puternică activitate inhibitoare împotriva Panc-1 (cancerul pancreatic) cu efect mic asupra corespondentului E6E7 (normal) al liniei celulare.

Structura și activitatea taxoizilor

Figura următoare redă structura paraclitaxelului și partenerilor săi conjugați antioxidanți:

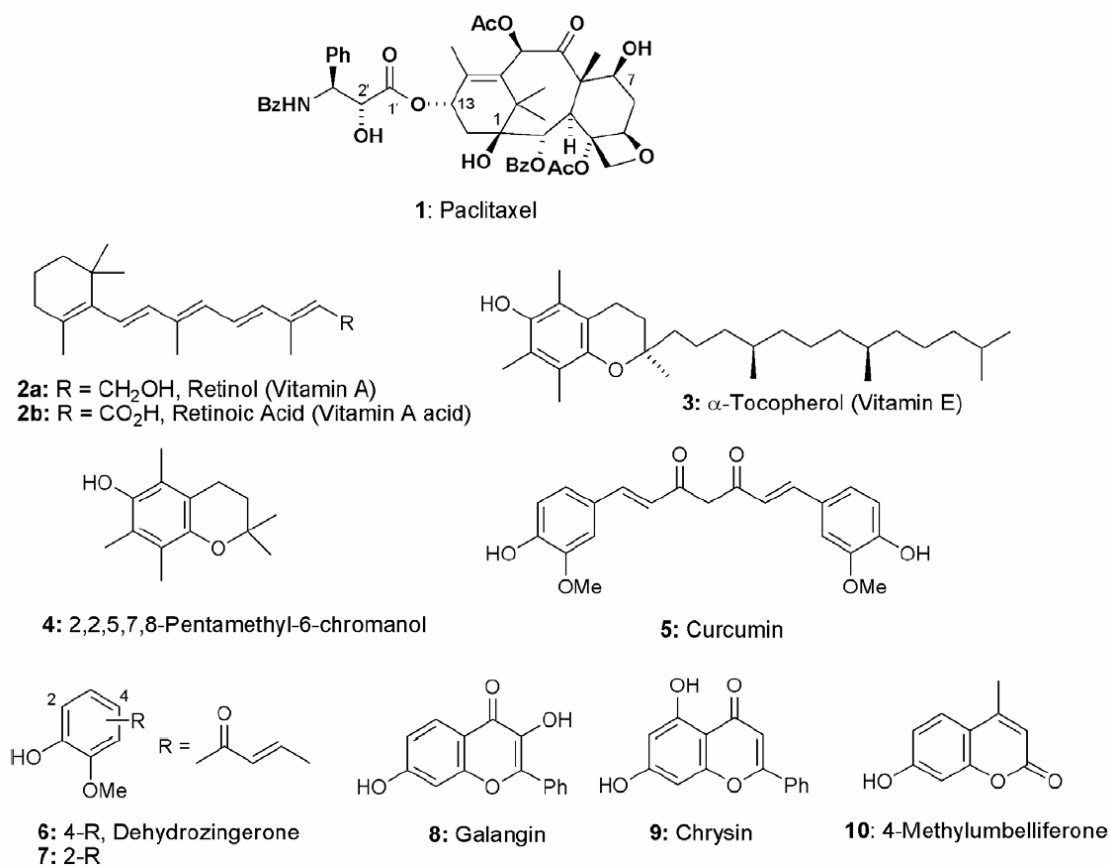
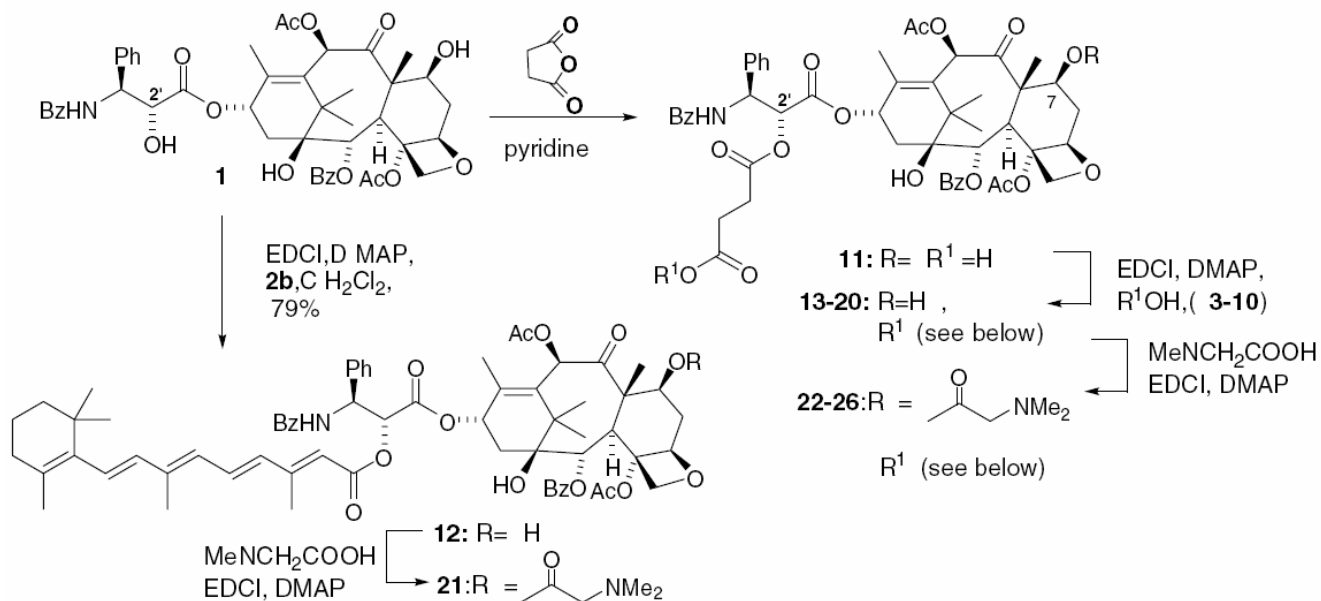
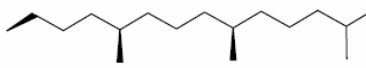
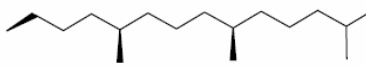
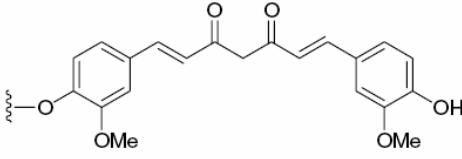
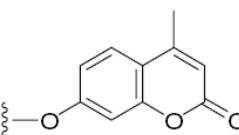


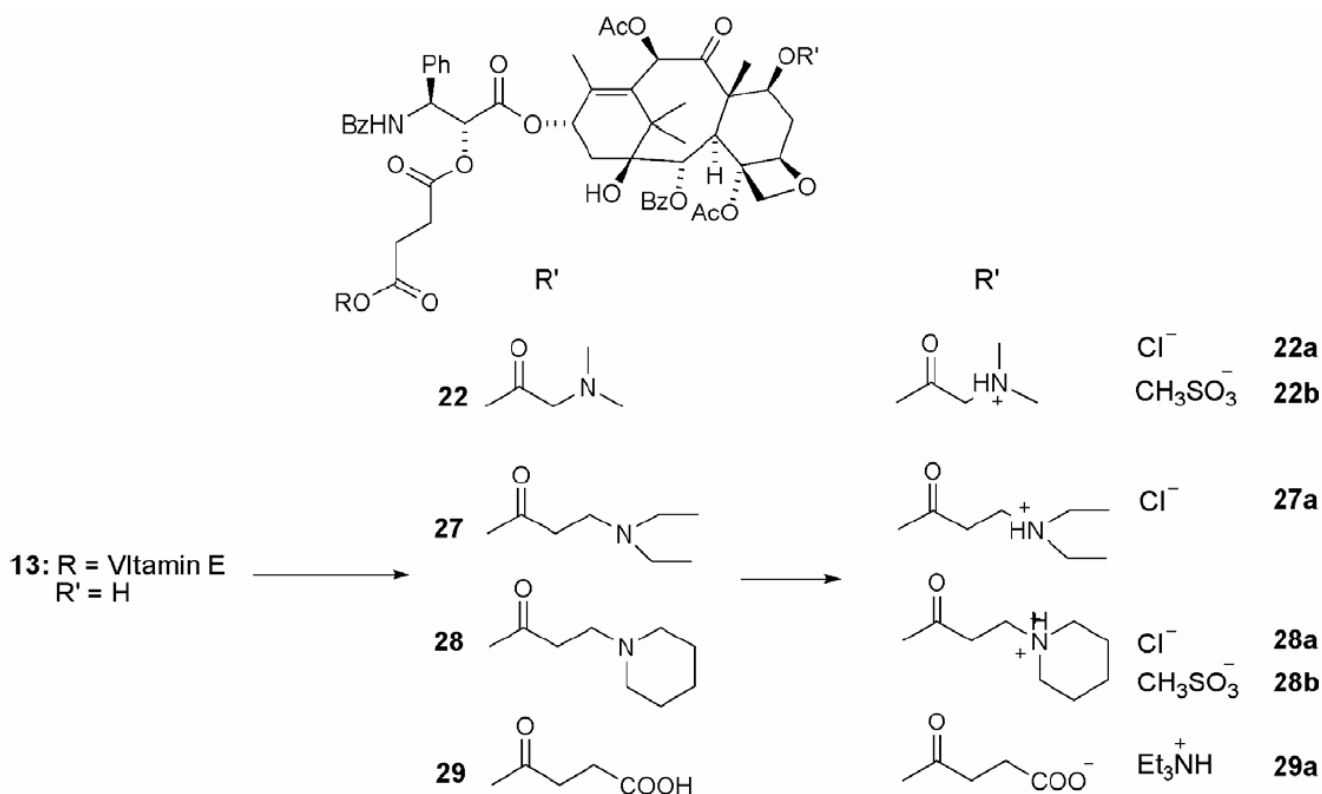
Figura următoare redă schema de sinteză a conjugățiilor 12-26:



Semnificațiile radicalilor din figura de mai sus sunt în tabelul de mai jos:

| Conjugate (R = -H) | Glycinate Ester (R = -COCH ₂ NMe ₂) | R ¹ |
|-----------------------|---|--|
| 13 | 22 |  R ² =  R ² = Me |
| 14 | 23 | |
| 15 | |  |
| 16 | 24 | R ² = 4-CH=CHCOCH ₃ |
| 17 | 25 | R ² = 2-CH=CHCOCH ₃ |
| 18 | 26 | R ² = OH, R ³ = H |
| 19 | | R ² = H, R ³ = OH |
| 20 | |  |

Schema din figura următoare redă sinteza sărilor solubile ale conjugatului 13:



Activitatea citotoxică a conjugărilor 12, 14, 16-21 și 23-26 (în unități de ED₅₀ [μg/mL]) sunt redată în tabelul de mai jos:

| | Cell line | |
|------------------------|------------------|-----------------|
| | 1A9 ^c | KB ^c |
| <i>Conjugate</i> | | |
| 12 | 0.49 | 0.20 |
| 14 | 0.20 | 0.94 |
| 16 | 0.005 | 0.005 |
| 17 | 0.14 | 0.39 |
| 18 | 0.13 | 0.26 |
| 19 | 0.20 | 0.47 |
| 20 | 0.005 | 0.14 |
| <i>Glycinate ester</i> | | |
| 21 | 4.47 | 7.65 |
| 23 | 1.01 | 9.79 |
| 24 | 0.63 | 0.70 |
| 25 | 0.19 | 0.34 |
| 26 | 0.11 | 0.09 |
| <i>Control</i> | | |
| 1 | 0.001 | 0.002 |

^a Cytotoxicity as ED₅₀ values for each cell line, the concentration of compound that caused 50% reduction in absorbance at 562 nm relative to untreated cells using the sulforhodamine B assay.

^b All compounds tested did not reach 50% inhibition against human lung carcinoma (A549), breast cancer (MCF-7), human prostate carcinoma (LN-CAP, PC-3, DU-145), and multi-drug resistant KB variant expressing P-glycoprotein (KB-VIN).

^c Human ovarian carcinoma (1A9), human epidermoid carcinoma of the nasopharynx (KB).

Datele de activitate citotoxică pentru conjugății 13 și 15 (și compușii în legătură cu aceștia selectați) sunt redată în tabelul de mai jos:

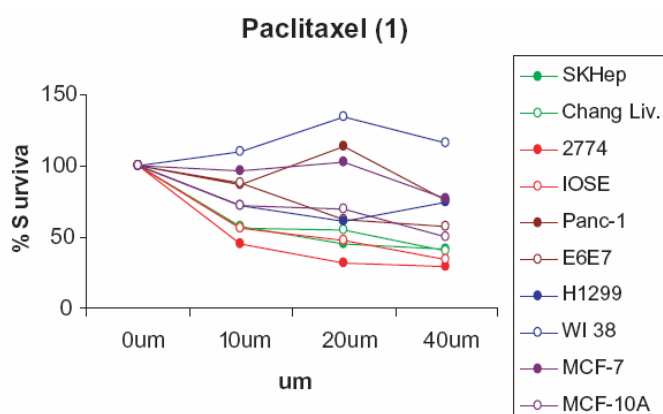
| Compound | ED ₅₀ (μg/mL) | | | | | |
|------------------------------|--------------------------|------|--------------------|-------------------|------|--------|
| | A549 | 1A9 | HCT-8 ^a | A431 ^a | KB | KB-VIN |
| 1 | 0.02 | 0.01 | 0.06 | 0.01 | 0.01 | 0.2 |
| 3 (Vitamin E) | NA ^b | NA | NA | 0.75 | NA | NA |
| 13 (1+3 Conjugate) | NA | NA | NA | 0.1 | NA | NA |
| 1+3 (1:1 Mixture) | 0.02 | 0.01 | 0.20 | 0.01 | 0.01 | 0.1 |
| 22 | NA | NA | NA | 0.2 | NA | NA |
| 22a, 22b | NA | NA | NA | 0.1 | NA | NA |
| 27, 27a, 28, 28a, 28b | NA | NA | NA | NA | NA | NA |
| 29 | NA | NA | NA | 0.1 | NA | NA |
| 29a | NA | NA | NA | 0.2 | NA | NA |
| 5 (Curcumin) | NA | NA | NA | 0.75 | NA | NA |
| 15 (1+5 Conjugate) | 0.03 | 0.03 | NA | 0.03 | 0.01 | 0.2 |
| 1+5 (1:1 Mixture) | 0.03 | 0.01 | 0.10 | 0.01 | 0.01 | 0.2 |

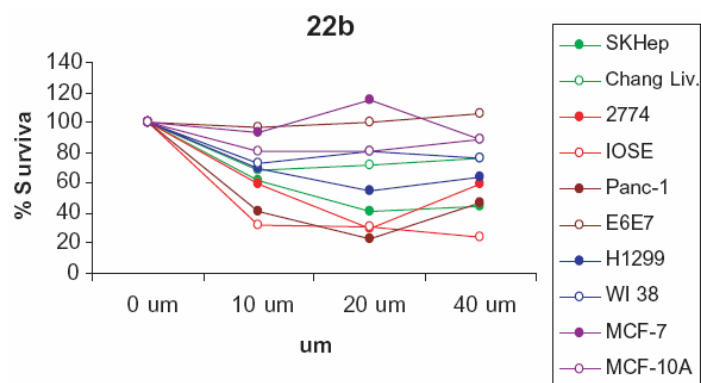
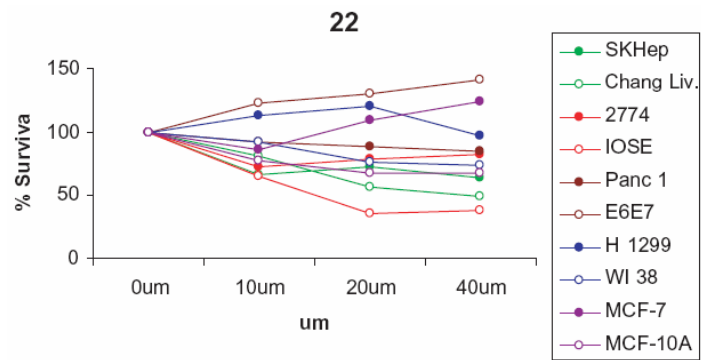
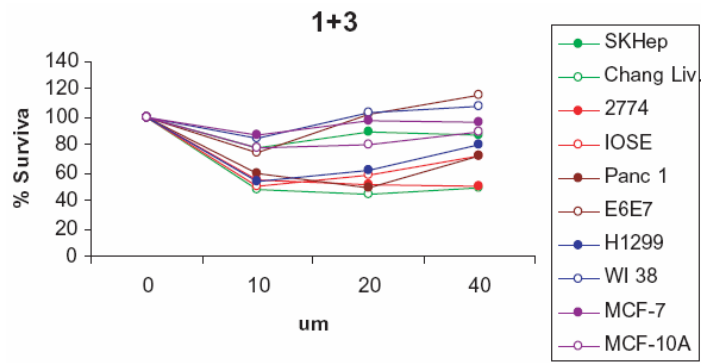
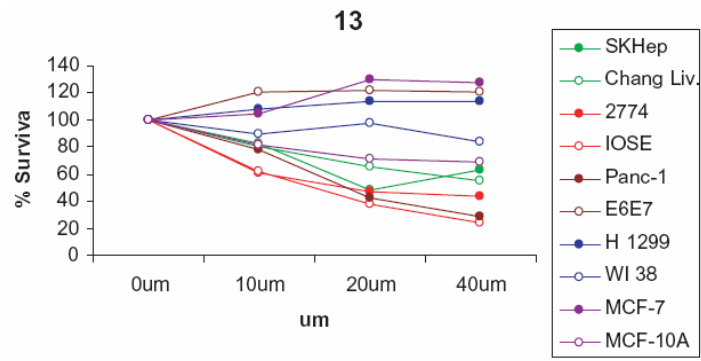
^a Epidermoid skin carcinoma (A431), colon adenocarcinoma (HCT-8); for other cell lines, see Table 1.

^b NA, not active.

Diagramele următoare prezintă activitatea inhibitoare a conjugăților pe:

- ÷ SKHep - hepatomacarcinoma ficatului uman
- ÷ Chang Liv. - celule normale epiteliale de ficat
- ÷ 2774 - carcinom ovariană umană
- ÷ IOSE - suprafața normală a celulei epiteliale ovariene
- ÷ Panc 1 - celula pancreatică canceroasă umană
- ÷ E6E7 - celula epitelială ovariană normală
- ÷ H1299 - celula canceroasă a plămânului uman
- ÷ W138 - fibroplastul uman normal
- ÷ MCF-7 - cancerul de sân
- ÷ MCF-10A - celula epitelială a sânului normal



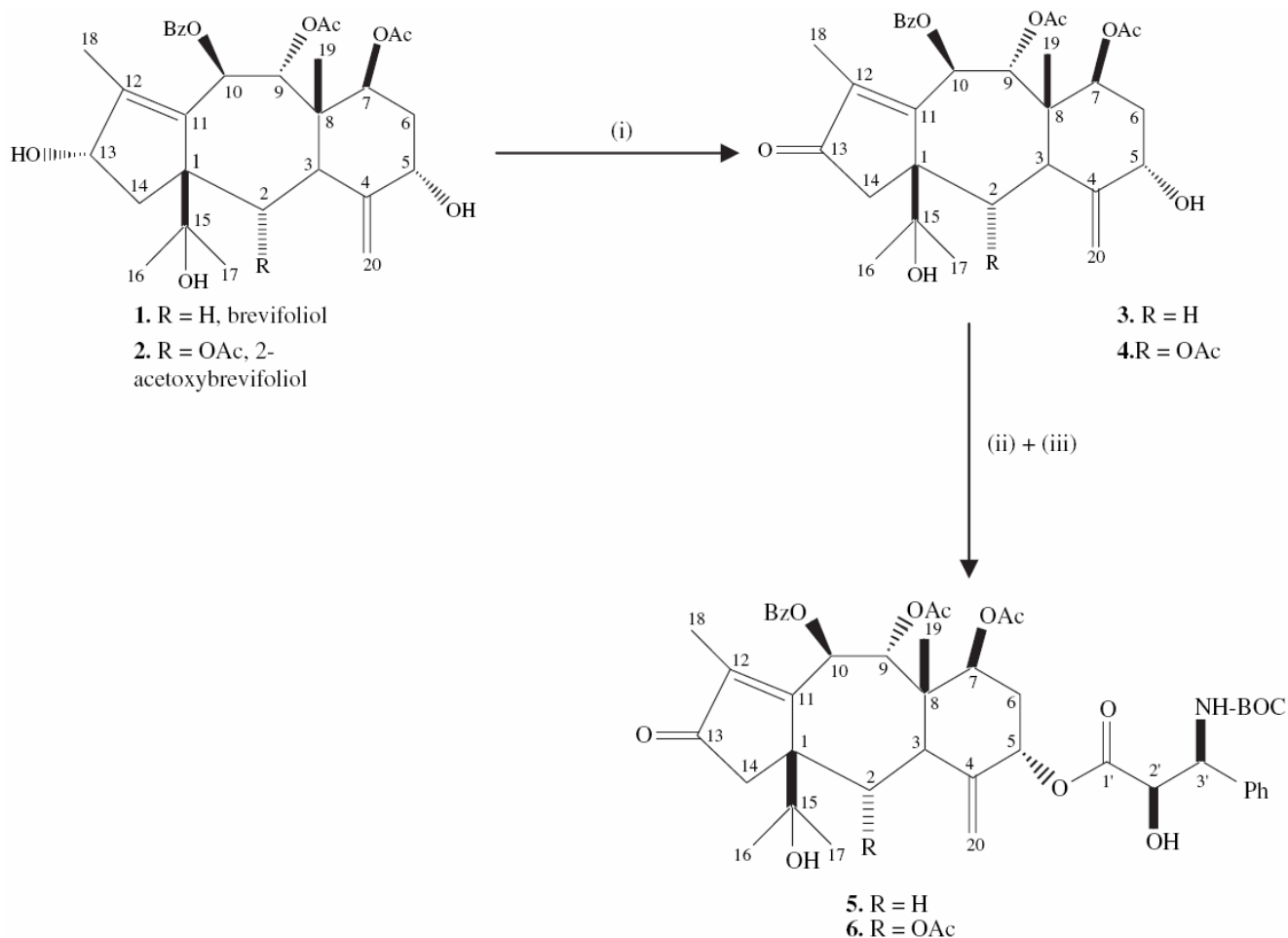


1.10. Setul de taxoizi (4)

- ÷ Lucrare: "Syntheses and cytotoxicities of the analogues of the taxoid brevifoliol"
- ÷ Autori: Sunil K. CHATTOPADHYAY, Sarita TRIPATHI, Mahendra P. DAROKAR, Uzma FARIDI, Brijesh SISODIA, Shubhra NEGI, J. Kotesch KUMAR, Suman P.S. KHANUJ
- ÷ Referință: European Journal of Medicinal Chemistry, doi:10.1016/j.ejmech.2007.09.002
- ÷ Rezumat: Șapte noi analogi de brevifoliol au fost sintetizați prin cuplarea brevifoliolului și acidului carboxilic 2-monosubstituit-4-fenil-1,3-oxazolidină după înlăturarea grupării protectoare prin tratare acidă. Brevifoliolul și analogii săi sintetizați au fost testați pentru activitate citotoxică pe 4 linii de cancer uman diferite, oral (KB), sân (MCF-7), colon (CaCO₂) și ficat (HepG-2) așa cum au fost determinate de testul MTT. Brevifoliolul C-13 oxidat a prezentat o activitate semnificativă. Dintre cei 7 analogi sintetizați, analogul brevifoliol-5-[N-terț-butoxicarbonil] amino-[2'R,3'S]-3'-fenil izoserină] C-13 oxidat a fost de interes atâta timp cât a prezentat citotoxicitate selectivă și potență împotriva celulelor canceroase de ficat într-un mod predominant.

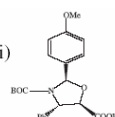
Structura și activitatea taxoizilor

Următoarea este schema de sinteză (BOC=tBuOCO):



Reactanții și condițiile de reacție sunt redată mai jos:

Scheme 1. Reagents and conditions: (i) MnO₂, acetone (5 ml), 25 °C, 30 h, 80%; (ii)



DCC (3 equiv), DMAP (0.2 equiv), toluene (3 ml), 25 °C, 1 h,

70%; (iii) *p*-TSA (1.1 equiv), MeOH (5 ml), 25 °C, 2 h, 51%.

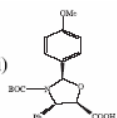
Citotoxicitatea brevifoliolului și analogilor săi pe celulele canceroase umane folosind testul MTT (valori în μg/mL) este redată în tabelul de mai jos:

| Compound | KB | | MCF-7 | |
|------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | IC ₅₀ ± s.d. | IC ₉₀ ± s.d. | IC ₅₀ ± s.d. | IC ₉₀ ± s.d. |
| 1 | 0.0031 ± 0.00047 | 0.6 ± 0.071 | 0.86 ± 0.023 | 6.16 ± 0.471 |
| 2 | 0.316 ± 0.0471 | 7.5 ± 0.816 | 58.3 ± 4.71 | 91.6 ± 2.35 |
| 3 | 0.0078 ± 0.00023 | 1.13 ± 0.262 | 5.33 ± 0.471 | 22.6 ± 2.054 |
| 4 | 100 ± 4.08 | 100 ± 4.08 | 100 ± 4.08 | 100 ± 4.08 |
| 5 | 100 ± 4.08 | 100 ± 4.08 | 8.5 ± 1.08 | 43.33 ± 4.71 |
| 6 | 0.105 ± 0.075 | 5.33 ± 0.471 | 100 ± 4.08 | 100 ± 4.08 |
| 7 | 100 ± 4.08 | 100 ± 4.08 | 33.3 ± 2.35 | 55 ± 4.08 |
| 8 | 100 ± 4.08 | 100 ± 4.08 | 48.3 ± 4.71 | 68.3 ± 4.71 |
| 9 | 2.73 ± 0.21 | 33.3 ± 2.35 | 100 ± 4.08 | 100 ± 4.08 |
| Paclitaxel | 0.0013 ± 0.0004 | 0.048 ± 0.0012 | 0.008 ± 0.0016 | 0.86 ± 0.023 |

| Compound | CaCO2 | | HepG2 | |
|------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | IC ₅₀ ± s.d. | IC ₉₀ ± s.d. | IC ₅₀ ± s.d. | IC ₉₀ ± s.d. |
| 1 | 0.0025 ± 0.00041 | 1.32 ± 0.259 | 0.055 ± 0.0071 | 1.13 ± 0.262 |
| 2 | 6.16 ± 1.027 | 21 ± 2.94 | 96.6 ± 2.35 | 100 ± 4.08 |
| 3 | 0.008 ± 0.0016 | 21 ± 2.94 | 100 ± 4.08 | 100 ± 4.08 |
| 4 | 100 ± 4.08 | 100 ± 4.08 | 96.6 ± 2.35 | 100 ± 4.08 |
| 5 | 100 ± 4.08 | 100 ± 4.08 | 0.065 ± 0.016 | 5.33 ± 0.85 |
| 6 | 100 ± 4.08 | 100 ± 4.08 | 100 ± 4.08 | 100 ± 4.08 |
| 7 | 100 ± 4.08 | 100 ± 4.08 | 100 ± 4.08 | 100 ± 4.08 |
| 8 | 100 ± 4.08 | 100 ± 4.08 | 78.33 ± 4.71 | 91.6 ± 4.71 |
| 9 | 100 ± 4.08 | 100 ± 4.08 | 96.6 ± 2.35 | 100 ± 4.08 |
| Paclitaxel | 0.0078 ± 0.00023 | 0.065 ± 0.016 | 0.008 ± 0.0016 | 1.13 ± 0.262 |

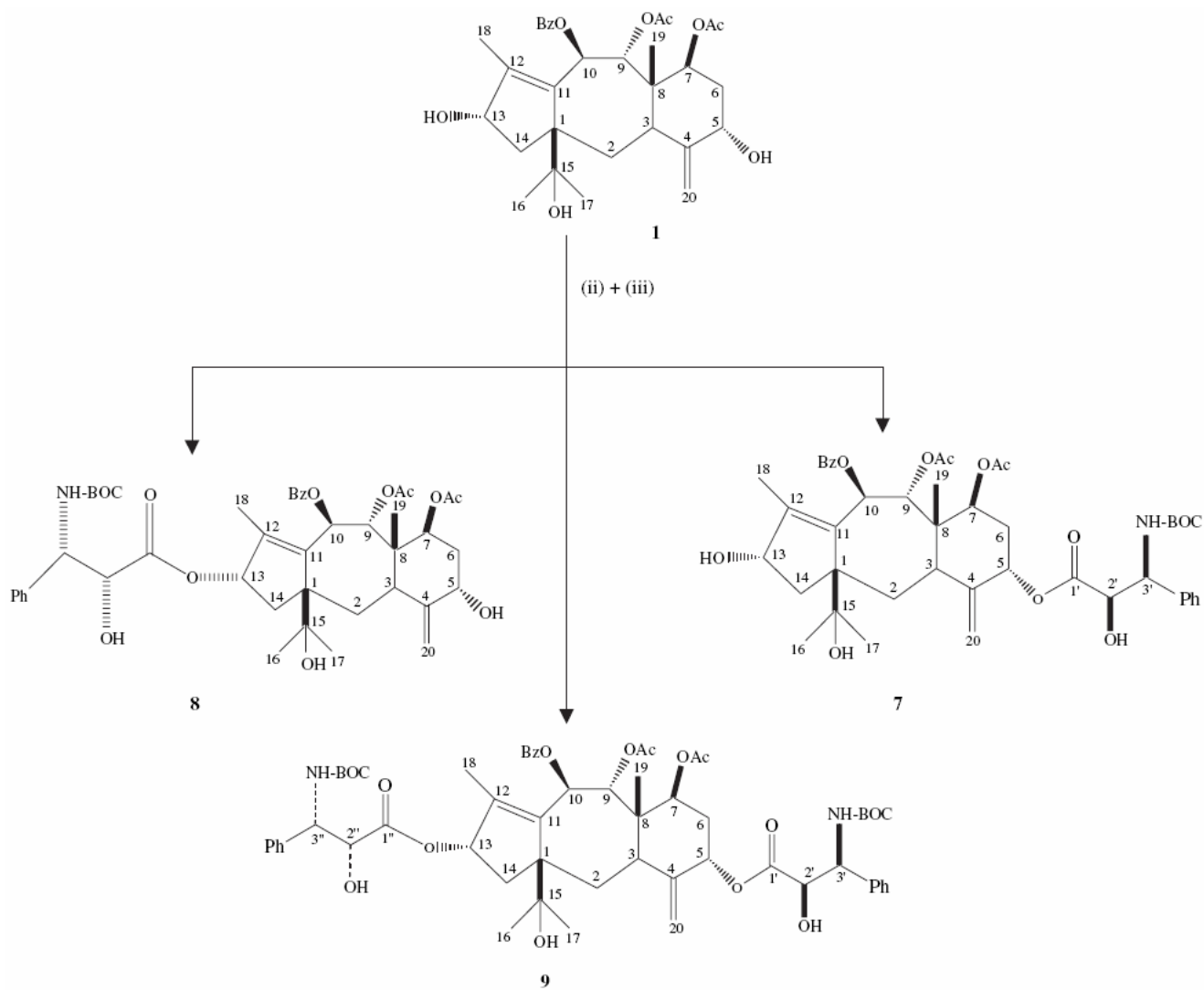
Următoarea schemă de sinteză pentru compușii obținuți folosește reactivii și condițiile de reacție următoare:

Scheme 2. Reagents and conditions: (i)



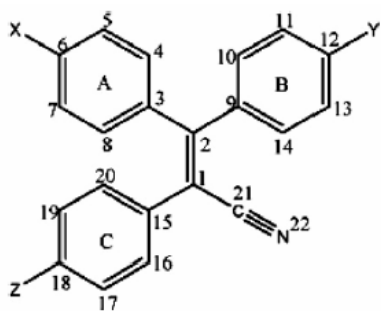
DCC (3 equiv), DMAP (0.2 equiv), toluene (3 ml), 25 °C, 1 h, 70%; (ii) *p*-TSA (1.1 equiv), MeOH (5 ml),

25 °C, 2 h, 51%.



1.11. Trifenilacrilonitrililor

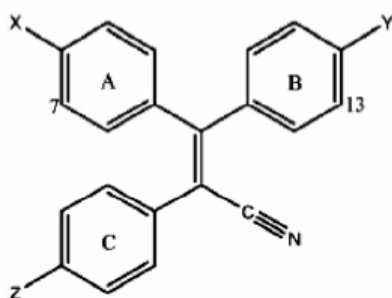
- ÷ Lucrare: “QSAR Studies with E-State Index: Predicting Pharmacophore Signals for Estrogen Receptor Binding Affinity of Triphenylacrylonitriles”
- ÷ Autori: Subhendu MUKHERJEE, Arup MUKHERJEE, Achintya SAHA
- ÷ Referință: Biological and Pharmaceutical Bulletin, 28 (1), 2005, pp. 154-157
- ÷ Rezumat: În conexiune cu descoperirea analogilor estrogenului non-steroidi, studiul de față explorează farmacognozia trifenilacrilonitrililor pentru afinitatea de legătură la receptorul de estrogen utilizând indici de stare electrotopologică (E-State) ale atomilor constitutivi. Analiza arată eficacitatea indicelui E-State în dezvoltarea unui model acceptabil statistic, care definește mediul electronic și stările topologice ale diversilor atomi într-o moleculă. Investigația concluzionează că substituțiile electrofilice la C6 și C18 ale inelelor fenilice (inelele A și respectiv C) atașate la C1 și C2 ale zonei etilenice, împreună cu prezența substituției hidroxilice la C12 (inelul B) și numărul de atomi terminali liberi de hidrogen ale moleculei au influență în afinitatea de legătură a receptorului estrogenic.



÷ Figura legendă:

Structura și activitatea trifenilacrilonitrililor

Figura următoare redă structura generică a fenilacrilonitrililor:



Tabelul următor conține valorile măsurate și valorile calculate pentru activitatea fenilacrilonitrililor:

| Comp. No. | Substitutions | | | Log RBA (competition for [³ H] E ₂ binding) | | | |
|-----------|------------------|------------------|-----------------|--|--------------------|--------------------|---------------------|
| | X | Y | Z | Obs. | Cal. ^{a)} | Cal. ^{b)} | Pred. ^{b)} |
| 1 | H | H | H | -1.046 | -0.107 | -0.928 | -0.916 |
| 2 | H | OH | H | 1.556 | 1.256 | 1.008 | 0.873 |
| 3 | OH | H | H | 0.342 | 0.695 | 0.148 | 0.123 |
| 4 | H | H | OH | 0.519 | 0.299 | -0.280 | -0.488 |
| 5 | OH | OH | H | 1.792 | 2.060 | 2.085 | 2.160 |
| 6 | H | OH | OH | 1.869 | 1.663 | 1.657 | 1.579 |
| 7 | OH | H | OH | 0.785 | 1.102 | 0.798 | 0.801 |
| 8 | OH | OH | OH | 2.220 | 2.468 | 2.735 | 2.929 |
| 9 | CH ₃ | OH | H | 1.447 | 1.095 | 1.186 | 1.143 |
| 10 | OH | CH ₃ | H | 0.398 | 0.135 | -0.046 | -0.104 |
| 11 | OH | OH | CH ₃ | 1.968 | 1.781 | 2.136 | 2.168 |
| 12 | CH ₃ | OH | OH | 1.892 | 1.503 | 1.836 | 1.819 |
| 13 | OH | CH ₃ | OH | 0.959 | 0.543 | 0.604 | 0.480 |
| 14 | OH | OCH ₃ | H | -0.180 | -0.822 | -0.938 | -1.017 |
| 15 | OCH ₃ | OH | H | 1.230 | 1.458 | 1.525 | 1.568 |
| 16 | OH | OiPr | H | -0.444 | -0.417 | -0.233 | -0.200 |
| 17 | OiPr | OH | H | 0.806 | 0.894 | 1.327 | 1.401 |
| 18 | OiPr | OiPr | H | -2.000 | -1.582 | -0.992 | -0.799 |
| 19 | OH | DEAE | H | 0.531 | -0.418 | -0.235 | -0.354 |
| 20 | DEAE | OH | H | 2.033 | 0.887 | 1.320 | 1.218 |
| 21 | DEAE | DEAE | H | -0.398 | -1.591 | -1.000 | -1.115 |
| 22 | H | H | DEAI | -2.000 | -0.698 | -0.855 | -0.616 |
| 23 | DMA | DMA | H | -1.398 | -1.906 | -1.293 | -1.266 |
| 24 | OH | OH | H | -2.000 | -0.250 | — | — |
| | 7 DMAM 13 DMAM | | | | | | |
| 25 | OCH ₃ | OCH ₃ | H | -1.398 | -0.563 | -0.697 | -0.644 |

Legendă: Obs.- observed value; Cal. - calculated value; Pred. - predicted value; *a)* From Eq. (1); *b)* From Eq. (2). OiPr - OCH(CH₃)₂; DEAE - O(CH₂)₂N(C₂H₅)₂; DMA - N(CH₃)₂; DMAM - CH₂N(CH₃)₂.

Modelele SAR pentru trifenilacrilonitrili

Următorul este un model SAR pentru afinitatea de legătură estrogen-receptor a trifenilacrilonitriilor:

$$\text{Log RBA} = 1.261 (\pm 0.491) + 1.906 (\pm 0.317) I_{12\text{-OH}} - 0.239 (\pm 0.077) S_6 \\ - 0.169 (\pm 0.065) S_{18} - 0.559 (\pm 0.143) N_t$$

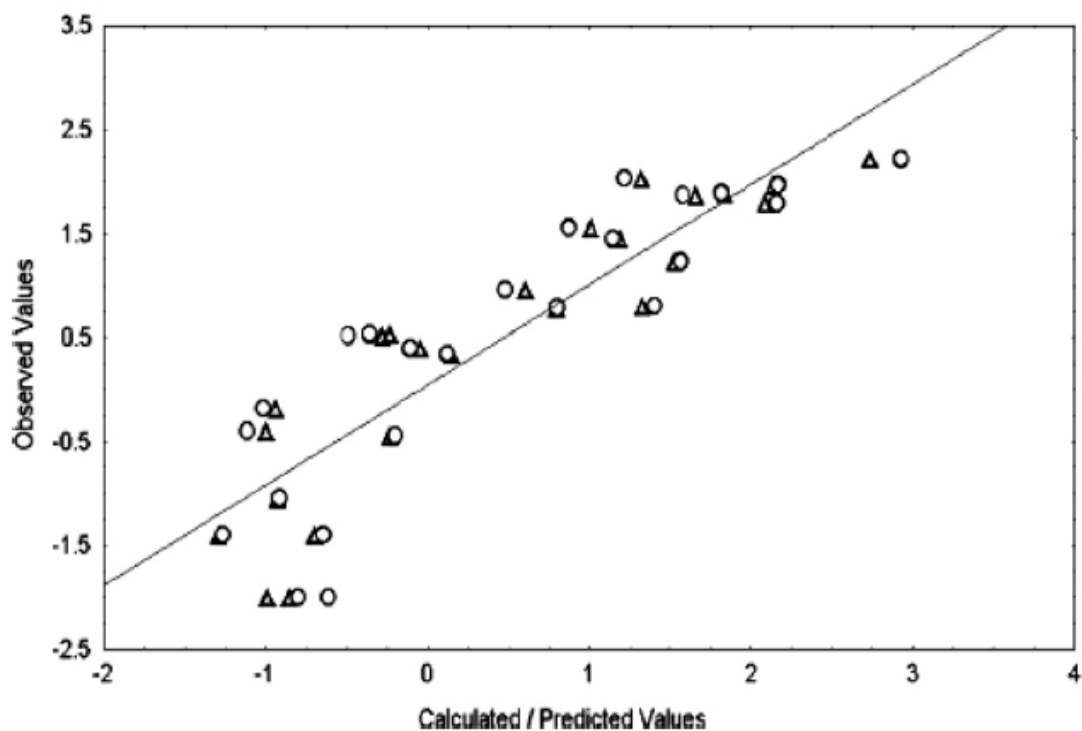
$$n = 25, R = 0.860, R^2 = 0.740, EV = 68.811\%, F = 14.238 (df 4, 20), \\ s = 0.774, AVRES = 0.542, PRESS = 21.751, SDEP = 0.933, \\ \text{Pres}_{\text{av}} = 0.697, Q^2 = 0.529.$$

Următorul este un model SAR pentru afinitatea de legătură estrogen-receptor a trifenilacrilonitriilor:

$$\text{Log RBA} = 2.114(\pm 0.243)I_{12\text{-OH}} - 0.223(\pm 0.059)S_6 \\ - 0.147(\pm 0.050)S_{18} - 0.193(\pm 0.052)N_t$$

$n=24$, $R=0.919$, $R^2=0.845$, $EV=81.415\%$, $F=27.284$ (df 4, 20),
 $s=0.595$, $AVRES=0.450$, $PRESS=10.021$, $SDEP=0.646$,
 $Pres_{av}=0.515$, $Q^2=0.751$.

Următorul grafic prezintă capacitatea estimativă a celui de-al doilea model obținut pentru legătura între structura și afinitatea de legătură estrogen-receptor a trifenilacrilonitrililor:



Activitatea 3. Achiziție, instalare, testare și configurare aparatură suport

Achiziție echipamente

În conformitate cu propunerea de proiect, a fost achiziționat un server și 2 unități de calcul și un sistem de achiziție de date (FEBE) - la care cheltuielile au fost suportate doar parțial din prezentul contract.

Activitatea de achiziție echipamente a cuprins selectarea celor mai reprezentative echipamente oferite de piață. Lista acestora este redată în continuare:

■ Server HP DL360 G5 5355 470064-466

- ÷ Preț: 12118.11 RON
- ÷ Producător: HP
- ÷ Model procesor: Intel Xeon 5355 Dual Core
- ÷ Fast System Bus: 1333 Mhz
- ÷ Model placă de bază: Intel 5000P
- ÷ Memorie maximă: 32768 Mb
- ÷ Unitate optică: DVD-RW
- ÷ Componentă de rețea: 10/100/1000 Mbps
- ÷ Componentă video: ATI ES1000 32MB SDRAM
- ÷ Memorie rapidă procesor: 2×4096 Mb
- ÷ Număr maxim de procesoare: 2
- ÷ Memorie instalată: 4096 Mb (2×2048 Mb)
- ÷ Matrice Rapid Access Interface Device: HP Smart Array P400i/256MB BBWC
- ÷ Sisteme de operare compatibile (conform dealer):
 - Microsoft Windows Server 2000
 - Microsoft Windows Server 2003
 - Novell NetWare
 - Red Hat Enterprise Linux
 - SUSE Linux Enterprise Server
 - SCO UnixWare, OpenServer
 - VMware Virtualization Software
 - Solaris 10 32/64-bit

■ Server HP Compaq Proliant ML370G5 470064-385

- ÷ Preț: 8836.67 RON
- ÷ Producător: HP
- ÷ Model procesor: Intel Xeon 5335 Quad Core

- ÷ Tip memorie: PC2-5300 Fully Buffered DIMMs (DDR2-667)
- ÷ Model placă de bază: Intel 5000P
- ÷ Memorie maximă: 65536 Mb
- ÷ Unitate optică: DVD-RW
- ÷ Componentă video: ATI ES1000 32MB SDRAM
- ÷ Sloturi de memorie: 8
- ÷ Memorie instalată: 2048 Mb (1×1024 Mb)
- ÷ Harddisk: 146 GB, 10K SAS
- HP Compaq dx5750 Athlon64 X2 DualCore 3200+ 1G 80G XP
- ÷ Procesor: Athlon 64 X2 Dual-Core 3800+
- ÷ Frecvență processor: 2.0-GHz
- ÷ Memorie rapidă procesor: 2x512K L2 cache
- ÷ Fast System Bus: 2000-MHz
- ÷ Componentă periferice: ATI Express 1150 3 bay-uri: 2 externe (1 - 5.25", 1 - 3.5"), 1 intern - 3.5"
- ÷ Placa de bază: 2 low profile PCI, 1 LP PCIe x1 1 LP PCIe x16 porturi: Spate: 6 USB 2.0, 1 standard serial, 1 parallel, 2 PS/2, 1 RJ-45, 1 VGA, 1 DVI-D, audio in/out Fata: 2 USB 2.0, casti si microfon sursa 240W 2004 standard keyboard PS/2, PS/2 Optical Mouse;
- ÷ Memorie internă: 1024MB (2x512) DDR SDRAM PC 5300, 667MHz non ECC, expandabila la 4 GB (4 sloturi);
- ÷ HardDisk: 80GB Serial ATA, 7.200 rpm; DVD-CDRW 48x/32x combo;
- ÷ Componentă audio: Controler audio integrat Realtek 2 channel ALC260, difuzor intern;
- ÷ Componentă rețea: Broadcom NetXtreme Gigabit Ethernet;
- ÷ Adaptor video integrat Radeon X300;
- ÷ Sistem de operare: MS Windows XP Professional

Instalare echipamente

Sistemele tip server au fost instalate cu FreeBSD v. 6.2 platforma amd64 (cele echipate cu procesor AMD) și respectiv ia64 (cele echipate cu procesor Intel).

În continuare se redau câteva din componentele instalate pentru un server rulând pe platforma amd64:

#dmesg

Copyright (c) 1992-2007 The FreeBSD Project.

Copyright (c) 1979, 1980, 1983, 1986, 1988, 1989, 1991, 1992, 1993, 1994

The Regents of the University of California. All rights reserved.

FreeBSD is a registered trademark of The FreeBSD Foundation.

FreeBSD 6.2-RELEASE #3: Wed Nov 31 15:58:36 EET 2007

Timecounter "i8254" frequency 1193182 Hz quality 0
CPU: Intel(R) Pentium(R) 4 CPU 3.20GHz (3194.56-MHz K8-class CPU)
Origin = "GenuineIntel" Id = 0xf49 Stepping = 9
Features=0xbfebfbff<FPU,VME,DE,PSE,TSC,MSR,PAE,MCE,CX8,APIC,SEP,MTRR,PGE,MCA,CMOV,PAT,PS
E36,CLFLUSH,DTS,ACPI,MMX,FXSR,SSE,SSE2,SS,HTT,TM,PBE>
Features2=0x641d<SSE3,RSVD2,MON,DS_CPL,CNTX-ID,CX16,<b14>>
AMD Features=0x20100800<SYSCALL,NX,LM>
AMD Features2=0x1<LAHF>
Logical CPUs per core: 2
real memory = 2145320960 (2045 MB)
avail memory = 2064601088 (1968 MB)
ACPI APIC Table: <IntelR AWRDACPI>
FreeBSD/SMP: Multiprocessor System Detected: 2 CPUs
cpu0 (BSP): APIC ID: 0
cpu1 (AP): APIC ID: 1
ioapic0: Changing APIC ID to 4
ioapic0 <Version 2.0> irqs 0-23 on motherboard
acpi0: <IntelR AWRDACPI> on motherboard
acpi0: Power Button (fixed)
Timecounter "ACPI-fast" frequency 3579545 Hz quality 1000
acpi_timer0: <24-bit timer at 3.579545MHz> port 0x408-0x40b on acpi0
cpu0: <ACPI CPU> on acpi0
cpu1: <ACPI CPU> on acpi0
acpi_button0: <Power Button> on acpi0
pcib0: <ACPI Host-PCI bridge> port 0xcf8-0xcff on acpi0
pci0: <ACPI PCI bus> on pcib0
pci0: <display, VGA> at device 2.0 (no driver attached)
pci0: <serial bus, USB> at device 29.0 (no driver attached)
pci0: <serial bus, USB> at device 29.1 (no driver attached)
pci0: <serial bus, USB> at device 29.2 (no driver attached)
pci0: <serial bus, USB> at device 29.3 (no driver attached)
pci0: <serial bus, USB> at device 29.7 (no driver attached)
pcib1: <ACPI PCI-PCI bridge> at device 30.0 on pci0
pci1: <ACPI PCI bus> on pcib1
rl0: <RealTek 8139 10/100BaseTX> port 0xd000-0xd0ff mem 0xf8400000-0xf84000ff irq 23 at device 3.0 on pci1
miibus0: <MII bus> on rl0
rlphy0: <RealTek internal media interface> on miibus0
rlphy0: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, auto
rl0: Ethernet address: 00:15:58:68:cf:7f
isab0: <PCI-ISA bridge> at device 31.0 on pci0
isa0: <ISA bus> on isab0

atapci0: <Intel ICH5 UDMA100 controller> port 0x1f0-0x1f7,0x3f6,0x170-0x177,0x376,0xf00-0xf00f at device 31.1 on pci0
ata0: <ATA channel 0> on atapci0
ata1: <ATA channel 1> on atapci0
atapci1: <Intel ICH5 SATA150 controller> port 0xe500-0xe507,0xe600-0xe603,0xe700-0xe707,0xe800-0xe803,0xe900-0xe90f irq 18 at device 31.2 on pci0
ata2: <ATA channel 0> on atapci1
ata3: <ATA channel 1> on atapci1
pci0: <serial bus, SMBus> at device 31.3 (no driver attached)
pcm0: <Intel ICH5 (82801EB)> port 0xeb00-0xebff,0xec00-0xec3f mem 0xf8581000-0xf85811ff,0xf8582000-0xf85820ff irq 17 at device 31.5 on pci0
pcm0: primary codec not ready!
pcm0: <Avance Logic ALC655 AC97 Codec>
acpi_tz0: <Thermal Zone> on acpi0
speaker0: <PC speaker> port 0x61 on acpi0
atkbd0: <Keyboard controller (i8042)> port 0x60,0x64 irq 1 on acpi0
atkbd0: <AT Keyboard> irq 1 on atkbd0
atkbd0: [GIANT-LOCKED]
orm0: <ISA Option ROM> at iomem 0xc0000-0xc9fff on isa0
sc0: <System console> at flags 0x100 on isa0
sc0: VGA <4 virtual consoles, flags=0x300>
vga0: <Generic ISA VGA> at port 0x3c0-0x3df iomem 0xa0000-0xbffff on isa0
Timecounters tick every 1.000 msec
ad0: 476940MB <Seagate ST3500630A 3.AAE> at ata1-slave UDMA100
ad1: 157066MB <Hitachi HDS721616PLA380 P22OA70A> at ata2-master SATA150
SMP: AP CPU #1 Launched!

#pkg_info -a

Information for apache-1.3.37_2:

Comment:

The extremely popular Apache http server. Very fast, very clean

Required by:

php5-5.2.0

php5-mysql-5.2.0

Description:

Apache is an HTTP server designed as a plug-in replacement for the NCSA server version 1.3 (or 1.4). It fixes numerous bugs in the NCSA server and includes many frequently requested new features, and has an API which allows it to be extended to meet users' needs more easily.

WWW: <http://httpd.apache.org/>

Information for autoconf-2.59_2:

Comment:

Automatically configure source code on many Un*x platforms (2.59)

Description:

Autoconf is an extensible package of m4 macros that produce shell scripts to automatically configure software source code packages. These scripts can adapt the packages to many kinds of UNIX-like systems without manual user intervention. Autoconf creates a configuration script for a package from a template file that lists the operating system features that the package can use, in the form of m4 macro calls.

WWW: <http://www.gnu.org/software/autoconf/>

Information for bash-3.1.17:

Comment:

The GNU Project's Bourne Again SHell

Description:

This is GNU Bash, version 3.0. Bash is the GNU Project's Bourne Again SHell, a complete implementation of the POSIX.2 shell spec, but also with interactive command line editing, job control on architectures that support it, csh-like features such as history substitution and brace expansion, and a slew of other features.

WWW: <http://cnswww.cns.cwru.edu/~chet/bash/bashtop.html>

Information for cvsup-without-gui-16.1h_2:

Comment:

General network file distribution system optimized for CVS (non-GUI version)

Description:

CVSup is a software package for distributing and updating collections of files across a network. It can efficiently and accurately mirror all types of files, including sources, binaries, hard links, symbolic links, and even device nodes. CVSup's streaming communication protocol and multithreaded architecture make it most likely the fastest mirroring tool in existence today. In addition to being a great general-purpose mirroring tool, CVSup includes special features and optimizations specifically tailored to CVS repositories. This port of CVSup omits the GUI and does not require X11. For a version that includes the GUI, use the "net/cvsup" port.

WWW: <http://www.cvsup.org/>

jdp@polstra.com

Information for expat-2.0.0_1:

Comment:

XML 1.0 parser written in C

Required by:

apache-1.3.37_2

php5-5.2.0

php5-mysql-5.2.0

p5-XML-Parser-2.34_2

intltool-0.35.4

Description:

Expat is an XML 1.0 parser written in C. It aims to be fully conforming. It is currently not a validating XML processor.

WWW: <http://sourceforge.net/projects/expat/>

Information for gawk-3.1.1_1:

Comment:

The GNU version of Awk

Required by:

mc-light-4.1.40.p9_6

Description:

This is GNU Awk. It should be upwardly compatible with the Bell Labs research version of awk. It is almost completely compliant with the 1993 POSIX 1003.2 standard for awk.

Information for gettext-0.14.5_2:

Comment:

GNU gettext package

Required by:

gawk-3.1.1_1

mc-light-4.1.40.p9_6

bash-3.1.17

p5-gettext-1.05_1

gmake-3.81_1

help2man-1.36.4_1

popt-1.7_2

glib-2.12.9

Description:

This is the GNU gettext package. It is interesting for authors or maintainers of other packages or programs which they want to see internationalized. As one step the handling of messages in different languages should be implemented. For this task GNU gettext provides the needed tools and library functions.

WWW: <http://www.gnu.org/software/gettext/>

Information for glib-2.12.9:

Comment:

Some useful routines of C programming (current stable version)

Description:

GLib provides a simple, abstract, and highly portable set of C support routines such as lists, trees, hashes, memory allocation, and much, much more. It is a foundation for writing portable applications.

WWW: <http://www.gtk.org/>

Information for gmake-3.81_1:

Comment:

GNU version of 'make' utility

Description:

GNU make is a tool that controls the generation of executables and other non-source files from source files. Its purpose is the same as that of the utility make(1).

WWW: <http://www.gnu.org/software/make/make.html>

Information for help2man-1.36.4_1:

Comment:

Automatically generating simple manual pages from program output

Description:

help2man is a tool for automatically generating simple manual pages from program output.

It is intended to provide an easy way for software authors to include a manual page in their distribution without having to maintain that document.

Given a program which produces reasonably standard --help and --version outputs, help2man will attempt to re-arrange that output into something which resembles a manual page.

WWW: <http://www.gnu.org/software/help2man/>

Information for icu-3.6:

Comment:

International Components for Unicode (from IBM)

Required by:

glib-2.12.9

Description:

The International Components for Unicode (ICU) is a C and C++ library providing Unicode support (p.ex. character set conversions, locales, normalization, transliteration, etc.).

More information on ICU can be found on the homepage

WWW: <http://icu.sourceforge.net/>

WWW: <http://ibm.com/software/globalization/icu>

- Bjoern A. Zeeb

bzeeb+freebsdports@zabbadoz.net

Information for intltool-0.35.4:

Comment:

Tools to internationalize various kinds of data files

Description:

The module contains some utility scripts and assorted auto* magic for internationalizing various kinds of XML files.

*** Features**

- o Extract translatable strings from various source files (.xml.in, .glade, .desktop.in, .server.in, .oaf.in).
- o Collect the extracted strings together with messages from traditional source files (.c, .h) in po/\$(PACKAGE).pot.
- o Merge back the translations from .po files into .xml, .desktop and .oaf files. This merge step will happen at build resp. installation time.

WWW: <http://www.gnome.org/>

Information for libiconv-1.9.2_2:

Comment:

A character set conversion library

Required by:

gettext-0.14.5_2

gawk-3.1.1_1

mc-light-4.1.40.p9_6

bash-3.1.17

p5-gettext-1.05_1

gmake-3.81_1

help2man-1.36.4_1

libxml2-2.6.27

php5-5.2.0

php5-mysql-5.2.0

intltool-0.35.4

popt-1.7_2

glib-2.12.9

Description:

This library provides an iconv() implementation, for use on systems which don't have one, or whose implementation cannot convert from/to Unicode. It can convert from any of these encodings to any other, through Unicode conversion. It has also some limited support for transliteration, i.e. when a character cannot be represented in the target character set, it can

be approximated through one or several similarly looking characters.
libiconv is for you if your application needs to support multiple character encodings, but that support lacks from your system.

See in either README or website for the list of support encodings.

WWW: <http://www.gnu.org/software/libiconv/>

Information for libtool-1.5.22_3:

Comment:

Generic shared library support script

Description:

This is GNU Libtool, a generic library support script. Libtool hides the complexity of using shared libraries behind a consistent, portable interface.

To use libtool, add the new generic library building commands to your Makefile, Makefile.in, or Makefile.am.

WWW: <http://www.gnu.org/software/libtool/>

Information for libxml2-2.6.27:

Comment:

XML parser library for GNOME

Required by:

php5-5.2.0

php5-mysql-5.2.0

Description:

Libxml2 is the XML C parser and toolkit developed for the Gnome project (but usable outside of the Gnome platform). XML itself is a metalanguage to design markup languages, i.e. text language where semantic and structure are added to the content using extra "markup" information enclosed between angle brackets.

WWW: <http://xmlsoft.org/>

Information for m4-1.4.8_1:

Comment:

GNU m4

Required by:

autoconf-2.59_2

Description:

GNU `m4' is an implementation of the traditional UNIX macro processor. It is mostly SVR4 compatible, although it has some extensions (for example, handling more than 9 positional parameters to macros). `m4' also has builtin functions for including files, running shell commands, doing arithmetic, etc.

WWW: <http://www.gnu.org/software/m4/m4.html>

Information for mc-light-4.1.40.p9_6:

Comment:

A lightweight Midnight Commander clone

Description:

Advanced Midnight Commander 4.1.40 by Olegarch. The goal of this project is creating a stable, well-working, usefull console-only version of well-known Midnight Commander, without bugs and garbage, like tk, xv and gnome.

WWW: <http://mc.linuxinside.com/>

Information for memtest-4.0.6:

Comment:

Utility to test for faulty memory subsystem

Description:

memtest is a utility for testing the memory subsystem in a computer to determine if it is faulty. The original source was by Simon Kirby <sim@stormix.com>. The program has been rewritten by Charles Cazabon and many additional tests were added to help catch borderline memory. He also rewrote the original tests (which catch mainly memory bits which are stuck permanently high or low) so that they run approximately an order of magnitude faster.

WWW: <http://pyropus.ca/software/memtester/>

Information for mysql-client-5.1.14:

Comment:

Multithreaded SQL database (client)

Required by:

php5-mysql-5.2.0

mysql-server-5.1.14

Description:

MySQL is a very fast, multi-threaded, multi-user and robust SQL (Structured Query Language) database server.

WWW: <http://www.mysql.com/>

- Alex Dupre

ale@FreeBSD.org

Information for mysql-server-5.1.14:

Comment:

Multithreaded SQL database (server)

Description:

MySQL is a very fast, multi-threaded, multi-user and robust SQL (Structured Query Language) database server.

WWW: <http://www.mysql.com/>

- Alex Dupre

ale@FreeBSD.org

Information for p5-XML-Parser-2.34_2:

Comment:

Perl extension interface to James Clark's XML parser, expat

Required by:

intltool-0.35.4

Description:

Perl extension interface to James Clark's XML parser, expat.

WWW: <http://search.cpan.org/dist/XML-Parser/>

Information for p5-gettext-1.05_1:

Comment:

Message handling functions

Required by:

help2man-1.36.4_1

Description:

This module permits from perl to the gettext() family of functions for retrieving message strings from databases constructed to internationalize software.

--Dima.

dima@Chg.RU

Information for perl-5.8.8:

Comment:

Practical Extraction and Report Language

Required by:

apache-1.3.37_2

p5-gettext-1.05_1

help2man-1.36.4_1

autoconf-2.59_2

php5-5.2.0

php5-mysql-5.2.0

p5-XML-Parser-2.34_2

intltool-0.35.4

glib-2.12.9

Description:

Perl is a language that combines some of the features of C, sed, awk and

shell. See the manual page for more hype. There are also many books published by O'Reilly & Assoc. See pod/perlbook.pod for more information.

WWW: <http://www.perl.org/>

- MarkM

Information for php5-5.2.0:

Comment:

PHP Scripting Language (Apache Module and CLI)

Required by:

php5-mysql-5.2.0

Description:

PHP, which stands for "PHP: Hypertext Preprocessor" is a widely-used Open Source general-purpose scripting language that is especially suited for Web development and can be embedded into HTML. Its syntax draws upon C, Java, and Perl, and is easy to learn. The main goal of the language is to allow web developers to write dynamically generated webpages quickly, but you can do much more with PHP.

WWW: <http://www.php.net/>

- Alex Dupre

ale@FreeBSD.org

Information for php5-mysql-5.2.0:

Comment:

The mysql shared extension for php

Description:

PHP, which stands for "PHP: Hypertext Preprocessor" is a widely-used Open Source general-purpose scripting language that is especially suited for Web development and can be embedded into HTML. Its syntax draws upon C, Java, and Perl, and is easy to learn. The main goal of the language is to allow web developers to write dynamically generated webpages quickly, but you can do much more with PHP.

WWW: <http://www.php.net/>

- Alex Dupre

ale@FreeBSD.org

Information for pkg-config-0.21:

Comment:

A utility to retrieve information about installed libraries

Required by:

mc-light-4.1.40.p9_6

libxml2-2.6.27

php5-5.2.0

php5-mysql-5.2.0

glib-2.12.9

Description:

The pkg-config program is used to retrieve information about installed libraries in the system. It is typically used to compile and link against one or more libraries. It's more useful to the compilation process than to the end-user.

pkg-config retrieves information about packages from special metadata files.

These files are named after the package, with the extension .pc. By default,

pkg-config looks for these files in the following directories:

`${PREFIX}/libdata/pkgconfig`, `${PREFIX}/lib/pkgconfig`,

`${LOCALBASE}/libdata/pkgconfig`, `${LOCALBASE}/lib/pkgconfig`,

`${X11BASE}/libdata/pkgconfig`, and `${X11BASE}/lib/pkgconfig`; it will also look in the list of directories specified by the `PKG_CONFIG_PATH` environment variable.

WWW: <http://pkgconfig.freedesktop.org/wiki/>

Information for popt-1.7_2:

Comment:

A getopt(3) like library with a number of enhancements, from Redhat

Description:

This is the popt command line option parsing library. While it is similiar to getopt(3), it contains a number of enhancements, including:

- 1) popt is fully reentrant
- 2) popt can parse arbitrary argv[] style arrays while getopt(2) makes this quite difficult
- 3) popt allows users to alias command line arguments
- 4) popt provides convience functions for parsing strings into argv[] style arrays

WWW: <http://freshmeat.net/projects/popt/>

Testare echipamente

Sistemele server au fost testate cu aplicația memtest. În continuare este redată documentația (manualul) acestei aplicații:

\$ man memtest

memtester(8) Maintenance Commands memtester(8)

NAME

memtester - stress test to find memory subsystem faults.

SYNOPSIS

memtester <MEMORY> [ITERATIONS]

DESCRIPTION

memtester is an effective userspace tester for stress-testing the memory subsystem. It is very effective at finding intermittent and non-deterministic faults. Note that problems in other hardware areas (overheating CPU, out-of-specification power supply, etc.) can cause memory faults.

memtester will **malloc(3)** the amount of memory specified, if possible. If this fails, it will decrease the amount of memory requested until it succeeds. It will then attempt to **mlock(3)** this memory; if it cannot do so, testing will be slower and less effective.

OPTIONS

MEMORY the amount of memory to allocate and test, in megabytes

ITERATIONS

(optional) number of loops to iterate through. Default is infinite.

NOTE

memtester must be run with root privileges to **mlock(3)** its pages.

Testing memory without locking the pages in place is mostly pointless and slow.

EXIT CODE

memtester's exit code is 0 when everything works properly. Otherwise, it is the logical OR of the following values:

- x01** error allocating or locking memory, or invocation error
- x02** error during stuck address test
- x04** error during one of the other tests

Rezultatele câtorva testări cu memtest sunt expuse în tabelul următor:

| Nr | Comanda | Rezultat |
|----|-----------------|--|
| 1 | \$ memtest 16 2 | memtester version 4.0.6 (64-bit) Copyright (C) 2006 Charles Cazabon. Licensed under the GNU General Public License version 2 (only). pagesize is 4096 pagesizemask is 0xfffffffffff000 want 16MB (16777216 bytes) got 16MB (16777216 bytes), trying mlock ...insufficient permission. Trying again, unlocked: got 16MB (16777216 bytes) Continuing with unlocked memory; testing will be slower and less reliable. Loop 1/2: Stuck Address : ok Random Value : ok Compare XOR : ok Compare SUB : ok Compare MUL : ok Compare DIV : ok Compare OR : ok |

| | | |
|---|---------------|---|
| | | <p> Compare AND : ok Sequential Increment: ok Solid Bits : ok Block Sequential : ok Checkerboard : ok Bit Spread : ok Bit Flip : ok Walking Ones : ok Walking Zeroes : ok Loop 2/2: Stuck Address : ok Random Value : ok Compare XOR : ok Compare SUB : ok Compare MUL : ok Compare DIV : ok Compare OR : ok Compare AND : ok Sequential Increment: ok Solid Bits : ok Block Sequential : ok Checkerboard : ok Bit Spread : ok Bit Flip : ok Walking Ones : ok Walking Zeroes : ok Done. </p> |
| 2 | #memtest 16 2 | <p> Copyright (C) 2006 Charles Cazabon. Licensed under the GNU General Public License version 2 (only). pagesize is 4096 pagesizemask is 0xfffffffffff000 want 16MB (16777216 bytes) got 16MB (16777216 bytes), trying mlock ...locked. Loop 1/2: Stuck Address : ok Random Value : ok Compare XOR : ok Compare SUB : ok Compare MUL : ok Compare DIV : ok Compare OR : ok Compare AND : ok Sequential Increment: ok Solid Bits : ok Block Sequential : ok Checkerboard : ok Bit Spread : ok Bit Flip : ok Walking Ones : ok Walking Zeroes : ok Loop 2/2: Stuck Address : ok </p> |

| | | |
|---|--------------|---|
| | | <pre> Random Value : ok Compare XOR : ok Compare SUB : ok Compare MUL : ok Compare DIV : ok Compare OR : ok Compare AND : ok Sequential Increment: ok Solid Bits : ok Block Sequential : ok Checkerboard : ok Bit Spread : ok Bit Flip : ok Walking Ones : ok Walking Zeroes : ok Done. </pre> |
| 3 | #memtest 4 1 | <pre> memtester version 4.0.6 (64-bit) Copyright (C) 2006 Charles Cazabon. Licensed under the GNU General Public License version 2 (only). pagesize is 4096 pagesizemask is 0xfffffffffff000 want 4MB (4194304 bytes) got 4MB (4194304 bytes), trying mlock ...locked. Loop 1/1: Stuck Address : ok Random Value : ok Compare XOR : ok Compare SUB : ok Compare MUL : ok Compare DIV : ok Compare OR : ok Compare AND : ok Sequential Increment: ok Solid Bits : ok Block Sequential : ok Checkerboard : ok Bit Spread : ok Bit Flip : ok Walking Ones : ok Walking Zeroes : ok Done. </pre> |

Sistemele server au fost testate cu aplicația netstat. În continuare este redată documentația (manualul) acestei aplicații:

NETSTAT(1) FreeBSD General Commands Manual NETSTAT(1)

NAME

netstat -- show network status

DESCRIPTION

The netstat command symbolically displays the contents of various network-related data structures. There are a number of output formats, depending on the options for the information presented.

**netstat [-AaLnSW] [-f protocol_family | -p protocol] [-M core]
[-N system]**

Display a list of active sockets (protocol control blocks) for each network protocol, for a particular protocol_family, or for a single protocol. If -A is also present, show the address of a protocol control block (PCB) associated with a socket; used for debugging. If -a is also present, show the state of all sockets; normally sockets used by server processes are not shown. If -L is also present, show the size of the various listen queues. The first count shows the number of unaccepted connections, the second count shows the amount of unaccepted incomplete connections, and the third count is the maximum number of queued connections. If -S is also present, show network addresses as numbers (as with -n) but show ports symbolically.

**netstat -i | -I interface [-abdhtW] [-f address_family] [-M core]
[-N system]**

Show the state of all network interfaces or a single interface which have been auto-configured (interfaces statically configured into a system, but not located at boot time are not shown). An asterisk (*') after an interface name indicates that the interface is 'down'. If -a is also present, multicast addresses currently in use are shown for each Ethernet interface and for each IP interface address. Multicast addresses are shown on separate lines following the interface address with which they are associated. If -b is also present, show the number of bytes in and out. If -d is also present, show the number of dropped packets. If -h is also present, print all counters in human readable form. If -t is also present, show the contents of watchdog timers. If -W is also present, print interface names using a wider field size.

netstat -w wait [-I interface] [-d] [-M core] [-N system]

At intervals of wait seconds, display the information regarding packet traffic on all configured network interfaces or a single interface. If -d is also present, show the number of dropped packets.

**netstat -s [-s] [-z] [-f protocol_family | -p protocol] [-M core]
[-N system]**

Display system-wide statistics for each network protocol, for a particular protocol_family, or for a single protocol. If -s is repeated, counters with a value of zero are suppressed. If -z is also present, reset statistic counters after displaying them.

netstat -i | -I interface -s [-f protocol_family | -p protocol] [-M core]

[-N system]

Display per-interface statistics for each network protocol, for a particular protocol_family, or for a single protocol.

netstat -m [-M core] [-N system]

Show statistics recorded by the memory management routines (mbuf(9)). The network manages a private pool of memory buffers.

netstat -r [-AanW] [-f address_family] [-M core] [-N system]

Display the contents of all routing tables, or a routing table for a particular address_family. If -A is also present, show the contents of the internal Patricia tree structures; used for debugging. If -a is also present, show protocol-cloned routes (routes generated by an RTF_PRCLONING parent route); normally these routes are not shown. When -W is also present, show the path MTU for each route, and print interface names with a wider field size.

netstat -rs [-s] [-M core] [-N system]

Display routing statistics. If -s is repeated, counters with a value of zero are suppressed.

netstat -g [-W] [-f address_family] [-M core] [-N system]

Show information related to multicast (group address) routing. By default, show the IP Multicast virtual-interface and routing tables, and multicast group memberships.

netstat -gs [-s] [-f address_family] [-M core] [-N system]

Show multicast routing statistics. If -s is repeated, counters with a value of zero are suppressed.

Some options have the general meaning:

-f address_family, -p protocol

Limit display to those records of the specified address_family or a single protocol. The following address families and protocols are recognized:

| Family | Protocols |
|----------------------------|--|
| inet (AF_INET) | bdg, divert, icmp, igmp, ip, ipsec, pim, tcp, udp |
| inet6 (AF_INET6) | bdg, icmp6, ip6, ipsec6, rip6, tcp, udp |
| pfkey (PF_KEY) | pfkey |
| atalk (AF_APPLETALK) | ddp |
| netgraph, ng (AF_NETGRAPH) | ctrl, data |
| ipx (AF_IPX) | ipx, spx |
| unix (AF_UNIX) | |
| link (AF_LINK) | |

The program will complain if protocol is unknown or if there is no statistics routine for it.

- M** Extract values associated with the name list from the specified core instead of the default /dev/kmem.
- N** Extract the name list from the specified system instead of the default, which is the kernel image the system has booted from.
- n** Show network addresses and ports as numbers. Normally netstat attempts to resolve addresses and ports, and display them symbolically.
- W** In certain displays, avoid truncating addresses even if this causes some fields to overflow.

The default display, for active sockets, shows the local and remote addresses, send and receive queue sizes (in bytes), protocol, and the internal state of the protocol. Address formats are of the form ```host.port``` or ```network.port``` if a socket's address specifies a network but no specific host address. When known, the host and network addresses are displayed symbolically according to the databases `hosts(5)` and `networks(5)`, respectively. If a symbolic name for an address is unknown, or if the `-n` option is specified, the address is printed numerically, according to the address family. For more information regarding the Internet IPv4 ```dot format```, refer to `inet(3)`. Unspecified, or ```wildcard```, addresses and ports appear as ```*```.

The interface display provides a table of cumulative statistics regarding packets transferred, errors, and collisions. The network addresses of the interface and the maximum transmission unit (```mtu```) are also displayed.

The routing table display indicates the available routes and their status. Each route consists of a destination host or network, and a gateway to use in forwarding packets. The flags field shows a collection of information about the route stored as binary choices. The individual flags are discussed in more detail in the `route(8)` and `route(4)` manual pages. The mapping between letters and flags is:

- 1** RTF_PROTO1 Protocol specific routing flag #1
- 2** RTF_PROTO2 Protocol specific routing flag #2
- 3** RTF_PROTO3 Protocol specific routing flag #3
- B** RTF_BLACKHOLE Just discard pkts (during updates)
- b** RTF_BROADCAST The route represents a broadcast address
- C** RTF_CLONING Generate new routes on use
- c** RTF_PRCLONING Protocol-specified generate new routes on use
- D** RTF_DYNAMIC Created dynamically (by redirect)
- G** RTF_GATEWAY Destination requires forwarding by intermediary
- H** RTF_HOST Host entry (net otherwise)
- L** RTF_LLINFO Valid protocol to link address translation
- M** RTF_MODIFIED Modified dynamically (by redirect)

R RTF_REJECT Host or net unreachable
S RTF_STATIC Manually added
U RTF_UP Route usable
W RTF_WASCLONED Route was generated as a result of cloning
X RTF_XRESOLVE External daemon translates proto to link address

Direct routes are created for each interface attached to the local host; the gateway field for such entries shows the address of the outgoing interface. The refcnt field gives the current number of active uses of the route. Connection oriented protocols normally hold on to a single route for the duration of a connection while connectionless protocols obtain a route while sending to the same destination. The use field provides a count of the number of packets sent using that route. The interface entry indicates the network interface utilized for the route.

When netstat is invoked with the **-w** option and a wait interval argument, it displays a running count of statistics related to network interfaces. An obsolescent version of this option used a numeric parameter with no option, and is currently supported for backward compatibility. By default, this display summarizes information for all interfaces. Information for a specific interface may be displayed with the **-I** option. The **bpf(4)** flags displayed when netstat is invoked with the **-B** option represents the underlying parameters of the bpf peer. Each flag is represented as a single lower case letter. The mapping between the letters and flags in order of appearance are:

p Set if listening promiscuously
i BIOCIIMMEDIATE has been set on the device
f BIOCGHDRAMPLT status: source link addresses are being filled automatically
s BIOCGSEESSENT status: see packets originating locally and remotely on the interface.
a Packet reception generates a signal
l BIOLOCK status: descriptor has been locked

For more information about these flags, please refer to **bpf(4)**.

SEE ALSO

fstat(1), **nfsstat(1)**, **ps(1)**, **sockstat(1)**, **bpf(4)**, **inet(4)**, **route(4)**, **unix(4)**, **hosts(5)**, **networks(5)**, **protocols(5)**, **services(5)**, **iostat(8)**, **route(8)**, **trpt(8)**, **vmstat(8)**, **mbuf(9)**

HISTORY

The netstat command appeared in 4.2BSD.

IPv6 support was added by WIDE/KAME project.

BUGS

The notion of errors is ill-defined.

Câteva rezultate ale execuției aplicației de testate netstat sunt redată în tabelul de mai jos:

| Nr | Comanda | Rezultat |
|----|--------------|--|
| 1 | # netstat -a | <pre>Active Internet connections (including servers) Proto Recv-Q Send-Q Local Address Foreign Address (state) tcp4 0 52 db.ssh 172.27.72.1.2455 ESTABLISHED tcp4 0 0 *.ftp *.* LISTEN tcp6 0 0 *.ftp *.* LISTEN tcp4 0 0 *.ssh *.* LISTEN tcp6 0 0 *.ssh *.* LISTEN tcp4 0 0 db.http *.* LISTEN tcp4 0 0 *.3306 *.* LISTEN udp4 0 0 *.syslog *.* . udp6 0 0 *.syslog *.* . Active UNIX domain sockets Address Type Recv-Q Send-Q Inode Conn Refs Nextref Addr fffff006123e780 stream 0 0 0 fffff006123eb40 0 0 fffff006123eb40 stream 0 0 0 fffff006123e780 0 0 fffff006123ea80 stream 0 0 fffff005dc5f9b0 0 0 0 /tmp/mysql.sock fffff006103f000 stream 0 0 fffff006135d3e0 0 0 0 /var/run/devd.pipe fffff006123e840 dgram 0 0 0 fffff006123ec00 0 0 fffff006123ec00 dgram 0 0 fffff00605703e0 0 fffff006123e840 0 /var/run/logpriv fffff006123ecc0 dgram 0 0 fffff00605705d0 0 0 0 /var/run/log</pre> |
| 2 | # netstat -i | <pre>Name Mtu Network Address Ipkts Ierrs Opkts Oerrs Coll dc0 1500 <Link#1> 00:15:58:68:cf:7f 28038 0 20648 0 0 dc0 1500 fe80:1::215:5 fe80:1::215:58ff: 0 - 5 - - dc0 1500 172.27.72/25 db 26359 - 20591 - - lo0 16384 <Link#2> 46 0 46 0 0 lo0 16384 ::1 ::1 0 - 0 - - lo0 16384 fe80:2::1 fe80:2::1 0 - 0 - - lo0 16384 your-net 127.0.0.1 46 - 46 - -</pre> |
| 3 | # netstat -m | <pre>2/268/270 mbufs in use (current/cache/total) 0/134/134/25600 mbuf clusters in use (current/cache/total/max) 0/128 mbuf+clusters out of packet secondary zone in use (current/cache) 0/0/0/0 4k (page size) jumbo clusters in use (current/cache/total/max) 0/0/0/0 9k jumbo clusters in use (current/cache/total/max) 0/0/0/0 16k jumbo clusters in use (current/cache/total/max) 0K/335K/335K bytes allocated to network (current/cache/total) 0/0/0 requests for mbufs denied (mbufs/clusters/mbuf+clusters) 0/0/0 requests for jumbo clusters denied (4k/9k/16k) 0/0/0 sbufs in use (current/peak/max) 0 requests for sbufs denied 0 requests for sbufs delayed 39 requests for I/O initiated by sendfile 0 calls to protocol drain routines</pre> |

Configurare sisteme server

Sistemele server au fost configurate pentru a funcționa optim. Configurarea a inclus configurarea nucleului sistemului de operare (kernel), a serviciului de baze de date (mysql), a serviciului de web

(apache), a interpretorului de limbaj de nivel înalt (php) și a altor servicii conexe acestora. În continuare sunt redate doar parte din acestea:

| kernel | |
|-------------------|---|
| machine | amd64 |
| cpu | HAMMER |
| ident | SORI |
| options | IPI_PREEMPTION |
| options | SMP |
| device | mptable |
| device | speaker |
| device | cpufreq |
| options | ACPI_NO_SEMAPHORES |
| options | ACPICA_PEDANTIC |
| device | io |
| options | MAXCONS=4 |
| options | SC_DISABLE_REBOOT |
| options | SC_NO_SYSMOUSE |
| device | sound |
| device | snd_ich |
| options | SCHED_4BSD # 4BSD scheduler |
| options | PREEMPTION # Enable kernel thread preemption |
| options | INET # InterNETworking |
| options | INET6 # IPv6 communications protocols |
| options | FFS # Berkeley Fast Filesystem |
| options | SOFTUPDATES # Enable FFS soft updates support |
| options | UFS_ACL # Support for access control lists |
| options | UFS_DIRHASH # Improve performance on big directories |
| options | MD_ROOT # MD is a potential root device |
| options | NFSCIENT # Network Filesystem Client |
| options | NFS_ROOT # NFS usable as /, requires NFSCIENT |
| options | NTFS # NT File System |
| options | PROCFS # Process filesystem (requires PSEUDOFS) |
| options | PSEUDOFS # Pseudo-filesystem framework |
| options | GEOM_GPT # GUID Partition Tables. |
| #options | COMPAT_43 # Needed by COMPAT_LINUX32 |
| #options | COMPAT_IA32 # Compatible with i386 binaries |
| #options | COMPAT_FREEBSD4 # Compatible with FreeBSD4 |
| #options | COMPAT_FREEBSD5 # Compatible with FreeBSD5 |
| options | KTRACE # ktrace(1) support |
| options | SYSVSHM # SYSV-style shared memory |
| options | SYSVMSG # SYSV-style message queues |
| options | SYSVSEM # SYSV-style semaphores |
| options | _KPOSIX_PRIORITY_SCHEDULING # POSIX P1003_1B real-time extensions |
| options | ADAPTIVE_GIANT # Giant mutex is adaptive. |
| device | acpi |
| device | pci |
| device | ata |
| device | atadisk # ATA disk drives |
| device | atkbdc # AT keyboard controller |
| device | atkbd # AT keyboard |
| device | vga # VGA video card driver |
| device | sc |
| device | miibus # MII bus support |
| device | dc # DEC/Intel 21143 and various workalikes |
| #device | fxp # Intel EtherExpress PRO/100B (82557, 82558) |
| #device | re # RealTek 8139C+/8169/8169S/8110S |
| #device | rl # RealTek 8129/8139 |
| #device | xl # 3Com 3c90x ("Boomerang", "Cyclone") |
| # Pseudo devices. | |
| device | loop # Network loopback |
| device | random # Entropy device |
| device | ether # Ethernet support |
| device | sl # Kernel SLIP |
| device | ppp # Kernel PPP |
| device | tun # Packet tunnel. |
| device | pty # Pseudo-ttys (telnet etc) |
| device | md # Memory "disks" |

| | | |
|--|-------|---------------------------------------|
| device | gif | # IPv6 and IPv4 tunneling |
| device | faith | # IPv6-to-IPv4 relaying (translation) |
| mysql | | |
| <pre> # Example mysql config file for very large systems. # This is for large system with memory of 1G-2G where the system runs mainly # MySQL. # You can copy this file to # /etc/my.cnf to set global options, # mysql-data-dir/my.cnf to set server-specific options (in this # installation this directory is /var/db/mysql) or # ~/.my.cnf to set user-specific options. # One can in this file use all long options that the program supports. # If you want to know which options a program support, run the program # with --help option. # The following options will be passed to all MySQL clients [client] #password = your_password port = 3306 socket = /tmp/mysql.sock # Here follows entries for some specific programs # The MySQL server [mysqld] port = 3306 socket = /tmp/mysql.sock skip-locking set-variable = key_buffer=384M set-variable = max_allowed_packet=128M set-variable = table_cache=512 set-variable = sort_buffer=2M set-variable = record_buffer=2M set-variable = thread_cache=8 # Try number of CPU's*2 for thread_concurrency set-variable = thread_concurrency=4 set-variable = myisam_sort_buffer_size=64M #log-bin server-id = 1 # Point the following paths to different dedicated disks #tmpdir = /tmp/ #log-update = /path-to-dedicated-directory/hostname # Uncomment the following if you are using BDB tables #set-variable = bdb_cache_size=384M #set-variable = bdb_max_lock=100000 # Uncomment the following if you are using InnoDB tables #innodb_data_home_dir = /var/db/mysql/ #innodb_data_file_path = ibdata1:2000M;ibdata2:10M:autoextend #innodb_log_group_home_dir = /var/db/mysql/ #innodb_log_arch_dir = /var/db/mysql/ # You can set .._buffer_pool_size up to 50 - 80 % # of RAM but beware of setting memory usage too high #set-variable = innodb_buffer_pool_size=384M #set-variable = innodb_additional_mem_pool_size=20M # Set .._log_file_size to 25 % of buffer pool size #set-variable = innodb_log_file_size=100M #set-variable = innodb_log_buffer_size=8M #innodb_flush_log_at_trx_commit=1 #set-variable = innodb_lock_wait_timeout=50 [mysqldump] quick set-variable = max_allowed_packet=16M [mysql] no-auto-rehash # Remove the next comment character if you are not familiar with SQL #safe-updates [isamchk] set-variable = key_buffer=256M set-variable = sort_buffer=256M set-variable = read_buffer=2M set-variable = write_buffer=2M [myisamchk] </pre> | | |

```
set-variable = key_buffer=256M
set-variable = sort_buffer=256M
set-variable = read_buffer=2M
set-variable = write_buffer=2M
[mysqlhotcopy]
interactive-timeout
```

apache

```
## httpd.conf -- Apache HTTP server configuration file
# Based upon the NCSA server configuration files originally by Rob McCool.
# This is the main Apache server configuration file. It contains the
# configuration directives that give the server its instructions.
# See <URL:http://httpd.apache.org/docs/> for detailed information about
# the directives.
# Do NOT simply read the instructions in here without understanding
# what they do. They're here only as hints or reminders. If you are unsure
# consult the online docs. You have been warned.
# After this file is processed, the server will look for and process
# /usr/local/etc/apache/srm.conf and then /usr/local/etc/apache/access.conf
# unless you have overridden these with ResourceConfig and/or
# AccessConfig directives here.
# The configuration directives are grouped into three basic sections:
# 1. Directives that control the operation of the Apache server process as a
# whole (the 'global environment').
# 2. Directives that define the parameters of the 'main' or 'default' server,
# which responds to requests that aren't handled by a virtual host.
# These directives also provide default values for the settings
# of all virtual hosts.
# 3. Settings for virtual hosts, which allow Web requests to be sent to
# different IP addresses or hostnames and have them handled by the
# same Apache server process.
# Configuration and logfile names: If the filenames you specify for many
# of the server's control files begin with "/" (or "drive:/" for Win32), the
# server will use that explicit path. If the filenames do *not* begin
# with "/", the value of ServerRoot is prepended -- so "logs/foo.log"
# with ServerRoot set to "/usr/local/apache" will be interpreted by the
# server as "/usr/local/apache/logs/foo.log".
#### Section 1: Global Environment
# The directives in this section affect the overall operation of Apache,
# such as the number of concurrent requests it can handle or where it
# can find its configuration files.
# ServerType is either inetd, or standalone. Inetd mode is only supported on
# Unix platforms.
ServerType standalone
# ServerRoot: The top of the directory tree under which the server's
# configuration, error, and log files are kept.
#
# NOTE! If you intend to place this on an NFS (or otherwise network)
# mounted filesystem then please read the LockFile documentation
# (available at <URL:http://www.apache.org/docs/mod/core.html#lockfile>);
# you will save yourself a lot of trouble.
ServerRoot "/usr/local"
# The LockFile directive sets the path to the lockfile used when Apache
# is compiled with either USE_FCNTL_SERIALIZED_ACCEPT or
# USE_FLOCK_SERIALIZED_ACCEPT. This directive should normally be left at
# its default value. The main reason for changing it is if the logs
# directory is NFS mounted, since the lockfile MUST BE STORED ON A LOCAL
# DISK. The PID of the main server process is automatically appended to
# the filename.
#LockFile /var/run/httpd.lock
# PidFile: The file in which the server should record its process
# identification number when it starts.
PidFile /var/run/httpd.pid
# ScoreBoardFile: File used to store internal server process information.
# Not all architectures require this. But if yours does (you'll know because
# this file will be created when you run Apache) then you *must* ensure that
# no two invocations of Apache share the same scoreboard file.
ScoreBoardFile /var/run/httpd.scoreboard
# In the standard configuration, the server will process httpd.conf (this
# file, specified by the -f command line option), srm.conf, and access.conf
```

```

# in that order. The latter two files are now distributed empty, as it is
# recommended that all directives be kept in a single file for simplicity.
# The commented-out values below are the built-in defaults. You can have the
# server ignore these files altogether by using "/dev/null" (for Unix) or
# "nul" (for Win32) for the arguments to the directives.
#
#ResourceConfig /usr/local/etc/apache/srm.conf
#AccessConfig /usr/local/etc/apache/access.conf
# Timeout: The number of seconds before receives and sends time out.
Timeout 300
# KeepAlive: Whether or not to allow persistent connections (more than
# one request per connection). Set to "Off" to deactivate.
KeepAlive On
# MaxKeepAliveRequests: The maximum number of requests to allow
# during a persistent connection. Set to 0 to allow an unlimited amount.
# We recommend you leave this number high, for maximum performance.
MaxKeepAliveRequests 100
# KeepAliveTimeout: Number of seconds to wait for the next request from the
# same client on the same connection.
KeepAliveTimeout 15
# Server-pool size regulation. Rather than making you guess how many
# server processes you need, Apache dynamically adapts to the load it
# sees --- that is, it tries to maintain enough server processes to
# handle the current load, plus a few spare servers to handle transient
# load spikes (e.g., multiple simultaneous requests from a single
# Netscape browser).
# It does this by periodically checking how many servers are waiting
# for a request. If there are fewer than MinSpareServers, it creates
# a new spare. If there are more than MaxSpareServers, some of the
# spares die off. The default values are probably OK for most sites.
MinSpareServers 3
MaxSpareServers 6
# Number of servers to start initially --- should be a reasonable ballpark
# figure.
StartServers 3
# Limit on total number of servers running, i.e., limit on the number
# of clients who can simultaneously connect --- if this limit is ever
# reached, clients will be LOCKED OUT, so it should NOT BE SET TOO LOW.
# It is intended mainly as a brake to keep a runaway server from taking
# the system with it as it spirals down...
MaxClients 100
# MaxRequestsPerChild: the number of requests each child process is
# allowed to process before the child dies. The child will exit so
# as to avoid problems after prolonged use when Apache (and maybe the
# libraries it uses) leak memory or other resources. On most systems, this
# isn't really needed, but a few (such as Solaris) do have notable leaks
# in the libraries. For these platforms, set to something like 10000
# or so; a setting of 0 means unlimited.
# NOTE: This value does not include keepalive requests after the initial
# request per connection. For example, if a child process handles
# an initial request and 10 subsequent "keptalive" requests, it
# would only count as 1 request towards this limit.
MaxRequestsPerChild 0
# Listen: Allows you to bind Apache to specific IP addresses and/or
# ports, instead of the default. See also the <VirtualHost>
# directive.
#Listen 3000
Listen 172.27.211.1:80
# BindAddress: You can support virtual hosts with this option. This directive
# is used to tell the server which IP address to listen to. It can either
# contain "*", an IP address, or a fully qualified Internet domain name.
# See also the <VirtualHost> and Listen directives.
BindAddress 172.27.211.1
# Dynamic Shared Object (DSO) Support
# To be able to use the functionality of a module which was built as a DSO you
# have to place corresponding 'LoadModule' lines at this location so the
# directives contained in it are actually available _before_ they are used.
# Please read the file http://httpd.apache.org/docs/dso.html for more
# details about the DSO mechanism and run 'httpd -l' for the list of already
# built-in (statically linked and thus always available) modules in your httpd

```

```

# binary.
# Note: The order in which modules are loaded is important. Don't change
# the order below without expert advice.
# Example:
# LoadModule foo_module libexec/mod_foo.so
LoadModule mmap_static_module libexec/apache/mod_mmap_static.so
LoadModule vhost_alias_module libexec/apache/mod_vhost_alias.so
LoadModule env_module libexec/apache/mod_env.so
LoadModule config_log_module libexec/apache/mod_log_config.so
LoadModule mime_magic_module libexec/apache/mod_mime_magic.so
LoadModule mime_module libexec/apache/mod_mime.so
LoadModule negotiation_module libexec/apache/mod_negotiation.so
LoadModule status_module libexec/apache/mod_status.so
LoadModule info_module libexec/apache/mod_info.so
LoadModule includes_module libexec/apache/mod_include.so
LoadModule autoindex_module libexec/apache/mod_autoindex.so
LoadModule dir_module libexec/apache/mod_dir.so
LoadModule cgi_module libexec/apache/mod_cgi.so
LoadModule asis_module libexec/apache/mod_asis.so
LoadModule imap_module libexec/apache/mod_imap.so
LoadModule action_module libexec/apache/mod_actions.so
LoadModule speling_module libexec/apache/mod_speling.so
LoadModule userdir_module libexec/apache/mod_userdir.so
LoadModule alias_module libexec/apache/mod_alias.so
LoadModule rewrite_module libexec/apache/mod_rewrite.so
LoadModule access_module libexec/apache/mod_access.so
LoadModule auth_module libexec/apache/mod_auth.so
LoadModule anon_auth_module libexec/apache/mod_auth_anon.so
LoadModule db_auth_module libexec/apache/mod_auth_db.so
LoadModule digest_module libexec/apache/mod_digest.so
LoadModule proxy_module libexec/apache/libproxy.so
LoadModule cern_meta_module libexec/apache/mod_cern_meta.so
LoadModule expires_module libexec/apache/mod_expires.so
LoadModule headers_module libexec/apache/mod_headers.so
LoadModule usertrack_module libexec/apache/mod_usertrack.so
LoadModule log_forensic_module libexec/apache/mod_log_forensic.so
LoadModule unique_id_module libexec/apache/mod_unique_id.so
LoadModule setenvif_module libexec/apache/mod_setenvif.so
LoadModule php5_module libexec/apache/libphp5.so
# Reconstruction of the complete module list from all available modules
# (static and shared ones) to achieve correct module execution order.
# [WHENEVER YOU CHANGE THE LOADMODULE SECTION ABOVE UPDATE THIS, TOO]
ClearModuleList
AddModule mod_mmap_static.c
AddModule mod_vhost_alias.c
AddModule mod_env.c
AddModule mod_log_config.c
AddModule mod_mime_magic.c
AddModule mod_mime.c
AddModule mod_negotiation.c
AddModule mod_status.c
AddModule mod_info.c
AddModule mod_include.c
AddModule mod_autoindex.c
AddModule mod_dir.c
AddModule mod_cgi.c
AddModule mod_asis.c
AddModule mod_imap.c
AddModule mod_actions.c
AddModule mod_speling.c
AddModule mod_userdir.c
AddModule mod_alias.c
AddModule mod_rewrite.c
AddModule mod_access.c
AddModule mod_auth.c
AddModule mod_auth_anon.c
AddModule mod_auth_db.c
AddModule mod_digest.c
AddModule mod_proxy.c
AddModule mod_cern_meta.c

```

```

AddModule mod_expires.c
AddModule mod_headers.c
AddModule mod_usertrack.c
AddModule mod_log_forensic.c
AddModule mod_unique_id.c
AddModule mod_so.c
AddModule mod_setenvif.c
AddModule mod_php5.c
# ExtendedStatus controls whether Apache will generate "full" status
# information (ExtendedStatus On) or just basic information (ExtendedStatus
# Off) when the "server-status" handler is called. The default is Off.
#ExtendedStatus On
#### Section 2: 'Main' server configuration
# The directives in this section set up the values used by the 'main'
# server, which responds to any requests that aren't handled by a
# <VirtualHost> definition. These values also provide defaults for
# any <VirtualHost> containers you may define later in the file.
# All of these directives may appear inside <VirtualHost> containers,
# in which case these default settings will be overridden for the
# virtual host being defined.
# If your ServerType directive (set earlier in the 'Global Environment'
# section) is set to "inetd", the next few directives don't have any
# effect since their settings are defined by the inetd configuration.
# Skip ahead to the ServerAdmin directive.
# Port: The port to which the standalone server listens. For
# ports < 1023, you will need httpd to be run as root initially.
Port 80
# If you wish httpd to run as a different user or group, you must run
# httpd as root initially and it will switch.
#
# User/Group: The name (or #number) of the user/group to run httpd as.
# . On SCO (ODT 3) use "User nouser" and "Group nogroup".
# . On HP-UX you may not be able to use shared memory as nobody, and the
# suggested workaround is to create a user www and use that user.
# NOTE that some kernels refuse to setgid(Group) or semctl(IPC_SET)
# when the value of (unsigned)Group is above 60000;
# don't use Group "#-1" on these systems!
#
User www
Group www
# ServerAdmin: Your address, where problems with the server should be
# e-mailed. This address appears on some server-generated pages, such
# as error documents.
ServerAdmin sorana@j.academicdirect.ro
# ServerName allows you to set a host name which is sent back to clients for
# your server if it's different than the one the program would get (i.e., use
# "www" instead of the host's real name).
# Note: You cannot just invent host names and hope they work. The name you
# define here must be a valid DNS name for your host. If you don't understand
# this, ask your network administrator.
# If your host doesn't have a registered DNS name, enter its IP address here.
# You will have to access it by its address (e.g., http://123.45.67.89/)
# anyway, and this will make redirections work in a sensible way.
# 127.0.0.1 is the TCP/IP local loop-back address, often named localhost. Your
# machine always knows itself by this address. If you use Apache strictly for
# local testing and development, you may use 127.0.0.1 as the server name.
ServerName 172.27.211.1
# DocumentRoot: The directory out of which you will serve your
# documents. By default, all requests are taken from this directory, but
# symbolic links and aliases may be used to point to other locations.
DocumentRoot "/usr/home/www/data"
# Each directory to which Apache has access, can be configured with respect
# to which services and features are allowed and/or disabled in that
# directory (and its subdirectories).
# First, we configure the "default" to be a very restrictive set of
# permissions.
<Directory />
    Options FollowSymLinks
    AllowOverride None
</Directory>

```

```

# Note that from this point forward you must specifically allow
# particular features to be enabled - so if something's not working as
# you might expect, make sure that you have specifically enabled it
# below.
# This should be changed to whatever you set DocumentRoot to.
<Directory "/usr/home/www/data">
# This may also be "None", "All", or any combination of "Indexes",
# "Includes", "FollowSymLinks", "ExecCGI", or "MultiViews".
# Note that "MultiViews" must be named *explicitly* --- "Options All"
# doesn't give it to you.
  Options Indexes FollowSymLinks MultiViews
# This controls which options the .htaccess files in directories can
# override. Can also be "All", or any combination of "Options", "FileInfo",
# "AuthConfig", and "Limit"
  AllowOverride None
# Controls who can get stuff from this server.
  Order allow,deny
  Allow from all
</Directory>
# UserDir: The name of the directory which is appended onto a user's home
# directory if a ~user request is received.
<IfModule mod_userdir.c>
  UserDir html
</IfModule>
# Control access to UserDir directories. The following is an example
# for a site where these directories are restricted to read-only.
<Directory /home/*/html>
  AllowOverride FileInfo AuthConfig Limit
  Options MultiViews Indexes SymLinksIfOwnerMatch IncludesNoExec
  <Limit GET POST OPTIONS PROPFIND>
    Order allow,deny
    Allow from all
  </Limit>
  <LimitExcept GET POST OPTIONS PROPFIND>
    Order deny,allow
    Deny from all
  </LimitExcept>
</Directory>
# DirectoryIndex: Name of the file or files to use as a pre-written HTML
# directory index. Separate multiple entries with spaces.
<IfModule mod_dir.c>
  <IfModule mod_php5.c>
    DirectoryIndex index.php index.html
  </IfModule>
</IfModule>
<IfModule mod_dir.c>
  <IfModule mod_php3.c>
    <IfModule mod_php4.c>
      DirectoryIndex index.php index.php3 index.html
    </IfModule>
    <IfModule !mod_php4.c>
      DirectoryIndex index.php3 index.html
    </IfModule>
  </IfModule>
  <IfModule !mod_php3.c>
    <IfModule mod_php4.c>
      DirectoryIndex index.php index.html
    </IfModule>
    <IfModule !mod_php4.c>
      DirectoryIndex index.html
    </IfModule>
  </IfModule>
</IfModule>
# AccessFileName: The name of the file to look for in each directory
# for access control information.
AccessFileName .htaccess
# The following lines prevent .htaccess files from being viewed by
# Web clients. Since .htaccess files often contain authorization
# information, access is disallowed for security reasons. Comment
# these lines out if you want Web visitors to see the contents of

```

```

# .htaccess files. If you change the AccessFileName directive above,
# be sure to make the corresponding changes here.
# Also, folks tend to use names such as .htpasswd for password
# files, so this will protect those as well.
<Files ~ "^\.ht">
    Order allow,deny
    Deny from all
    Satisfy All
</Files>
# CacheNegotiatedDocs: By default, Apache sends "Pragma: no-cache" with each
# document that was negotiated on the basis of content. This asks proxy
# servers not to cache the document. Uncommenting the following line disables
# this behavior, and proxies will be allowed to cache the documents.
#CacheNegotiatedDocs
# UseCanonicalName: (new for 1.3) With this setting turned on, whenever
# Apache needs to construct a self-referencing URL (a URL that refers back
# to the server the response is coming from) it will use ServerName and
# Port to form a "canonical" name. With this setting off, Apache will
# use the hostname:port that the client supplied, when possible. This
# also affects SERVER_NAME and SERVER_PORT in CGI scripts.
UseCanonicalName Off
# TypesConfig describes where the mime.types file (or equivalent) is
# to be found.
<IfModule mod_mime.c>
    TypesConfig /usr/local/etc/apache/mime.types
</IfModule>
# DefaultType is the default MIME type the server will use for a document
# if it cannot otherwise determine one, such as from filename extensions.
# If your server contains mostly text or HTML documents, "text/plain" is
# a good value. If most of your content is binary, such as applications
# or images, you may want to use "application/octet-stream" instead to
# keep browsers from trying to display binary files as though they are
# text.
DefaultType text/plain
# The mod_mime_magic module allows the server to use various hints from the
# contents of the file itself to determine its type. The MIMEMagicFile
# directive tells the module where the hint definitions are located.
# mod_mime_magic is not part of the default server (you have to add
# it yourself with a LoadModule [see the DSO paragraph in the 'Global
# Environment' section], or recompile the server and include mod_mime_magic
# as part of the configuration), so it's enclosed in an <IfModule> container.
# This means that the MIMEMagicFile directive will only be processed if the
# module is part of the server.
<IfModule mod_mime_magic.c>
    MIMEMagicFile /usr/local/etc/apache/magic
</IfModule>
# HostnameLookups: Log the names of clients or just their IP addresses
# e.g., www.apache.org (on) or 204.62.129.132 (off).
# The default is off because it'd be overall better for the net if people
# had to knowingly turn this feature on, since enabling it means that
# each client request will result in AT LEAST one lookup request to the
# nameserver.
HostnameLookups Off
# ErrorLog: The location of the error log file.
# If you do not specify an ErrorLog directive within a <VirtualHost>
# container, error messages relating to that virtual host will be
# logged here. If you *do* define an error logfile for a <VirtualHost>
# container, that host's errors will be logged there and not here.
ErrorLog /var/log/httpd-error.log
# LogLevel: Control the number of messages logged to the error_log.
# Possible values include: debug, info, notice, warn, error, crit,
# alert, emerg.
LogLevel error
# The following directives define some format nicknames for use with
# a CustomLog directive (see below).
LogFormat "%h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-Agent}i\"" combined
LogFormat "%h %l %u %t \"%r\" %>s %b" common
LogFormat "%{Referer}i -> %U" referer
LogFormat "%{User-agent}i" agent
# The location and format of the access logfile (Common Logfile Format).

```

```

# If you do not define any access logfiles within a <VirtualHost>
# container, they will be logged here. Contrariwise, if you *do*
# define per-<VirtualHost> access logfiles, transactions will be
# logged therein and *not* in this file.
#CustomLog /var/log/httpd-access.log common
# If you would like to have agent and referer logfiles, uncomment the
# following directives.
#CustomLog /var/log/httpd-referer.log referer
#CustomLog /var/log/httpd-agent.log agent
# If you prefer a single logfile with access, agent, and referer information
# (Combined Logfile Format) you can use the following directive.
CustomLog /var/log/httpd-access.log combined
# Optionally add a line containing the server version and virtual host
# name to server-generated pages (error documents, FTP directory listings,
# mod_status and mod_info output etc., but not CGI generated documents).
# Set to "EMail" to also include a mailto: link to the ServerAdmin.
# Set to one of: On | Off | EMail
ServerSignature Off
# EBCDIC configuration:
# (only for mainframes using the EBCDIC codeset, currently one of:
# Fujitsu-Siemens' BS2000/OSD, IBM's OS/390 and IBM's TPF)!!
# The following default configuration assumes that "text files"
# are stored in EBCDIC (so that you can operate on them using the
# normal POSIX tools like grep and sort) while "binary files" are
# stored with identical octets as on an ASCII machine.
# The directives are evaluated in configuration file order, with
# the EBCDICConvert directives applied before EBCDICConvertByType.
# If you want to have ASCII HTML documents and EBCDIC HTML documents
# at the same time, you can use the file extension to force
# conversion off for the ASCII documents:
# > AddType text/html .ahtml
# > EBCDICConvert Off=InOut .ahtml
# EBCDICConvertByType On=InOut text/* message/* multipart/*
# EBCDICConvertByType On=In application/x-www-form-urlencoded
# EBCDICConvertByType On=InOut application/postscript model/vrml
# EBCDICConvertByType Off=InOut */*
# Aliases: Add here as many aliases as you need (with no limit). The format is
# Alias fakename realname
<IfModule mod_alias.c>
# Note that if you include a trailing / on fakename then the server will
# require it to be present in the URL. So "/icons" isn't aliased in this
# example, only "/icons/". If the fakename is slash-terminated, then the
# realname must also be slash terminated, and if the fakename omits the
# trailing slash, the realname must also omit it.
Alias /icons/ "/usr/home/www/icons/"
<Directory "/usr/home/www/icons">
    Options Indexes MultiViews
    AllowOverride None
    Order allow,deny
    Allow from all
</Directory>
# This Alias will project the on-line documentation tree under /manual/
# even if you change the DocumentRoot. Comment it if you don't want to
# provide access to the on-line documentation.
Alias /manual/ "/usr/local/share/doc/apache/"
<Directory "/usr/local/share/doc/apache">
    Options Indexes FollowSymlinks MultiViews
    AllowOverride None
    Order allow,deny
    Allow from all
</Directory>
# ScriptAlias: This controls which directories contain server scripts.
# ScriptAliases are essentially the same as Aliases, except that
# documents in the realname directory are treated as applications and
# run by the server when requested rather than as documents sent to the client.
# The same rules about trailing "/" apply to ScriptAlias directives as to
# Alias.
ScriptAlias /cgi-bin/ "/usr/home/www/cgi-bin/"
# "/usr/local/www/cgi-bin" should be changed to whatever your ScriptAliased
# CGI directory exists, if you have that configured.

```

```

<Directory "/usr/home/www/cgi-bin">
  AllowOverride None
  Options None
  Order allow,deny
  Allow from all
</Directory>
</IfModule>
# End of aliases.
# Redirect allows you to tell clients about documents which used to exist in
# your server's namespace, but do not anymore. This allows you to tell the
# clients where to look for the relocated document.
# Format: Redirect old-URI new-URL
# Directives controlling the display of server-generated directory listings.
<IfModule mod_autoindex.c>
  # FancyIndexing is whether you want fancy directory indexing or standard
  IndexOptions FancyIndexing
  # AddIcon* directives tell the server which icon to show for different
  # files or filename extensions. These are only displayed for
  # FancyIndexed directories.
  AddIconByEncoding (CMP,/icons/compressed.gif) x-compress x-gzip
  AddIconByType (TXT,/icons/text.gif) text/*
  AddIconByType (IMG,/icons/image2.gif) image/*
  AddIconByType (SND,/icons/sound2.gif) audio/*
  AddIconByType (VID,/icons/movie.gif) video/*
  AddIcon /icons/binary.gif .bin .exe
  AddIcon /icons/binhex.gif .hqx
  AddIcon /icons/tar.gif .tar
  AddIcon /icons/world2.gif .wrl .wrl.gz .vrml .vrm .iv
  AddIcon /icons/compressed.gif .Z .z .tgz .gz .zip
  AddIcon /icons/a.gif .ps .ai .eps
  AddIcon /icons/layout.gif .html .shtml .htm .pdf
  AddIcon /icons/text.gif .txt
  AddIcon /icons/c.gif .c
  AddIcon /icons/p.gif .pl .py
  AddIcon /icons/f.gif .for
  AddIcon /icons/dvi.gif .dvi
  AddIcon /icons/uuencoded.gif .uu
  AddIcon /icons/script.gif .conf .sh .shar .csh .ksh .tcl
  AddIcon /icons/tex.gif .tex
  AddIcon /icons/bomb.gif core
  AddIcon /icons/back.gif ..
  AddIcon /icons/hand.right.gif README
  AddIcon /icons/folder.gif ^^DIRECTORY^^
  AddIcon /icons/blank.gif ^^BLANKICON^^
  # DefaultIcon is which icon to show for files which do not have an icon
  # explicitly set.
  DefaultIcon /icons/unknown.gif
  # AddDescription allows you to place a short description after a file in
  # server-generated indexes. These are only displayed for FancyIndexed
  # directories.
  # Format: AddDescription "description" filename
  #AddDescription "GZIP compressed document" .gz
  #AddDescription "tar archive" .tar
  #AddDescription "GZIP compressed tar archive" .tgz
  # ReadmeName is the name of the README file the server will look for by
  # default, and append to directory listings.
  # HeaderName is the name of a file which should be prepended to
  # directory indexes.
  ReadmeName README.html
  HeaderName HEADER.html
  # IndexIgnore is a set of filenames which directory indexing should ignore
  # and not include in the listing. Shell-style wildcarding is permitted.
  IndexIgnore .??.* ~*# HEADER* README* RCS CVS *,v *,t
</IfModule>
# End of indexing directives.
# Document types.
<IfModule mod_mime.c>
  # AddLanguage allows you to specify the language of a document. You can
  # then use content negotiation to give a browser a file in a language
  # it can understand.

```

```

# Note 1: The suffix does not have to be the same as the language
# keyword --- those with documents in Polish (whose net-standard
# language code is pl) may wish to use "AddLanguage pl.po" to
# avoid the ambiguity with the common suffix for perl scripts.
# Note 2: The example entries below illustrate that in quite
# some cases the two character 'Language' abbreviation is not
# identical to the two character 'Country' code for its country,
# E.g. 'Danmark/dk' versus 'Danish/da'.
# Note 3: In the case of 'ltz' we violate the RFC by using a three char
# specifier. But there is 'work in progress' to fix this and get
# the reference data for rfc1766 cleaned up.
# Danish (da) - Dutch (nl) - English (en) - Estonian (ee)
# French (fr) - German (de) - Greek-Modern (el)
# Italian (it) - Korean (kr) - Norwegian (no) - Norwegian Nynorsk (nn)
# Portugese (pt) - Luxembourgish* (ltz)
# Spanish (es) - Swedish (sv) - Catalan (ca) - Czech(cs)
# Polish (pl) - Brazilian Portuguese (pt-br) - Japanese (ja)
# Russian (ru)
AddLanguage da .dk
AddLanguage nl .nl
AddLanguage en .en
AddLanguage et .ee
AddLanguage fr .fr
AddLanguage de .de
AddLanguage el .el
AddLanguage he .he
AddCharset ISO-8859-8 .iso8859-8
AddLanguage it .it
AddLanguage ja .ja
AddCharset ISO-2022-JP .jis
AddLanguage kr .kr
AddCharset ISO-2022-KR .iso-kr
AddLanguage nn .nn
AddLanguage no .no
AddLanguage pl .po
AddCharset ISO-8859-2 .iso-pl
AddLanguage pt .pt
AddLanguage pt-br .pt-br
AddLanguage ltz .lu
AddLanguage ca .ca
AddLanguage es .es
AddLanguage sv .sv
AddLanguage cs .cz .cs
AddLanguage ru .ru
AddLanguage zh-TW .zh-tw
AddCharset Big5 .Big5 .big5
AddCharset WINDOWS-1251 .cp-1251
AddCharset CP866 .cp866
AddCharset ISO-8859-5 .iso-ru
AddCharset KOI8-R .koi8-r
AddCharset UCS-2 .ucs2
AddCharset UCS-4 .ucs4
AddCharset UTF-8 .utf8
# LanguagePriority allows you to give precedence to some languages
# in case of a tie during content negotiation.
# Just list the languages in decreasing order of preference. We have
# more or less alphabetized them here. You probably want to change this.
<IfModule mod_negotiation.c>
    LanguagePriority en da nl et fr de el it ja kr no pl pt pt-br ru ltz ca es sv tw
</IfModule>
<IfModule mod_php3.c>
    AddType application/x-httpd-php3 .php3
    AddType application/x-httpd-php3-source .php3s
</IfModule>
<IfModule mod_php4.c>
    AddType application/x-httpd-php .php
    AddType application/x-httpd-php-source .phps
</IfModule>
<IfModule mod_php5.c>
    AddType application/x-httpd-php .php

```

```

    AddType application/x-httpd-php-source .phps
</IfModule>
# AddType allows you to tweak mime.types without actually editing it, or to
# make certain files to be certain types.
AddType application/x-tar .tgz
# AddEncoding allows you to have certain browsers uncompress
# information on the fly. Note: Not all browsers support this.
# Despite the name similarity, the following Add* directives have nothing
# to do with the FancyIndexing customization directives above.
AddEncoding x-compress .Z
AddEncoding x-gzip .gz .tgz
# If the AddEncoding directives above are commented-out, then you
# probably should define those extensions to indicate media types:
#AddType application/x-compress .Z
#AddType application/x-gzip .gz .tgz
# AddHandler allows you to map certain file extensions to "handlers",
# actions unrelated to filetype. These can be either built into the server
# or added with the Action command (see below)
# If you want to use server side includes, or CGI outside
# ScriptAliased directories, uncomment the following lines.
# To use CGI scripts:
#AddHandler cgi-script .cgi
# To use server-parsed HTML files
#AddType text/html .shtml
#AddHandler server-parsed .shtml
# Uncomment the following line to enable Apache's send-asis HTTP file
# feature
#AddHandler send-as-is asis
# If you wish to use server-parsed imagemap files, use
#AddHandler imap-file map
# To enable type maps, you might want to use
#AddHandler type-map var
</IfModule>
# End of document types.
# Action lets you define media types that will execute a script whenever
# a matching file is called. This eliminates the need for repeated URL
# pathnames for oft-used CGI file processors.
# Format: Action media/type /cgi-script/location
# Format: Action handler-name /cgi-script/location
# MetaDir: specifies the name of the directory in which Apache can find
# meta information files. These files contain additional HTTP headers
# to include when sending the document
#MetaDir .web
# MetaSuffix: specifies the file name suffix for the file containing the
# meta information.
#MetaSuffix .meta
# Customizable error response (Apache style)
# these come in three flavors
# 1) plain text
#ErrorDocument 500 "The server made a boo boo."
# n.b. the single leading (") marks it as text, it does not get output
# 2) local redirects
#ErrorDocument 404 /missing.html
# to redirect to local URL /missing.html
#ErrorDocument 404 /cgi-bin/missing_handler.pl
# N.B.: You can redirect to a script or a document using server-side-includes.
# 3) external redirects
#ErrorDocument 402 http://www.example.com/subscription_info.html
# N.B.: Many of the environment variables associated with the original
# request will *not* be available to such a script.
# Customize behaviour based on the browser
<IfModule mod_setenvif.c>
# The following directives modify normal HTTP response behavior.
# The first directive disables keepalive for Netscape 2.x and browsers that
# spoof it. There are known problems with these browser implementations.
# The second directive is for Microsoft Internet Explorer 4.0b2
# which has a broken HTTP/1.1 implementation and does not properly
# support keepalive when it is used on 301 or 302 (redirect) responses.
BrowserMatch "Mozilla/2" nokeepalive
BrowserMatch "MSIE 4.0b2;" nokeepalive downgrade-1.0 force-response-1.0

```



```

; Implicit flush tells PHP to tell the output layer to flush itself
; automatically after every output block. This is equivalent to calling the
; PHP function flush() after each and every call to print() or echo() and each
; and every HTML block. Turning this option on has serious performance
; implications and is generally recommended for debugging purposes only.
implicit_flush = Off
; The unserialize callback function will be called (with the undefined class'
; name as parameter), if the unserializer finds an undefined class
; which should be instantiated.
; A warning appears if the specified function is not defined, or if the
; function doesn't include/implement the missing class.
; So only set this entry, if you really want to implement such a
; callback-function.
unserialize_callback_func=
; When floats & doubles are serialized store serialize_precision significant
; digits after the floating point. The default value ensures that when floats
; are decoded with unserialize, the data will remain the same.
serialize_precision = 100
; Whether to enable the ability to force arguments to be passed by reference
; at function call time. This method is deprecated and is likely to be
; unsupported in future versions of PHP/Zend. The encouraged method of
; specifying which arguments should be passed by reference is in the function
; declaration. You're encouraged to try and turn this option Off and make
; sure your scripts work properly with it in order to ensure they will work
; with future versions of the language (you will receive a warning each time
; you use this feature, and the argument will be passed by value instead of by
; reference).
allow_call_time_pass_reference = Off
; Safe Mode
; SECURITY NOTE: The FreeBSD Security Officer strongly recommend that
; the PHP Safe Mode feature not be relied upon for security, since the
; issues Safe Mode tries to handle cannot properly be handled in PHP
; (primarily due to PHP's use of external libraries). While many bugs
; in Safe Mode has been fixed it's very likely that more issues exist
; which allows a user to bypass Safe Mode restrictions.
; For increased security we recommend to always install the Suhosin
; extension.
safe_mode = Off
; By default, Safe Mode does a UID compare check when
; opening files. If you want to relax this to a GID compare,
; then turn on safe_mode_gid.
safe_mode_gid = Off
; When safe_mode is on, UID/GID checks are bypassed when
; including files from this directory and its subdirectories.
; (directory must also be in include_path or full path must
; be used when including)
safe_mode_include_dir =
; When safe_mode is on, only executables located in the safe_mode_exec_dir
; will be allowed to be executed via the exec family of functions.
safe_mode_exec_dir =
; Setting certain environment variables may be a potential security breach.
; This directive contains a comma-delimited list of prefixes. In Safe Mode,
; the user may only alter environment variables whose names begin with the
; prefixes supplied here. By default, users will only be able to set
; environment variables that begin with PHP_ (e.g. PHP_FOO=BAR).
; Note: If this directive is empty, PHP will let the user modify ANY
; environment variable!
safe_mode_allowed_env_vars = PHP_
; This directive contains a comma-delimited list of environment variables that
; the end user won't be able to change using putenv(). These variables will be
; protected even if safe_mode_allowed_env_vars is set to allow to change them.
safe_mode_protected_env_vars = LD_LIBRARY_PATH
; open_basedir, if set, limits all file operations to the defined directory
; and below. This directive makes most sense if used in a per-directory
; or per-virtualhost web server configuration file. This directive is
; *NOT* affected by whether Safe Mode is turned On or Off.
open_basedir =
; This directive allows you to disable certain functions for security reasons.
; It receives a comma-delimited list of function names. This directive is
; *NOT* affected by whether Safe Mode is turned On or Off.

```

```

disable_functions =
; This directive allows you to disable certain classes for security reasons.
; It receives a comma-delimited list of class names. This directive is
; *NOT* affected by whether Safe Mode is turned On or Off.
disable_classes =
; Colors for Syntax Highlighting mode. Anything that's acceptable in
; <span style="color: ???????"> would work.
;highlight.string = #DD0000
;highlight.comment = #FF9900
;highlight.keyword = #007700
;highlight.bg = #FFFFFF
;highlight.default = #0000BB
;highlight.html = #000000
; If enabled, the request will be allowed to complete even if the user aborts
; the request. Consider enabling it if executing long request, which may end up
; being interrupted by the user or a browser timing out.
; ignore_user_abort = On
; Determines the size of the realpath cache to be used by PHP. This value should
; be increased on systems where PHP opens many files to reflect the quantity of
; the file operations performed.
; realpath_cache_size=16k
; Duration of time, in seconds for which to cache realpath information for a given
; file or directory. For systems with rarely changing files, consider increasing this
; value.
; realpath_cache_ttl=120
; Misc
; Decides whether PHP may expose the fact that it is installed on the server
; (e.g. by adding its signature to the Web server header). It is no security
; threat in any way, but it makes it possible to determine whether you use PHP
; on your server or not.
expose_php = On
; Resource Limits ;
max_execution_time = 72000 ; Maximum execution time of each script, in seconds
max_input_time = 3600 ; Maximum amount of time each script may spend parsing request data
memory_limit = 384M ; Maximum amount of memory a script may consume (16MB)
; Error handling and logging ;
; error_reporting is a bit-field. Or each number up to get desired error
; reporting level
; E_ALL - All errors and warnings (doesn't include E_STRICT)
; E_ERROR - fatal run-time errors
; E_RECOVERABLE_ERROR - almost fatal run-time errors
; E_WARNING - run-time warnings (non-fatal errors)
; E_PARSE - compile-time parse errors
; E_NOTICE - run-time notices (these are warnings which often result
; from a bug in your code, but it's possible that it was
; intentional (e.g., using an uninitialized variable and
; relying on the fact it's automatically initialized to an
; empty string)
; E_STRICT - run-time notices, enable to have PHP suggest changes
; to your code which will ensure the best interoperability
; and forward compatibility of your code
; E_CORE_ERROR - fatal errors that occur during PHP's initial startup
; E_CORE_WARNING - warnings (non-fatal errors) that occur during PHP's
; initial startup
; E_COMPILE_ERROR - fatal compile-time errors
; E_COMPILE_WARNING - compile-time warnings (non-fatal errors)
; E_USER_ERROR - user-generated error message
; E_USER_WARNING - user-generated warning message
; E_USER_NOTICE - user-generated notice message
; Examples:
; - Show all errors, except for notices and coding standards warnings
;error_reporting = E_ALL & ~E_NOTICE
; - Show all errors, except for notices
;error_reporting = E_ALL & ~E_NOTICE | E_STRICT
; - Show only errors
;error_reporting = E_COMPILE_ERROR|E_RECOVERABLE_ERROR|E_ERROR|E_CORE_ERROR

```

```

; - Show all errors, except coding standards warnings
error_reporting = E_ALL
; Print out errors (as a part of the output). For production web sites,
; you're strongly encouraged to turn this feature off, and use error logging
; instead (see below). Keeping display_errors enabled on a production web site
; may reveal security information to end users, such as file paths on your Web
; server, your database schema or other information.
display_errors = On
; Even when display_errors is on, errors that occur during PHP's startup
; sequence are not displayed. It's strongly recommended to keep
; display_startup_errors off, except for when debugging.
display_startup_errors = On
; Log errors into a log file (server-specific log, stderr, or error_log (below))
; As stated above, you're strongly advised to use error logging in place of
; error displaying on production web sites.
log_errors = Off
; Set maximum length of log_errors. In error_log information about the source is
; added. The default is 1024 and 0 allows to not apply any maximum length at all.
log_errors_max_len = 1024
; Do not log repeated messages. Repeated errors must occur in same file on same
; line until ignore_repeated_source is set true.
ignore_repeated_errors = Off
; Ignore source of message when ignoring repeated messages. When this setting
; is On you will not log errors with repeated messages from different files or
; source lines.
ignore_repeated_source = Off
; If this parameter is set to Off, then memory leaks will not be shown (on
; stdout or in the log). This has only effect in a debug compile, and if
; error reporting includes E_WARNING in the allowed list
report_memleaks = On
report_zend_debug = 0
; Store the last error/warning message in $php_errormsg (boolean).
track_errors = Off
; Disable the inclusion of HTML tags in error messages.
; Note: Never use this feature for production boxes.
html_errors = Off
; If html_errors is set On PHP produces clickable error messages that direct
; to a page describing the error or function causing the error in detail.
; You can download a copy of the PHP manual from http://www.php.net/docs.php
; and change docref_root to the base URL of your local copy including the
; leading '/'. You must also specify the file extension being used including
; the dot.
; Note: Never use this feature for production boxes.
docref_root = "/phpmanual/"
docref_ext = .html
; String to output before an error message.
error_prepend_string = "<font color=ff0000>"
; String to output after an error message.
error_append_string = "</font>"
; Log errors to specified file.
error_log = filename
; Log errors to syslog (Event Log on NT, not valid in Windows 95).
error_log = syslog
; Data Handling ;
; Note - track_vars is ALWAYS enabled as of PHP 4.0.3
; The separator used in PHP generated URLs to separate arguments.
; Default is "&".
arg_separator.output = "&"
; List of separator(s) used by PHP to parse input URLs into variables.
; Default is "&".
; NOTE: Every character in this directive is considered as separator!
arg_separator.input = ";&"
; This directive describes the order in which PHP registers GET, POST, Cookie,
; Environment and Built-in variables (G, P, C, E & S respectively, often
; referred to as EGPCS or GPC). Registration is done from left to right, newer
; values override older values.
variables_order = "GPCS"
; Whether or not to register the EGPCS variables as global variables. You may

```

```

; want to turn this off if you don't want to clutter your scripts' global scope
; with user data. This makes most sense when coupled with track_vars - in which
; case you can access all of the GPC variables through the $HTTP_*_VARS[],
; variables.
; You should do your best to write your scripts so that they do not require
; register_globals to be on; Using form variables as globals can easily lead
; to possible security problems, if the code is not very well thought of.
register_globals = Off
; Whether or not to register the old-style input arrays, HTTP_GET_VARS
; and friends. If you're not using them, it's recommended to turn them off,
; for performance reasons.
register_long_arrays = Off
; This directive tells PHP whether to declare the argv&argc variables (that
; would contain the GET information). If you don't use these variables, you
; should turn it off for increased performance.
register_argc_argv = Off
; When enabled, the SERVER and ENV variables are created when they're first
; used (Just In Time) instead of when the script starts. If these variables
; are not used within a script, having this directive on will result in a
; performance gain. The PHP directives register_globals, register_long_arrays,
; and register_argc_argv must be disabled for this directive to have any affect.
auto_globals_jit = On
; Maximum size of POST data that PHP will accept.
post_max_size = 8M
; Magic quotes
; Magic quotes for incoming GET/POST/Cookie data.
magic_quotes_gpc = Off
; Magic quotes for runtime-generated data, e.g. data from SQL, from exec(), etc.
magic_quotes_runtime = Off
; Use Sybase-style magic quotes (escape ' with '' instead of \').
magic_quotes_sybase = Off
; Automatically add files before or after any PHP document.
auto_prepend_file =
auto_append_file =
; As of 4.0b4, PHP always outputs a character encoding by default in
; the Content-type: header. To disable sending of the charset, simply
; set it to be empty.
; PHP's built-in default is text/html
default_mimetype = "text/html"
default_charset = "iso-8859-1"
; Always populate the $HTTP_RAW_POST_DATA variable.
always_populate_raw_post_data = On
;;;;;;;;;;;;;;;;;;;;;;;;;
; Paths and Directories ;
;;;;;;;;;;;;;;;;;;;;;;;;;
; UNIX: "/path1:/path2"
include_path = "./php/includes"
; Windows: "\path1;\path2"
include_path = ".;c:\php\includes"
; The root of the PHP pages, used only if nonempty.
; if PHP was not compiled with FORCE_REDIRECT, you SHOULD set doc_root
; if you are running php as a CGI under any web server (other than IIS)
; see documentation for security issues. The alternate is to use the
; cgi.force_redirect configuration below
doc_root =
; The directory under which PHP opens the script using /~username used only
; if nonempty.
user_dir =
; Directory in which the loadable extensions (modules) reside.
; extension_dir = "./"
; Whether or not to enable the dl() function. The dl() function does NOT work
; properly in multithreaded servers, such as IIS or Zeus, and is automatically
; disabled on them.
enable_dl = On
; cgi.force_redirect is necessary to provide security running PHP as a CGI under
; most web servers. Left undefined, PHP turns this on by default. You can
; turn it off here AT YOUR OWN RISK
; **You CAN safely turn this off for IIS, in fact, you MUST.**
; cgi.force_redirect = 1
; if cgi.nph is enabled it will force cgi to always sent Status: 200 with

```

```

; every request.
; cgi.nph = 1
; if cgi.force_redirect is turned on, and you are not running under Apache or Netscape
; (iPlanet) web servers, you MAY need to set an environment variable name that PHP
; will look for to know it is OK to continue execution. Setting this variable MAY
; cause security issues, KNOW WHAT YOU ARE DOING FIRST.
; cgi.redirect_status_env = ;
; FastCGI under IIS (on WINNT based OS) supports the ability to impersonate
; security tokens of the calling client. This allows IIS to define the
; security context that the request runs under. mod_fastcgi under Apache
; does not currently support this feature (03/17/2002)
; Set to 1 if running under IIS. Default is zero.
; fastcgi.impersonate = 1;
; Disable logging through FastCGI connection
; fastcgi.log = 0
; cgi.rfc2616_headers configuration option tells PHP what type of headers to
; use when sending HTTP response code. If it's set 0 PHP sends Status: header that
; is supported by Apache. When this option is set to 1 PHP will send
; RFC2616 compliant header.
; Default is zero.
;cgi.rfc2616_headers = 0
;;;;;;;;;;;;;;;;;
; File Uploads ;
;;;;;;;;;;;;;;;;;
; Whether to allow HTTP file uploads.
file_uploads = On
; Temporary directory for HTTP uploaded files (will use system default if not
; specified).
;upload_tmp_dir =
; Maximum allowed size for uploaded files.
upload_max_filesize = 8M
;;;;;;;;;;;;;;;;;
; Fopen wrappers ;
;;;;;;;;;;;;;;;;;
; Whether to allow the treatment of URLs (like http:// or ftp://) as files.
allow_url_fopen = On
; Whether to allow include/require to open URLs (like http:// or ftp://) as files.
allow_url_include = Off
; Define the anonymous ftp password (your email address)
;from="john@doe.com"
; Define the User-Agent string
; user_agent="PHP"
; Default timeout for socket based streams (seconds)
default_socket_timeout = 60
; If your scripts have to deal with files from Macintosh systems,
; or you are running on a Mac and need to deal with files from
; unix or win32 systems, setting this flag will cause PHP to
; automatically detect the EOL character in those files so that
; fgets() and file() will work regardless of the source of the file.
; auto_detect_line_endings = Off
;;;;;;;;;;;;;;;;;
; Dynamic Extensions ;
;;;;;;;;;;;;;;;;;
; If you wish to have an extension loaded automatically, use the following
; syntax:
; extension=modulename.extension
; For example, on Windows:
; extension=mysql.dll
; ... or under UNIX:
; extension=mysql.so
; Note that it should be the name of the module only; no directory information
; needs to go here. Specify the location of the extension with the
; extension_dir directive above.
; Windows Extensions
; Note that ODBC support is built in, so no dll is needed for it.
; Note that many DLL files are located in the extensions/ (PHP 4) ext/ (PHP 5)
; extension folders as well as the separate PECL DLL download (PHP 5).
; Be sure to appropriately set the extension_dir directive.
;extension=php_mbstring.dll

```

```

;extension=php_bz2.dll
;extension=php_curl.dll
;extension=php_dba.dll
;extension=php_dbase.dll
;extension=php_exif.dll
;extension=php_fdf.dll
;extension=php_filepro.dll
;extension=php_gd2.dll
;extension=php_gettext.dll
;extension=php_ifx.dll
;extension=php_imap.dll
;extension=php_interbase.dll
;extension=php_ldap.dll
;extension=php_mcrypt.dll
;extension=php_mhash.dll
;extension=php_mime_magic.dll
;extension=php_ming.dll
;extension=php_mssql.dll
;extension=php_mysql.dll
;extension=php_mysqli.dll
;extension=php_oci8.dll
;extension=php_openssl.dll
;extension=php_oracle.dll
;extension=php_pgsql.dll
;extension=php_shmop.dll
;extension=php_snmp.dll
;extension=php_sockets.dll
;extension=php_sqlite.dll
;extension=php_sybase_ct.dll
;extension=php_tidy.dll
;extension=php_xmldrpc.dll
;extension=php_xsl.dll
;
; Module Settings ;
;
[Date]
; Defines the default timezone used by the date functions
;date.timezone =
;date.default_latitude = 31.7667
;date.default_longitude = 35.2333
;date.sunrise_zenith = 90.583333
;date.sunset_zenith = 90.583333
[filter]
;filter.default = unsafe_raw
;filter.default_flags =

[iconv]
;iconv.input_encoding = ISO-8859-1
;iconv.internal_encoding = ISO-8859-1
;iconv.output_encoding = ISO-8859-1
[sqlite]
;sqlite.assoc_case = 0
[xmldrpc]
;xmldrpc_error_number = 0
;xmldrpc_errors = 0
[Pcre]
;pcre.recursion_limit=100000
;pcre.backtrack_limit=100000
[Syslog]
; Whether or not to define the various syslog variables (e.g. $LOG_PID,
; $LOG_CRON, etc.). Turning it off is a good idea performance-wise. In
; runtime, you can define these variables by calling define_syslog_variables().
define_syslog_variables = Off
[mail function]
; For Win32 only.
SMTP = localhost
smtp_port = 25
; For Win32 only.
;sendmail_from = me@example.com
; For Unix only. You may supply arguments as well (default: "sendmail -t -i").

```

```

;sendmail_path =
; Force the addition of the specified parameters to be passed as extra parameters
; to the sendmail binary. These parameters will always replace the value of
; the 5th parameter to mail(), even in safe mode.
;mail.force_extra_parameters =
[SQL]
sql.safe_mode = Off
[ODBC]
;odbc.default_db = Not yet implemented
;odbc.default_user = Not yet implemented
;odbc.default_pw = Not yet implemented
; Allow or prevent persistent links.
odbc.allow_persistent = On
; Check that a connection is still valid before reuse.
odbc.check_persistent = On
; Maximum number of persistent links. -1 means no limit.
odbc.max_persistent = -1
; Maximum number of links (persistent + non-persistent). -1 means no limit.
odbc.max_links = -1
; Handling of LONG fields. Returns number of bytes to variables. 0 means
; passthru.
odbc.defaultlrl = 4096
; Handling of binary data. 0 means passthru, 1 return as is, 2 convert to char.
; See the documentation on odbc_binmode and odbc_longreadlen for an explanation
; of uodbc.defaultlrl and uodbc.defaultbinmode
odbc.defaultbinmode = 1
[MySQL]
; Allow or prevent persistent links.
mysql.allow_persistent = On
; Maximum number of persistent links. -1 means no limit.
mysql.max_persistent = -1
; Maximum number of links (persistent + non-persistent). -1 means no limit.
mysql.max_links = -1
; Default port number for mysql_connect(). If unset, mysql_connect() will use
; the $MYSQL_TCP_PORT or the mysql-tcp entry in /etc/services or the
; compile-time value defined MYSQL_PORT (in that order). Win32 will only look
; at MYSQL_PORT.
mysql.default_port =
; Default socket name for local MySQL connects. If empty, uses the built-in
; MySQL defaults.
mysql.default_socket =
; Default host for mysql_connect() (doesn't apply in safe mode).
mysql.default_host =
; Default user for mysql_connect() (doesn't apply in safe mode).
mysql.default_user =
; Default password for mysql_connect() (doesn't apply in safe mode).
; Note that this is generally a *bad* idea to store passwords in this file.
; *Any* user with PHP access can run 'echo get_cfg_var("mysql.default_password")
; and reveal this password! And of course, any users with read access to this
; file will be able to reveal the password as well.
mysql.default_password =
; Maximum time (in seconds) for connect timeout. -1 means no limit
mysql.connect_timeout = 60
; Trace mode. When trace_mode is active (=On), warnings for table/index scans and
; SQL-Errors will be displayed.
mysql.trace_mode = Off
[MySQLi]
; Maximum number of links. -1 means no limit.
mysqli.max_links = -1
; Default port number for mysqli_connect(). If unset, mysqli_connect() will use
; the $MYSQL_TCP_PORT or the mysql-tcp entry in /etc/services or the
; compile-time value defined MYSQL_PORT (in that order). Win32 will only look
; at MYSQL_PORT.
mysqli.default_port = 3306
; Default socket name for local MySQL connects. If empty, uses the built-in
; MySQL defaults.
mysqli.default_socket =
; Default host for mysqli_connect() (doesn't apply in safe mode).
mysqli.default_host =
; Default user for mysqli_connect() (doesn't apply in safe mode).

```

```

mysql.default_user =
; Default password for mysqli_connect() (doesn't apply in safe mode).
; Note that this is generally a *bad* idea to store passwords in this file.
; *Any* user with PHP access can run 'echo get_cfg_var("mysqli.default_pw")
; and reveal this password! And of course, any users with read access to this
; file will be able to reveal the password as well.
mysqli.default_pw =
; Allow or prevent reconnect
mysqli.reconnect = Off
[MySQL]
; Allow or prevent persistent links.
mysql.allow_persistent = On
; Maximum number of persistent links. -1 means no limit.
mysql.max_persistent = -1
; Maximum number of links (persistent+non persistent). -1 means no limit.
mysql.max_links = -1
[PostgreSQL]
; Allow or prevent persistent links.
pgsql.allow_persistent = On
; Detect broken persistent links always with pg_pconnect().
; Auto reset feature requires a little overheads.
pgsql.auto_reset_persistent = Off
; Maximum number of persistent links. -1 means no limit.
pgsql.max_persistent = -1
; Maximum number of links (persistent+non persistent). -1 means no limit.
pgsql.max_links = -1
; Ignore PostgreSQL backends Notice message or not.
; Notice message logging require a little overheads.
pgsql.ignore_notice = 0
; Log PostgreSQL backends Noitce message or not.
; Unless pgsql.ignore_notice=0, module cannot log notice message.
pgsql.log_notice = 0
[Sybase]
; Allow or prevent persistent links.
sybase.allow_persistent = On
; Maximum number of persistent links. -1 means no limit.
sybase.max_persistent = -1
; Maximum number of links (persistent + non-persistent). -1 means no limit.
sybase.max_links = -1
;sybase.interface_file = "/usr/sybase/interfaces"
; Minimum error severity to display.
sybase.min_error_severity = 10
; Minimum message severity to display.
sybase.min_message_severity = 10
; Compatibility mode with old versions of PHP 3.0.
; If on, this will cause PHP to automatically assign types to results according
; to their Sybase type, instead of treating them all as strings. This
; compatibility mode will probably not stay around forever, so try applying
; whatever necessary changes to your code, and turn it off.
sybase.compatibility_mode = Off
[Sybase-CT]
; Allow or prevent persistent links.
sybct.allow_persistent = On
; Maximum number of persistent links. -1 means no limit.
sybct.max_persistent = -1
; Maximum number of links (persistent + non-persistent). -1 means no limit.
sybct.max_links = -1
; Minimum server message severity to display.
sybct.min_server_severity = 10
; Minimum client message severity to display.
sybct.min_client_severity = 10
[bcmath]
; Number of decimal digits for all bcmath functions.
bcmath.scale = 0
[browscap]
;browscap = extra/browscap.ini
[Informix]
; Default host for ifx_connect() (doesn't apply in safe mode).
ifx.default_host =
; Default user for ifx_connect() (doesn't apply in safe mode).

```

```

ifx.default_user =
; Default password for ifx_connect() (doesn't apply in safe mode).
ifx.default_password =
; Allow or prevent persistent links.
ifx.allow_persistent = On
; Maximum number of persistent links. -1 means no limit.
ifx.max_persistent = -1
; Maximum number of links (persistent + non-persistent). -1 means no limit.
ifx.max_links = -1
; If on, select statements return the contents of a text blob instead of its id.
ifx.textasvarchar = 0
; If on, select statements return the contents of a byte blob instead of its id.
ifx.byteasvarchar = 0
; Trailing blanks are stripped from fixed-length char columns. May help the
; life of Informix SE users.
ifx.charasvarchar = 0
; If on, the contents of text and byte blobs are dumped to a file instead of
; keeping them in memory.
ifx.blobinfile = 0
; NULL's are returned as empty strings, unless this is set to 1. In that case,
; NULL's are returned as string 'NULL'.
ifx.nullformat = 0
[Session]
; Handler used to store/retrieve data.
session.save_handler = files
; Argument passed to save_handler. In the case of files, this is the path
; where data files are stored. Note: Windows users have to change this
; variable in order to use PHP's session functions.
; As of PHP 4.0.1, you can define the path as:
; session.save_path = "N;/path"
; where N is an integer. Instead of storing all the session files in
; /path, what this will do is use subdirectories N-levels deep, and
; store the session data in those directories. This is useful if you
; or your OS have problems with lots of files in one directory, and is
; a more efficient layout for servers that handle lots of sessions.
; NOTE 1: PHP will not create this directory structure automatically.
; You can use the script in the ext/session dir for that purpose.
; NOTE 2: See the section on garbage collection below if you choose to
; use subdirectories for session storage
; The file storage module creates files using mode 600 by default.
; You can change that by using
; session.save_path = "N;MODE;/path"
; where MODE is the octal representation of the mode. Note that this
; does not overwrite the process's umask.
;session.save_path = "/tmp"
; Whether to use cookies.
session.use_cookies = 1
;session.cookie_secure =

; This option enables administrators to make their users invulnerable to
; attacks which involve passing session ids in URLs; defaults to 0.
;session.use_only_cookies = 1
; Name of the session (used as cookie name).
session.name = PHPSESSID
; Initialize session on request startup.
session.auto_start = 0
; Lifetime in seconds of cookie or, if 0, until browser is restarted.
session.cookie_lifetime = 0
; The path for which the cookie is valid.
session.cookie_path = /
; The domain for which the cookie is valid.
session.cookie_domain =
; Whether or not to add the httpOnly flag to the cookie, which makes it inaccessible to browser scripting languages such as
JavaScript.
session.cookie_httponly =
; Handler used to serialize data. php is the standard serializer of PHP.
session.serialize_handler = php
; Define the probability that the 'garbage collection' process is started
; on every session initialization.
; The probability is calculated by using gc_probability/gc_divisor,

```

```

; e.g. 1/100 means there is a 1% chance that the GC process starts
; on each request.
session.gc_probability = 1
session.gc_divisor = 1000
; After this number of seconds, stored data will be seen as 'garbage' and
; cleaned up by the garbage collection process.
session.gc_maxlifetime = 1440
; NOTE: If you are using the subdirectory option for storing session files
; (see session.save_path above), then garbage collection does *not*
; happen automatically. You will need to do your own garbage
; collection through a shell script, cron entry, or some other method.
; For example, the following script would be the equivalent of
; setting session.gc_maxlifetime to 1440 (1440 seconds = 24 minutes):
; cd /path/to/sessions; find -cmin +24 | xargs rm
; PHP 4.2 and less have an undocumented feature/bug that allows you to
; to initialize a session variable in the global scope, albeit register_globals
; is disabled. PHP 4.3 and later will warn you, if this feature is used.
; You can disable the feature and the warning separately. At this time,
; the warning is only displayed, if bug_compat_42 is enabled.
session.bug_compat_42 = 0
session.bug_compat_warn = 1
; Check HTTP Referer to invalidate externally stored URLs containing ids.
; HTTP_REFERER has to contain this substring for the session to be
; considered as valid.
session.referer_check =
; How many bytes to read from the file.
session.entropy_length = 0
; Specified here to create the session id.
session.entropy_file =
;session.entropy_length = 16
;session.entropy_file = /dev/urandom
; Set to {nocache,private,public,} to determine HTTP caching aspects
; or leave this empty to avoid sending anti-caching headers.
session.cache_limiter = nocache
; Document expires after n minutes.
session.cache_expire = 180
; trans sid support is disabled by default.
; Use of trans sid may risk your users security.
; Use this option with caution.
; - User may send URL contains active session ID
; to other person via. email/irc/etc.
; - URL that contains active session ID may be stored
; in publically accessible computer.
; - User may access your site with the same session ID
; always using URL stored in browser's history or bookmarks.
session.use_trans_sid = 0
; Select a hash function
; 0: MD5 (128 bits)
; 1: SHA-1 (160 bits)
session.hash_function = 0
; Define how many bits are stored in each character when converting
; the binary hash data to something readable.
; 4 bits: 0-9, a-f
; 5 bits: 0-9, a-v
; 6 bits: 0-9, a-z, A-Z, "-", ",",
session.hash_bits_per_character = 5
; The URL rewriter will look for URLs in a defined set of HTML tags.
; form/fieldset are special; if you include them here, the rewriter will
; add a hidden <input> field with the info which is otherwise appended
; to URLs. If you want XHTML conformity, remove the form entry.
; Note that all valid entries require a "=", even if no value follows.
url_rewriter.tags = "a:href,area:href,frame=src,input=src,form=fakeentry"
[MSSQL]
; Allow or prevent persistent links.
mssql.allow_persistent = On
; Maximum number of persistent links. -1 means no limit.
mssql.max_persistent = -1
; Maximum number of links (persistent+non persistent). -1 means no limit.
mssql.max_links = -1
; Minimum error severity to display.

```

```

mssql.min_error_severity = 10
; Minimum message severity to display.
mssql.min_message_severity = 10
; Compatibility mode with old versions of PHP 3.0.
mssql.compatibility_mode = Off
; Connect timeout
;mssql.connect_timeout = 5
; Query timeout
;mssql.timeout = 60
; Valid range 0 - 2147483647. Default = 4096.
;mssql.textlimit = 4096
; Valid range 0 - 2147483647. Default = 4096.
;mssql.textsize = 4096
; Limits the number of records in each batch. 0 = all records in one batch.
;mssql.batchsize = 0
; Specify how datetime and datetime4 columns are returned
; On => Returns data converted to SQL server settings
; Off => Returns values as YYYY-MM-DD hh:mm:ss
;mssql.datetimeconvert = On
; Use NT authentication when connecting to the server
mssql.secure_connection = Off
; Specify max number of processes. -1 = library default
; msdlib defaults to 25
; FreeTDS defaults to 4096
;mssql.max_procs = -1
; Specify client character set.
; If empty or not set the client charset from freetds.comf is used
; This is only used when compiled with FreeTDS
;mssql.charset = "ISO-8859-1"
[Assertion]
; Assert(expr); active by default.
;assert.active = On
; Issue a PHP warning for each failed assertion.
;assert.warning = On
; Don't bail out by default.
;assert.bail = Off
; User-function to be called if an assertion fails.
;assert.callback = 0
; Eval the expression with current error_reporting(). Set to true if you want
; error_reporting(0) around the eval().
;assert.quiet_eval = 0
[COM]
; path to a file containing GUIDs, IIDs or filenames of files with TypeLibs
;com.typelib_file =
; allow Distributed-COM calls
;com.allow_dcom = true
; autoregister constants of a components typlib on com_load()
;com.autoregister_typelib = true
; register constants casesensitive
;com.autoregister_casesensitive = false
; show warnings on duplicate constant registrations
;com.autoregister_verbose = true
[mbstring]
; language for internal character representation.
;mbstring.language = Japanese
; internal/script encoding.
; Some encoding cannot work as internal encoding.
; (e.g. SJIS, BIG5, ISO-2022-*)
;mbstring.internal_encoding = EUC-JP
; http input encoding.
;mbstring.http_input = auto
; http output encoding. mb_output_handler must be
; registered as output buffer to function
;mbstring.http_output = SJIS
; enable automatic encoding translation according to
; mbstring.internal_encoding setting. Input chars are
; converted to internal encoding by setting this to On.
; Note: Do_not_use automatic encoding translation for
; portable libs/applications.
;mbstring.encoding_translation = Off

```

```

; automatic encoding detection order.
; auto means
;mbstring.detect_order = auto
; substitute_character used when character cannot be converted
; one from another
;mbstring.substitute_character = none;
; overload(replace) single byte functions by mbstring functions.
; mail(), ereg(), etc are overloaded by mb_send_mail(), mb_ereg(),
; etc. Possible values are 0,1,2,4 or combination of them.
; For example, 7 for overload everything.
; 0: No overload
; 1: Overload mail() function
; 2: Overload str*() functions
; 4: Overload ereg*() functions
;mbstring.func_overload = 0
; enable strict encoding detection.
;mbstring.strict_encoding = Off
[FrontBase]
;fbsql.allow_persistent = On
;fbsql.autocommit = On
;fbsql.show_timestamp_decimals = Off
;fbsql.default_database =
;fbsql.default_database_password =
;fbsql.default_host =
;fbsql.default_password =
;fbsql.default_user = "_SYSTEM"
;fbsql.generate_warnings = Off
;fbsql.max_connections = 128
;fbsql.max_links = 128
;fbsql.max_persistent = -1
;fbsql.max_results = 128
[gd]
; Tell the jpeg decode to libjpeg warnings and try to create
; a gd image. The warning will then be displayed as notices
; disabled by default
;gd.jpeg_ignore_warning = 0
[exif]
; Exif UNICODE user comments are handled as UCS-2BE/UCS-2LE and JIS as JIS.
; With mbstring support this will automatically be converted into the encoding
; given by corresponding encode setting. When empty mbstring.internal_encoding
; is used. For the decode settings you can distinguish between motorola and
; intel byte order. A decode setting cannot be empty.
;exif.encode_unicode = ISO-8859-15
;exif.decode_unicode_motorola = UCS-2BE
;exif.decode_unicode_intel = UCS-2LE
;exif.encode_jis =
;exif.decode_jis_motorola = JIS
;exif.decode_jis_intel = JIS
[Tidy]
; The path to a default tidy configuration file to use when using tidy
;tidy.default_config = /usr/local/lib/php/default.tcfg
; Should tidy clean and repair output automatically?
; WARNING: Do not use this option if you are generating non-html content
; such as dynamic images
tidy.clean_output = Off
[soap]
; Enables or disables WSDL caching feature.
soap.wsdl_cache_enabled=1
; Sets the directory name where SOAP extension will put cache files.
soap.wsdl_cache_dir="/tmp"
; (time to live) Sets the number of second while cached file will be used
; instead of original one.
soap.wsdl_cache_ttl=86400
; Local Variables:
; tab-width: 4
; End:

```

Activitatea 5. Dobândire competențe complementare prin participări la manifestări științifice/stagii de documentare-cercetare

În a doua fază a dobândirii de competențe complementare prin participări la manifestări științifice / stagii de documentare-cercetare s-a planificat pentru perioada 6-14 decembrie participarea la următoarele:

- ÷ University of Oxford, Computational Biology Reseach Group, cursul de instruire intitulat "Introduction to Bioinformatics at CBRG", 7 decembrie 2007
- ÷ Dublin Molecular Medicine Centre, cursul de instruire intitulat "DMMC Course: Techniques and Strategies in Molecular Medicine", 10-13 decembrie 2007
- ÷ Trinity College Dublin, Centre for Synthesis & Chemical Biology, simpozionul intitulat "Recent Advances in Synthesis and Chemical Biology VI", 14 decembrie 2007

Programul evenimentelor planificate este redat în tabelul de mai jos:

"Introduction to Bioinformatics at CBRG", 7 decembrie 2007

This course is intended for new users and/or anyone not familiar with using their molbiol account for bioinformatics analysis. It will be held in the Medical Sciences Teaching Centre (behind the Dunn School of Pathology on South Parks Road).

The day-long course is designed to be run in two sessions:

The first session will introduce you to the many bioinformatics analysis tools that are now available on this web site. The course will also introduce you to EMBOSS Explorer - a suite of bioinformatics software intended largely as a replacement for GCG. You will be shown how to carry out some basic bioinformatics analyses - for example:

access the databases on our servers to retrieve sequence files

examine sequence file formats

run restriction analysis software

carry out sequence alignments and produce publishable images of the aligned sequences

search databases using BLAST

It will also introduce you to other more specialised tools - BASE and MASCOT - that will allow you to analyse your microarray and proteomics data respectively.

Second session: Some of the bioinformatics tools are only available on our Unix servers and the second session will introduce the Unix computing facilities available via your molbiol account. We will show you how to connect to the CBRG Unix machines and introduce you to the Unix environment in general.

You will be introduced to some of the features of the programs from the morning session that are

not available via the web. There will be a brief introduction to the Staden package of software used for the analysis of dna sequencing chromatograms.

The course is not intended to be a comprehensive guide to all the bioinformatics packages available at the CBRG. Instead it is designed to show you the kind of analysis software that is available to you via your account. It should ensure that you also know where to look to find similar software on the CBRG web site and on orac and to find help with any of the software on the system.

"DMMC Course: Techniques and Strategies in Molecular Medicine", 10-13 decembrie 2007

This course, running over four mornings, is designed to give bioscientists and clinicians a broad overview of research techniques and their application. Basic molecular biology laboratory experience is assumed, but you should not need prior knowledge of the techniques covered in the course.

This course, running over four mornings, is designed to give bioscientists and clinicians a broad overview of research techniques and their application. Basic molecular biology laboratory experience is assumed, but you should not need prior knowledge of the techniques covered in the course.

The objective for postgraduate teaching is to give students a broad basic knowledge of bioscience research techniques and technologies, including those not currently used in their own project/laboratory that may be of future use. We recommend that postgraduate students attend the whole course (DMMC Course Attendance Certificates are only provided for complete attendance).

Follow the links below to read an abstract of each lecture, together with supplementary reading in some cases.

ANALYSING GENES (Mon 10 Dec; 0930-1300)

Session Chair: Dr Ross McManus, TCD

0930 RNA Detection and quantitation Dr Shane Duggan, TCD

The protein components of the cell are derived by numerous processes indirectly interpreted from a genetic element known as the "gene" which is coded in the cellular DNA. This element is interpreted by the cell in a process called "transcription" where the genetic code for a particular gene is converted into a molecular code known as messenger RNA (mRNA). This mRNA molecule can now be utilised in the creation of a new protein via the translation process. In this lecture the nature and analysis of Ribonucleic acid (RNA) in biological systems will be explored. The understanding of this has allowed the laboratory scientist to interrogate and explore gene expression as it may relate to diseases or cell signalling. Extraction and quantitation of good quality RNA will be discussed as they are the first step in any investigation of gene expression. Standard techniques in common use such as Northern blotting and cycle limited RT-PCR shall also be described as well as more modern techniques such as real time RT-PCR analysis. This lecture will allow the interpretation of published literature utilising these techniques and introduce the steps involved in performing RNA related techniques in your laboratory.

1015 Differential gene expression: overview of relevant methods Prof William Gallagher, UCD

This lecture will summarise the main approaches used to determine alterations in gene expression at the RNA level.

Emphasis will be placed in this context on global approaches that attempt to map differences in the transcriptome, i.e. entire complement of transcripts in a cell. Methodologies that will be addressed include differential display, subtractive hybridization, high-throughput sequencing (ESTs and SAGE), and DNA microarray technologies. Key examples from the literature will be utilised to illustrate examples of investigators applying these technologies to understand biological phenomena, with a focus on disease-related processes. An indication of the relevant infrastructure and expertise to carry out this work within the DMMC will be presented.

Review articles

Lennon, G. G. (2000). High-throughput gene expression analysis for drug discovery. *Drug Discovery Today*, 5, 59-66.

Schulze, A. and Downward, J. (2001). Navigating gene expression using microarrays – a technology review. *Nature Cell Biology*, 3, E190-E195.

1100 Coffee/Tea

1130 Mutation detection, SNP analysis and genetic linkage Prof Denis Shields, UCD

Different strategies are required to identify rare and common genetic variants underlying both rare and common diseases. For common genetic variants, there is now a very rich dataset of identified common single nucleotide polymorphisms (SNPs). These can be investigated in disease groups (compared to controls) in candidate genes, or by whole genome association analysis. Analysis of these genes requires careful attention to the patterns of association of SNPs that are chromosomally adjacent (in linkage disequilibrium). Linkage analysis (tracking in families the disease co-inheritance with widely spaced gene markers) is the traditional approach of choice for rare mutations that have strong phenotypic effects. High throughput sequencing of candidate regions (and in future whole genomes) are accelerating the rate of data accumulation.

1215 Model organisms Dr Breandán Kennedy, UCD

The goal of this lecture is to discuss animal models that are routinely applied to biomedical research. The advantages of using *Drosophila* (fly), *Xenopus* (frog), *Danio* (zebrafish), *Gallus* (chicken) and *Mus Musculus* (mouse) as model organisms will be described. The life-cycle, generation time, embryo development and amenability of these organisms to genetic manipulation will be discussed. An emphasis will be placed on describing mutagenesis screens. This technique, in which the genes in the genome are randomly inactivated, has been extensively applied to the fly/fish models and has accelerated our understanding of gene function (functional genomics).

MANIPULATING GENES (Tue 11 Dec; 0930-1300)

Session Chair: Dr Ross McManus, TCD

0930 DNA cloning strategies Dr Ross McManus, TCD

Even in the post genome era, DNA cloning is essential to the manipulation and stable propagation of genetic material. This talk will cover the basic aspects of DNA cloning, ranging from the anatomy of cloning vectors to the choice of vectors based on the cloning strategy employed. The strategy employed will depend on the overall objectives of the project and the nature of the starting information or material available. Thus different choices and approaches would be employed for a sequencing project compared with a genome mapping project or production of RNA or protein. I will discuss a number of basic and specialised cloning strategies to illustrate some of the options and possibilities available.

1015 RNA interference Dr Jane Farrar, TCD

1100 Coffee/Tea

1130 Transgenics and knockouts Dr Derek Brazil, UCD

This lecture will provide a broad overview of the strategies used to generate both transgenic and knockout mice, starting

from the generation of the DNA constructs using cDNAs or genomic DNA, and proceeding through embryonic stem cell biology, to aggregation and chimeric mouse generation. Details on genotyping of transgenic animals, as well as phenotype characterization will be discussed. Specific examples such as the IRS-2 knockout will be cited.

1215 Molecular therapies - false hope or the future of medicine? Dr Ruth Foley, TCD

Molecular medicine has allowed the identification of new targets and new approaches to treat human disease. The lecture will focus on some of these strategies, including targeting signal transduction pathways, use of monoclonal antibody based approaches and the numerous gene therapy strategies that are currently being evaluated. An overview of the different methodologies will be presented and the current situation on the clinical application of these approaches will be considered. Focusing particularly on cancer therapies as a model system, the advances and challenges of the different approaches will be presented and discussed.

PROTEINS (Wed 12 Dec; 0930-1300)

Session Chair: Dr Niamh Moran, RCSI

0930 Protein expression and purification

Dr Henry Windle, TCD

This lecture will cover the basics of protein expression and purification. Emphasis will be placed on alternative strategies and issues that should be considered prior to selection of specific expression systems and purification strategies. As protein purification methodologies are generally well described and accessible, only a brief overview of these will be given but with emphasis on common problems that can arise, particularly for those about to attempt purification for the first time. The following books from The Practical Approach series by IRL Press are an invaluable aid with detailed and reliable protocols: Protein Purification Applications; Protein Purification Methods (2001, Editor Simon Roe).

Gallus (chicken) and Mus Musculus (mouse) as model organisms will be described. The life-cycle, generation time, embryo development and amenability of these organisms to genetic manipulation will be discussed. An emphasis will be placed on describing mutagenesis screens. This technique, in which the genes in the genome are randomly inactivated, has been extensively applied to the fly/fish models and has accelerated our understanding of gene function (functional genomics).

1015 Determining protein: protein interactions in biology Dr Niamh Moran, RCSI

During the past two decades, mass spectrometry has become a major technique for the identification, characterisation and quantification of biological molecules and bioactive drugs. In particular, the impact of mass spectrometry on proteomics and metabolomics has been phenomenal.

This lecture will outline some of the applications of mass spectrometry in modern life sciences and introduce course participants to basic vocabulary and concepts in biological mass spectrometry.

1100 Coffee/Tea

1130 Introduction to Mass Spectrometry

Dr Achim Treumann, RCSI

During the past two decades, mass spectrometry has become a major technique for the identification, characterisation and quantification of biological molecules and bioactive drugs. In particular, the impact of mass spectrometry on proteomics and metabolomics has been phenomenal.

This lecture will outline some of the applications of mass spectrometry in modern life sciences and introduce course participants to basic vocabulary and concepts in biological mass spectrometry.

1215 Proteomic technologies Prof Steve Pennington, UCD

CELLS & TISSUES (Thu 13 Dec; 0930-1300)

Session Chair: Dr William Watson, UCD

0930 Immunodetection methods on cell and tissue extracts Dr Leonie Young, RCSI

The use of antibodies to detect and characterise proteins has been well established. With the development of high through-put techniques such as tissue microarrays (TMA), a real challenge now exists to determine the cellular location, level of expression and the function of these identified proteins. In this lecture, principles fundamental to immunodetection will be outlined. Common pitfalls and measures to avoid these will be discussed. Applications of immunodetection in a modern molecular context will be illustrated, including: western blotting, ELISA, immunohistochemistry/ immunofluorescence, tissue microarrays, co-immunoprecipitation, Electromobility shift assays, chromatin immunoprecipitation (ChIP) and antibody arrays.

1015 Cell imaging and sorting - flow cytometry

Dr William Watson, UCD

Flow cytometry is a method for quantitating components or structural features of cells, primarily by optical means. Although it makes measurements on one cell at a time, it can process thousands of cells in a few seconds. Since cell types can be distinguished by quantitating structural features, flow cytometry can be used to count prokaryotic or eukaryotic cells of different types in complex mixtures.

1100 Coffee/Tea

1130 High Content Analysis of nanoparticle/cell interactions Dr Yuri Volkov, TCD

Fluorescent organic tags have represented one of the major tools in the arsenal of researchers working in the biomedical sciences for more than two decades. The progress in development of new fast and efficient research and diagnostic methods is largely dependent on the availability of fluorescent probes with desired cell receptor- and organelle specificity and optimised experimental protocols for their utilization.

A unique opportunity to generate a wide spectrum of such probes suitable for applications in living cells is offered by semiconductor quantum dots (QDs). As fluorescent probes QDs have several advantages over organic dyes, including wide absorption profiles, tunable emission spectra, and superior photostability. QDs have been shown to readily distribute across animal cells, tissues and organs. Today, QDs with different physico-chemical properties and functionalities are readily available worldwide. However, further exploitation of QDs in biomedical studies has been hindered by the absence of adequate technological platforms capable of performing multi-parametric quantitative analysis of individual responses in specific cell types.

Recent years have witnessed a rapid progress in the development of novel methods permitting high-resolution visualisation of cell receptor dynamics and intracellular biochemical processes utilizing fluorescent probes. Among these, high content screening (HCS) technology allows to perform analysis of molecular interactions in individual cells and their populations at the sub-cellular level under physiological conditions. This technology not only facilitates development of a better understanding of the true functionality of target molecules in the living cells, but it can also promote designing of highly informative screens for novel therapeutic drugs, including inhibitors at small molecule and gene therapy level applicable in inflammation and cancer.

Here we will discuss the data demonstrating the influence of such factors as QDs size, charge and selective functionalisation on their membrane and subcellular localisation specificity and present an overview of advantages and hurdles on the way of merging nanotechnology and high content cell analysis.

1215 Laser Capture Microdissection and in situ hybridisation Dr Orla Sheils, TCD

Laser Capture Microdissection is a method for procuring pure cells from specific microscopic regions of tissue sections. Under the microscope, tissues are heterogeneous complicated structures with hundreds of different cell types locked in morphologic units. In disease pathologies, the diseased cells of interest are surrounded by these heterogeneous tissue elements. Laser Capture Microdissection constitutes an essential upstream technology to molecular analysis methods studying evolving disease lesions in actual tissue.

In Situ Hybridization techniques allow the demonstration of specific nucleic acid sequences within their cellular environment. A logical extension of early in situ hybridization (ISH) techniques, which exploited the ability to label DNA with high-energy fluorophores, is FISH. This technique is now applied in an increasing number of molecular diagnostic areas, including karyotype analysis, gene mapping, disease diagnosis, and therapeutic targeting.

Course Instructors

Dr Derek Brazil (UCD Conway Institute of Biomolecular & Biomedical Research)

Dr Shane Duggan (Institute of Molecular Medicine & TCD)

Dr Jane Farrar (TCIN & TCD)

Dr Ruth Foley (Institute of Molecular Medicine, TCD)

Prof William Gallagher (UCD Conway Institute of Biomolecular & Biomedical Research)

Prof Dr Breandán Kennedy (UCD Conway Institute of Biomolecular & Biomedical Research)

Prof Mark Lawler (Institute of Molecular Medicine, St James's Hospital & TCD)

Dr Ross McManus (Institute of Molecular Medicine, St James's Hospital & TCD)

Dr Niamh Moran (Institute of Biopharmaceutical Sciences, RCSI)

Prof Steve Pennington (UCD Conway Institute of Biomolecular & Biomedical Research)

Dr Orla Sheils (Institute of Molecular Medicine, TCD)

Denis Shields (UCD Conway Institute of Biomolecular & Biomedical Research)

Dr Achim Treumann (Institute of Biopharmaceutical Sciences, RCSI)

Dr Yuri Volkov (Institute of Molecular Medicine, St James's Hospital & TCD)

Dr William Watson (UCD Conway Institute of Biomolecular & Biomedical Research)

Dr Henry Windle (Institute of Molecular Medicine, St James's Hospital & TCD)

Dr Leonie Young (Royal College of Surgeons in Ireland & Beaumont Hospital)

"Recent Advances in Synthesis and Chemical Biology VI", 14 decembrie 2007

8.45am Opening session

Introductory remarks: Professor Thorri Gunnlaugsson

Speaker: Professor John Hegarty, Provost, TCD

9.00am-10.00am Chairperson: Professor Paul Murphy

Professor Peter Seeberger

'From Microreactors to a Malaria Vaccine'

10.00am-11.00am Chairperson: Professor Kevin Nolan

Professor Thomas Carell

‘Synthetic and Crystallographic studies of Transcription and Replication through DNA lesions’

11.00am-11.30am Coffee/Tea Break

11.30am-12.30pm Chairperson: Dr Donal O’Shea

GLAXOSMITHKLINE LECTURE

Professor Michael J. Krische

‘Formation of C-C Bonds via Catalytic Hydrogenation and Transfer Hydrogenation’

12.30pm-1.15pm Lunch Break

1.15pm-2.15pm Chairperson: Professor Mathias Senge

ELI LILLY LECTURE

Professor Chris Schofield

‘The Chemistry of Oxygen Sensing’

2.15pm-3.15pm Poster Session. Coffee/Tea Break

3.15pm-4.15pm Chairperson: Professor John M. Kelly

INSTITUT DE RECHERCHES SERVIER LECTURE

Professor André Kirsch - De Mesmaeker

‘Ru(II) complexes under illumination. Can they be used in Chemical Biology?’

4.15pm-5.15pm Chairperson: Dr Marc Devocelle

Professor Mark Bradley

‘Interfacing Combinatorial Chemistry with Microarrays’

5.15pm Closing Remarks followed by half-hour Wine Reception

Professor Pat Guiry

Activitatea 6. Elaborare specificații și metodologie model matematic

O serie de lucrări au constituit referința pentru elaborarea specificațiilor și metodologiei modelului matematic:

| Nr | Lucrare | Metoda | Rezultate |
|----|---|---|---|
| 1 | QSPR modeling of stability constants of diverse 15-crown-5 ethers complexes using best multiple linear regression, J Incl Phenom Macrocycl Chem, (2007+), DOI 10.1007/s10847-007-9383-3, Jahanbakhsh Ghasemi, Saadi Saaidpour JInclPhenomMacrocyclChem1.pdf | <p>÷ Molecule desenate în HyperChem 7.5</p> <p>÷ Geometrie optimizată în AMPAC 8.16 cu AM1 (Austin Model 1) - metodă cuantică semiempirică</p> <p>÷ Fișierele HyperChem *.mol și fișierele output AMPAC au fost folosite în programul CODESSA pentru a calcula mai mult de 575 descriptori de structură pentru cei 58 eteri coroană</p> | <p>Număr de compuși: 58</p> <p>Număr de descriptori: 5</p> <p>Coefficient de determinare: 0.9452</p> <p>TvT Experiment: Da (46 Training, 12 Test)</p> <p>Cei 5 descriptori sunt:</p> <ol style="list-style-type: none"> 1. $Q_{\max} = \max(Q_H^+)$ 2. ${}^kIC = -\sum_{i=1}^k n_i/n \cdot \log_2(n_i/n)$, k număr de straturi atomice în sfera de coordinare a unui atom dat 3. $E_{nn}(C-H) = Z_C Z_H / R_{CH}$ 4. $E_{nn}(C-C) = Z_C Z_H / R_{CC}$ 5. $WNSA1 =$ $PNSA1 \cdot TMSA / 1000$ |
| 2 | Representation of molecular structure using quantum topology with inductive logic programming in structure-activity relationships, J Comput Aided Mol Des, (2006) 20:361–373, DOI 10.1007/s10822-006-9058-y, Bard Buttingsrud, Einar Ryeng, Ross D. King, Bjørn K. Alsberg JComputAidedMolDes1.pdf | <p>÷ inductive logic programming (ILP) ref [Nienhuys-Cheng SH, de Wolf R (1997) Foundations of inductiv logic programming, volume 1228 of Lecture notes in artificial intelligence. Springer-Verlag, Berlin]</p> <p>÷ SAR method based on</p> | <p>÷ Mutagenesis</p> <p>Număr de compuși: 230</p> <p>Număr de descriptori: N/A</p> <p>Coefficient de determinare: 0.84, 0.86; 0.91, 0.91</p> <p>TvT Experiment: Da (188 Training, 42 Test)</p> <p>÷ Factor Xa</p> <p>Număr de compuși: 230</p> <p>Număr de descriptori: N/A</p> <p>Coefficient de determinare: 0.88, 0.84, 0.88</p> |

| | | | |
|--|--|---|--|
| | | ILP ref [King RD, Muggleton SH, Srinivasan A, Sternberg JE (1996), Proc Natl Acad Sci USA 93:438] | TvT Experiment: Da (188 Training, 42 Test) |
|--|--|---|--|

Elaborare specificații și metodologie la identificarea dependenței liniare

| Coeficient de corelație | Definiție | Formule de calcul |
|---|---|--|
| 1. Coeficient de corelație Pearson | Masă a intensității și direcției relației lineare a două variabile cantitative; | $r_{Prs} = \frac{\sum (Y_{m-i} - \bar{Y}_m)(Y_{est-i} - \bar{Y}_{est})}{\sqrt{(\sum (Y_{m-i} - \bar{Y}_m)^2)(\sum (Y_{est-i} - \bar{Y}_{est})^2)}}$ unde Y_{m-i} sunt valorile măsurate, iar \bar{Y}_m este valoarea medie măsurată; |
| 2. Coeficient de corelație Spearman | Masă (neparametrică) a corelației între două variabile (care nu sunt neaparat cantitative) despre a căror distribuții de frecvențe nu se face nici o presupunere; | $r_{Spm} = \frac{\sum (R_{Y_{m-i}} - \bar{R}_{Y_m})(R_{Y_{est-i}} - \bar{R}_{Y_{est}})}{\sqrt{(\sum (R_{Y_{m-i}} - \bar{R}_{Y_m})^2)(\sum (R_{Y_{est-i}} - \bar{R}_{Y_{est}})^2)}}$ unde $R_{Y_{m-i}}$ este rangul activității măsurate pe componenta i , iar $\bar{R}_{Y_{est-i}}$ este media activităților estimate; |
| 3. Coeficient de corelație Spearman-Pearson | Masă (neparametrică) a corelației între două variabile semi-cantitative; | $r_{sq} = \frac{\sum (Y_{m-i} - \bar{Y}_m)(Y_{est-i} - \bar{Y}_{est})}{\sqrt{(\sum (Y_{m-i} - \bar{Y}_m)^2)(\sum (Y_{est-i} - \bar{Y}_{est})^2)}} \cdot \frac{\sum (R_{Y_{m-i}} - \bar{R}_{Y_m})(R_{Y_{est-i}} - \bar{R}_{Y_{est}})}{\sqrt{(\sum (R_{Y_{m-i}} - \bar{R}_{Y_m})^2)(\sum (R_{Y_{est-i}} - \bar{R}_{Y_{est}})^2)}}$ |
| 4. Coeficient de corelație Kendall-tau | Masă (neparametrică) a corelației între două variabile ordinale; -exista 3 coeficienți de corelație Kendall: tau-a, tau-b si tau-c | $\tau_{Ken,a} = (C-D)/[n(n-1)/2];$ $\tau_{Ken,b} = (C-D)/\sqrt{[(n(n-1)/2-t)(n(n-1)/2-u)]}$ $\tau_{Ken,c} = 2(C-D)/n^2,$ unde C=numărul perechi concordante; D= numărul perechi disconcordante; E=numărul perechi egale; |

| | | |
|----------------------------------|--|---|
| 5. Coeficient de corelație Gamma | Măsură a asocierii între două variabile, care conțin valori lipsă; | $\Gamma = (C-D)/(C+D)$, unde C = numărul perechi concordante (perechile $(Y_{m-j}; Y_{est-j})$ și $(Y_{m-i}; Y_{est-i})$ sunt concordante dacă diferențele $Y_{m-j} - Y_{m-i}$ și $Y_{est-j} - Y_{est-i}$ au același semn); D = numărul perechi disconcordante (dacă diferențele $Y_{m-j} - Y_{m-i}$ și $Y_{est-j} - Y_{est-i}$ au semne diferite); |
|----------------------------------|--|---|

Interpretarea coeficientului de corelație Pearson:

- ÷ $-1 \leq r_{Prs} \leq 1$;
- ÷ $r_{Prs} \geq 0.80$ indica o legătură puternică de directă proporționalitate între cele două variabile după cum $r_{Prs} \leq 0.5$ indică o corelație slabă între variabile;
- ÷ coeficientul de determinare r_{Prs}^2 oferă informații despre procentul în care variația variabilei dependente se datorează relației lineare;

Semnificația testului:

Testul Student poate fi utilizat pentru a determina dacă valoarea coeficientului de corelație este semnificativă statistic, cu nivelul de semnificație de 0.01 (sau 0.05);

- ÷ Ipoteza nulă $H_0: r_{Prs} = 0$ (nu există corelație între variabile)
- ÷ Ipoteza alternativă $H_1: r_{Prs} < 0$ (nu există corelație între variabile)

Dacă pentru un nivel de semnificație de 0.01, p-value asociat lui $t_{Prs,df}$ este mai mic decât 0.01 înseamnă că se respinge ipoteza nulă, adică acceptăm ipoteza alternativă. Există corelație între variabilele studiate;

Formula folosită pentru calculul parametrului testului Student este:

$$t = \sqrt{n-2} \cdot \frac{r^2}{\sqrt{1-r^2}}$$

Formula folosită pentru calculul parametrului testului Fisher este:

$$F = \frac{(n-2) \cdot r^2}{1-r^2}$$

Interpretarea coeficientului de corelație Spearman:

Pentru calculul coeficientului Spearman se definesc rangul asociat fiecărei valori măsurate ($RankY_m$) și rangul asociat fiecărei valori estimate ($RankY_{est}$), ranguri ce corespund poziției respectivei valori în seria ordonată. În cadrul acestei etape de definire a rangurilor, se vor lua în considerare cel mai mare și

cel mai mic rang , cu observația că pentru două valori egale, rangul este definit ca fiind media rangurilor corespunzătoare celor două valori.

Interpretarea coeficientului de corelație Spearman este identică cu cea a coeficientului de corelație Pearson.

Semnificația testului:

Testul Student poate fi utilizat pentru a determina dacă valoarea coeficientului de corelație este semnificativă statistic, cu nivelul de semnificație de 0.01

÷ Ipoteza nulă $H_0: r_{Spm} = 0$ (nu există corelație între perechi de ranguri)

÷ Ipoteza alternativă $H_1: r_{Spm} \neq 0$ (există corelație între perechi de ranguri)

Dacă pentru un nivel de semnificație de 0.01, p-value asociat lui $t_{Prs,df}$ este mai mic decât 0.05 înseamnă că se respinge ipoteza nulă, adică acceptăm ipoteza alternativă. Există corelație între variabilele studiate.

Formula folosită pentru calculul parametrului testului Student este:

$$t = \frac{r_{Sp} \cdot \sqrt{n-2}}{\sqrt{1-r_{sp}^2}}$$

Interpretarea coeficientului de corelație Pearson-Spearman(Spearman-semicantitativ):

-identică cu cea a coeficientului de corelație Pearson;

Semnificația testului:

Testul Student poate fi utilizat pentru a determina dacă valoarea coeficientului de corelație este semnificativă statistic, cu nivelul de semnificație de 0.01

÷ Ipoteza nulă $H_0: r_{sQ} = 0$ (nu există corelație între perechi de ranguri)

÷ Ipoteza alternativă $H_1: r_{sQ} \neq 0$ (nu există corelație între perechi de ranguri)

Dacă pentru un nivel de semnificație de 0.01, p-value asociat lui $t_{Prs,df}$ este mai mic decât 0.01 înseamnă că se respinge ipoteza nulă, adică acceptăm ipoteza alternativă. Există corelație între variabilele studiate;

Formula folosită pentru calculul parametrului testului Student este:

$$t = \frac{r_{p_s} \cdot \sqrt{n-2}}{\sqrt{1-r_{p_s}^2}}$$

Interpretarea coeficientului de corelație Kendall

÷ dacă există un acord între două ranguri și cele două ranguri sunt identice atunci coeficientul de corelație este egal cu 1;

÷ dacă există un dezacord între două ranguri și unul din cele două ranguri este inversul celuilalt atunci coeficientul de corelație este egal cu -1;

- ÷ pentru toate celelalte situații, coeficientul de corelație va avea valori cuprinse între -1 și 1, valori pozitive obținându-se pentru acord crescător;
- ÷ dacă rangurile sunt independente atunci coeficientul de corelație este egal cu 0;

Semnificația testului:

Testul Z poate fi utilizat pentru a determina dacă valoarea coeficientului de corelație este semnificativă statistic, cu nivelul de semnificație de 0.05.

Coeficientul de corelație Kendall tau-a:

- ÷ Ipoteza nulă H_0 : $\tau_{\text{Ken},a} = 0$ (nu există corelație)
- ÷ Ipoteza alternativă H_1 : $\tau_{\text{Ken},a} \neq 0$ (există corelație)

Formula folosită pentru calculul parametrului testului Z în cazul coeficientului de corelație Kendall tau-a este:

$$Z_{\tau_{\text{Ken},a}} = \frac{C - D}{\sqrt{S_{\tau_{\text{Ken},a}}^2}}$$

unde variația este dată de formula:

$$S_{\tau_{\text{Ken},a}}^2 = \frac{n(n-1)(n+5)}{9}$$

Coeficientul de corelație Kendall tau-b:

- ÷ Ipoteza nulă H_0 : $\tau_{\text{Ken},b} = 0$ (nu există corelație)
- ÷ Ipoteza alternativă H_1 : $\tau_{\text{Ken},b} \neq 0$ (există corelație)

Formula folosită pentru calculul parametrului testului Z în cazul coeficientului de corelație Kendall tau-b este:

$$Z_{\tau_{\text{Ken},b}} = \frac{C - D}{\sqrt{S_{\tau_{\text{Ken},b}}^2}}$$

unde variația este dată de formula:

$$S_{\tau_{\text{Ken},b}}^2 = \sqrt{S_{\tau_{\text{Ken},a}}^2 - Tied_{\tau_{\text{Ken},b}}}$$

Coeficientul de corelație Kendall tau-c:

- ÷ Ipoteza nulă H_0 : $\tau_{\text{Ken},c} = 0$ (nu există corelație)
- ÷ Ipoteza alternativă H_1 : $\tau_{\text{Ken},c} \neq 0$ (există corelație)

Elaborare specificații și metodologie la identificarea regresiei simple

Următorul tabel sumarizează metodele de regresie simplă:

| Nr | Metoda | Varianta | Ecuția | Formula de calcul |
|----|---------|----------|--------|-------------------|
| 1 | liniară | 1 | $Y=aX$ | $a = M(Y)/M(X)$ |
| 2 | liniară | 2 | $Y=aX$ | $(Y-aX)^2 = \min$ |

| | | | | |
|--|-------------|---|--------------------|---|
| 3 | liniară | 3 | $Y=aX$ | $(X-Y/a)^2 = \min$ |
| 4 | liniară | 4 | $Y=aX$ | $a = \sqrt{M(Y^2)/M(X^2)}$ |
| 5 | liniară | 5 | $Y=aX+b$ | $(aX/b-Y/b+1)^2 = \min$ |
| 6 | liniară | 6 | $Y=aX+b$ | $(Y-aX-b)^2 = \min$ |
| 7 | liniară | 7 | $Y=aX+b$ | $(X-Y/a+b/a)^2 = \min$ |
| 8 | liniară | 8 | $Y=aX+b$ | $a = \sqrt{M(Y^2)/M(X^2)}$ $b= M(Y)-aM(X)$ |
| 9 | hiperbolică | 1 | $(X+a)(Y+b)-c=0$ | a, b și c după formulele de mai jos |
| $C(X, Y) = M(XY) - M(X)M(Y)$, $c = ab + aM(Y) + bM(X) + M(XY)$ $a = \frac{M(XY)C(X, XY) - M(X^2)C(Y, XY) + M(X)M(Y)(M(XY^2) - M(X^2Y))}{M(Y^2)C(X, X) - M(XY)C(X, Y) + M(X)M(Y)(M(XY) - M(X^2))}$ $b = \frac{M(XY)C(Y, XY) - M(Y^2)C(X, XY) + M(X)M(Y)(M(X^2Y) - M(XY^2))}{M(Y^2)C(X, X) - M(XY)C(X, Y) + M(X)M(Y)(M(XY) - M(X^2))}$ | | | | |
| 10 | hiperbolică | 2 | $Y=a+bX^{-1}$ | $(X(Y-a)-b)^2 = \min$ |
| 11 | hiperbolică | 3 | $Y^{-1}=a+bX^{-1}$ | $(Y^{-1}-a-bX^{-1})^2 = \min$ |

Elaborare specificații și metodologie la predicție bazată pe regresie (metoda leave-one-out):

Metoda constă în următoarele etape:

- ÷ determinarea valorilor estimate (YY) pe baza ecuației modelului de regresie lineară: $Y=aX+b$
- ÷ eliminarea pe rând a câte unei valori din mulțimea valorilor observate;
- ÷ determinarea valorilor prezise pe baza ecuației noului model de regresie lineară, în care coeficienții (a_i) și (b_i), $1 \leq i \leq n$, n = numărul de valori observate (măsurate) sunt obținuți prin minimizarea sumei pătratelor distanțelor între valorile observate și cele estimate:

$$YYY_i = a_iX_i+b_i, (a_i, b_i) \text{ din } \sum_{\substack{j=1 \\ j \neq i}}^n (Y_j - a_iX_j - b_i)^2 = \min, i = 1..n$$

Elaborare specificații și metodologie la metoda de calcul a intervalului de încredere pentru variabile distribuite binomial

Definiția formală a intervalului de confidență este aceea că un interval de confidență oferă un șir de valori estimate, șir care este cel mai probabil să includă un parametru necunoscut al populației, valorile estimate fiind calculate pe un eșantion.

Dacă din populație se extrag in mod repetat, eșantioane independente și se calculează câte un interval de confidență pentru fiecare eșantion, atunci un anumit procent al intervalelor va conține

parametrul necunoscut al populației. În general intervalele de confidență sunt calculate astfel încât procentul este 95%. Totuși, procentul poate fi generic α .

Se știe ca variabilele cantitative și calitative ce rezultă din observabile pot fi clasificate în două distribuții teoretice. Variabilele continue urmează de regulă, distribuția normală (Laplace-Gauss) în timp ce variabilele discrete urmează o distribuție binomială.

Distribuția normală a fost introdusă de De Moivre în contextul aproximării unei distribuții binomiale cu un n mare. Mai târziu rezultatele au fost extinse de Laplace în teorema care acum este cunoscută sub numele de Teorema lui De Moivre-Laplace.

Estimarea intervalelor de încredere pentru proporții utilizând o distribuție normală a fost foarte utilizată în simulare, mai ales din cauza faptului că în practică era mai ușor ca aproximarea normală să fie utilizată pentru compararea cu alți estimatori apropiați.

Dacă pe un eșantion de talie n considerăm variabila X ce urmează o distribuție binomială atunci probabilitatea de a obține valoarea Y ($0 \leq Y \leq n$) este dată de formula:

$$P_B(n, X, Y) = \frac{n!}{Y!(n-Y)!} \cdot \frac{X^Y (n-X)^{n-Y}}{n^n}$$

Media, respectiv variația distribuției binomiale sunt:

$$M(n, X) = X,$$

$$\text{Var}(n, X) = \frac{X \cdot (n - X)}{n}$$

Probabilitatea de a obține variabila normală Y care are media $M(n, X)$ și variația $\text{Var}(n, X)$ este:

$$P_N(n, X, Y) = \frac{1}{\sqrt{2\pi \text{Var}(n, X)}} \cdot e^{-\frac{(Y - M(n, X))^2}{2 \text{Var}(n, X)}}$$

Întrucât noua variabilă normală depinde doar de medie și dispersie, înlocuind media și dispersia din formula distribuției binomiale s-a obținut:

$$P_N(n, X, Y) = \frac{(Y - X)^2}{\sqrt{2X(n - X)/n}} \cdot e^{-\frac{(Y - X)^2}{2X(n - X)/n}}$$

Eroarea de aproximare a distribuției binomiale a variabilei Y printr-o distribuție normală este dată de formula:

$$\text{Err}(n, X, Y) = P_B(n, X, Y) - P_N(n, X, Y)$$

Având în vedere că probabilitatea lui Y descrește odată cu creșterea dispersiei lui X , funcția de eroare poate fi discretizată prin:

$$\text{Err}_c(n, X, Y) = \begin{cases} P_B(n, X, Y) - P_N(n, X, Y), & P_B(n, X, Y) > 1/n \\ 0, & P_B(n, X, Y) < 1/n \end{cases}$$

Evaluarea intervalelor de încredere pentru variabile distribuite binomial se pot evalua cu metodele date în tabelul următor:

| Metoda | Formula |
|--------|---|
| AvgOEA | $\left 100\alpha - \sum_{X=A}^{n-A} \frac{\epsilon_{X,n}^M}{n+1-2A} \right $ |
| StDOEA | $\sqrt{\sum_{X=A}^{n-A} \frac{(\epsilon_{X,n}^M - \text{AvgOEA})^2}{n-2A}}$ |
| SiDOEA | $\sqrt{\sum_{X=A}^{n-A} \frac{(\epsilon_{X,n}^M - 100\alpha)^2}{n+1-2A}}$ |
| AvADAA | $\sum_{X=A}^{n-A} \frac{ \epsilon_{X,n}^M - \text{AvgOEA} }{n-2A}$ |
| AvADSA | $\sum_{X=A}^{n-A} \frac{ \epsilon_{X,n}^M - 100\alpha }{n+1-2A}$ |
| S8DOEA | $\sqrt[8]{\sum_{X=A}^{n-A} \frac{(\epsilon_{X,n}^M - 100\alpha)^8}{n+1-2A}}$ |

Eroarea experimentală pentru distribuția binomială corespunzătoare unui eșantion de volum n este dată de formula:

$$\text{Err}(n, X, Y) = \frac{\sum_{CI_L^M(n, X, Y) > X} d\text{Bin}(n, X, Y) + \sum_{CI_U^M(n, X, Y) > X} d\text{Bin}(n, X, Y)}{\sum_{Y=1}^{n-1} d\text{Bin}(n, X, Y)}$$

unde $d\text{Bin}(X, Y)$ este probabilitatea distribuției binomiale definită prin relația (1), CI_L^M este limita inferioară a intervalului de încredere determinat prin metoda M , iar CI_U^M este limita inferioară a intervalului de încredere determinat prin metoda M .

Elaborare specificații și metodologie la designul experimental

Tehnica experimentului se bazează pe noțiunea de observație. Experimentele sunt întotdeauna empirice și cu toate acestea uneori doar măsurătorile nu formează un experiment.

Experimentele implică mai mult stabilirea unui anumit nivel de control și manipularea unuia sau a mai multor factori de interes decât stabilirea cauzei și efectului.

Definiția din dicționar a experimentului este aceea de a determina eficacitatea unui lucru neîncercat anterior sau aceea de a examina validitatea unei ipoteze sau de a demonstra un adevăr cunoscut.

Prima etapă în formularea unui experiment este definirea câtorva termeni, cum ar fi:

÷ Propoziții, ipoteze

- ÷ Presupunerile ce trebuiesc făcute (ele determină scopul și identifică domeniile care nu pot fi investigate);
- ÷ Identificarea variabilei (variabilelor) dependente (sunt datele de ieșire, rezultate);
- ÷ Identificarea variabilelor independente (sunt datele de intrare care luate împreună formează spațiul experimental);
- ÷ Care din variabilele independente pot fi controlate;

În general, statisticianul nu studiază un caracter al populației pe întreaga mulțime de elemente, din mai multe motive, dintre care menționăm următoarele:

- ÷ Talia populației poate fi foarte mare sau chiar infinită ceea ce face imposibilă o “observare” exhaustivă a întregii populații;
- ÷ Eșantioanele pot fi studiate mai rapid decât populațiile;
- ÷ Studiul caracterului pe întreaga populație este frecvent imposibil, deoarece poate distruge populația.
- ÷ În anumite situații nu se mai pot obține informații decât despre o parte a populației.
- ÷ Rezultatele observațiilor pe eșantioane adesea sunt mai precise decât rezultatele bazate pe observarea populației în totalitate, deoarece la nivelul unui eșantion se controlează mai ușor procesul și tehnicile de observare, acestea menținându-se cu un efort mai mic în standardele de eroare acceptate.
- ÷ Costul și resursele necesare (umane, materiale, etc.) pentru observarea exhaustivă a unei populații pot de asemenea să fie un motiv pentru utilizarea eșantioanelor.

Acestea sunt câteva rațiuni pentru care o populație este studiată cu ajutorul unei submulțimi a ei de talie mai mică care să permită un studiu exhaustiv al ei.

Elaborare specificații și metodologie la metode de eșantionare

Un bun eșantion trebuie să constituie o imagine adecvată și fidelă a întregii populații pentru care se dorește studierea unui caracter. În caz contrar, se spune că eșantionul este nereprezentativ, sau cu "bias". Alegerea eșantionului și culegerea datelor constituie partea cea mai laborioasă. Modalitatea de alegere a eșantionului, principiile de eșantinare, criteriile de includere și excludere precum și modalitatea de culegere, colectare și stocare a datelor trebuie să se regăsească în protocolul cercetării. În scopul extrapolării rezultatelor la întreaga populație din cadrul căreia eșantionul a fost creat (obiectiv al statisticii inductive), acesta (adică eșantionul) trebuie să îndeplinească anumite criterii (trebuie să fie reprezentativ). Pentru a obține reprezentativitatea un eșantion trebuie să îndeplinească următoarele condiții:

- ÷ Reprezentativitatea prin taliei: condiție de ordin *cantitativ*. Volumul eșantionului trebuie să fie suficient de mare raportat la populația din care s-a extras.

÷ Reprezentativitatea prin caracteristici: condiție de ordin *calitativ*. Eșantionul trebuie *extras aleator* din populație pentru a cuprinde în mod proporțional toate caracteristicile populației din care face parte.

Cel mai bun mijloc de a asigura că un eșantion va permite inferențe corecte este utilizarea metodelor de *eșantionare probabilistică* în obținerea eșantionului. Această metodă, pentru fiecare subiect al populației este cunoscută probabilitatea (șansa) de a fi inclus în eșantion.

Patru metode de eșantionare sunt folosite pentru a obține eșantioane probabiliste:

- ÷ eșantionarea simplu randomizată
- ÷ eșantionarea sistematică
- ÷ eșantionarea stratificată
- ÷ eșantionarea cluster.

Eșantionarea simplă randomizată este o selecție formată din subiecți extrași la întâmplare din populația statistică. Fiecare subiect are aceeași șansă de a fi inclus în eșantion.

Astfel pentru obținerea unui eșantion aleator se poate utiliza o metodă de randomizare bazată pe procedee de generare de numere aleatoare (de exemplu, funcțiile RAND sau RANDBETWEEN din Microsoft EXCEL). Prin această metodă, fiecare element al populației primește un număr de identificare. Pentru eșantionare este disponibilă o listă de numere de identificare numită structură de eșantionare.

Eșantionarea sistematică include tot al k-lea element din populație în eșantion. Numărul k se obține împărțind talia populației la talia dorită a eșantionului.

Eșantionarea sistematică nu este indicat să fie folosită atunci când în structura de eșantionare ar putea apare o periodicitate.

Există și alte metode de metode de eșantionare mai complexe, cum ar fi, de exemplu, *eșantionarea stratificată*. Pentru aceasta, se împarte populația în mai multe subgrupe relevante numite *straturi* și se constituie eșantionul prin extrageri aleatoare din straturi. Fiecare strat trebuie să fie reprezentat în eșantion în funcție de importanța sa în populație.

Eșantionarea cluster. Un *eșantion aleator de clusteri* se obține printr-un procedeu în două etape. Într-o primă etapă se împarte populația în clusteri și ulterior se selectează aleator o submulțime de clusteri. În mod obișnuit alegerea clusterilor se bazează pe criterii geografice, iar acest procedeu este utilizat frecvent în studiile epidemiologice.

Eșantionarea nonprobabilistă este aceia în care nu este cunoscută probabilitatea ca o entitate din populație să fie selectată. Eșantioanele obținute în acest mod au frecvent bias de selecție.

Asignarea aleatoare. În anumite studii experimentale, prima dată, pe baza unor criterii subiecții sunt selectați pentru a fi incluși în studiu. Apoi fiecăruia dintre acești subiecți trebuie să li se atribuie un anumit tratament. Dacă această atribuire a tratamentului este aleatoare atunci procedeu se numește

atribuire aleatoare. Acest procedeu de asignare aleatoare are ca scop ca grupurile ce primesc diferite tratamente să fie cât mai asemănător posibil.

Populația țintă și populația de disponibilă. În anumite studii, subiecții incluși în eșantionul de studiu nu aparțin întotdeauna populației pentru care cercetătorul dorește să generalizeze concluziile studiului. În locul acestei populații, numită populație țintă, cercetătorul utilizează frecvent o populație de subiecți disponibili care verifică anumite condiții impuse. Această populație din care de fapt este prelevat eșantionul (sau eșantioanele) se numește *populația disponibilă*. Pentru a face inferențe relativ la populația țintă pe baza populației disponibile, aceasta din urmă trebuie să fie reprezentativă pentru populația țintă (cele mai importante caracteristici au aceeași distribuție în populația țintă și în populația disponibilă).

În anumite tipuri de studii sunt necesare cel puțin două eșantioane. Aceste eșantioane pot fi extrase sau prelevate din populația țintă în două feluri: dependent sau independent.

În cazul eșantioanelor independente:

÷ prelevarea unuia nu influențează prelevarea celuilalt,

÷ nu au, în general, același număr de unități.

Un caz al eșantioanelor dependente este cel implicând două eșantioane, care se spune că sunt eșantioane de observații perechi dacă au aceeași talie și unitățile unui eșantion sunt în corespondență bijectivă cu unitățile celuilalt eșantion.

Datele statistice au diverse proveniențe. Un prim tip este cel al *datelor obținute din măsurători*, care rezultă pe baza unor determinări cantitative ale unor proprietăți susceptibile să varieze, în principiu de o manieră continuă, cum ar fi, spre exemplu, înălțimea, greutatea, presiunea sangvină, glicemia.

Alte date statistice rezultă din enumerare, operație care furnizează în mod necesar date întregi. Aceste *date de enumerare* se obțin de regulă ca fiind numărul de indivizi care îndeplinesc o anumită caracteristică, regula fiind stabilită în urma unor operații de clasificare după criterii bine stabilite. Rezultatele datelor de enumerare se exprimă frecvent sub forma de procente: rata de piese care îndeplinesc condițiile calitative necesare vânzării a fost de 65,5%.

O altă categorie de date sunt *datele de înscriere (ordinale sau de ordonare)*, care reprezintă poziția unor obiecte "clasament" stabilit după anumite criterii. *Datele de ordonare* sunt frecvent utilizate, cum este de exemplu pentru descrierea calității unui produs.

Elaborare specificații și metodologie la colectarea datelor statistice

Clasificarea statistică a datelor poate fi realizată ținând seama de scalele de măsură utilizate.

Astfel se disting următoarele scale de măsură:

- ÷ Scala **nominală** este o scală pentru măsurarea variabilelor calitative ce pot lua un număr finit de valori. Scala nominală nu admite nici o proprietate aritmetică și nici ordonarea valorilor.

Datele evaluate după o scală nominală sunt numite *observații calitative*, deoarece ele descriu o calitate a unui obiect studiat. Unele dintre aceste scale au doar două valori și atunci observațiile sunt binare.

- ÷ Scala **ordinală** este o scală utilizată în cazul variabilelor care pot lua valori într-o mulțime discretă finită de valori. Nu are nici o proprietate aritmetică. Permite o ordonare a valorilor.
- ÷ Scala **interval** este o scală utilizată în cazul variabilelor cantitative continue (ce pot lua valori într-un interval) și pentru care diferența între două valori ale scalei are sens.
- ÷ Scala de tip **rație sau raport** este utilizată în cazul variabilelor cantitative continue pentru care atât diferența cât și câtul a oricăror două valori de pe scală au sens. Această scală are un zero absolut. Nu acceptă valori negative.

Elaborare specificații și metodologie la colectarea datelor

Culegerea datelor se realizează în funcție de:

- ÷ obiectivele propuse
- ÷ tipul studiului
- ÷ modalitățile de alegere a subiecților
- ÷ timpul disponibil
- ÷ resursele financiare și umane disponibile
- ÷ procedura folosită
- ÷ accesul la date.

În funcție de populația cuprinsă în studiu culegerea datelor poate fi:

- ÷ *Exhaustivă* - cuprinzând toți subiecții populației țintă. Din considerente financiare (costuri foarte ridicate, echipă numeroasă necesară pentru culegerea datelor) și/sau economice (distrugerea populației de interes) culegerea exhaustivă a datelor este practic imposibilă.
- ÷ *Prin eșantionare* - alegând din populația țintă sau din populația disponibilă un grup de subiecți - eșantionul. Pentru caracteristicile eșantionului reprezentativ vezi textul anterior.

În funcție de durata culegerii datelor poate fi:

- ÷ *Transversală*: studierea unui eșantion la un moment X determinat în timp și/sau spațiu.

÷ *Longitudinală*: studierea unui eșantion într-un interval de timp prestabilit. În culegerea de date de tip longitudinal în funcție de accesul la date culegerea poate fi:

- *Retrospectiv*: din înregistrările anterioare (acolo unde este posibil), de la un moment M la un moment N (unde M este posterior momentului N și anterior momentului în care protocolul de cercetare a fost stabilit).
- *Prospectiv*: cercetare sistematică a datelor care se vor întâmpla de la un moment Y la un moment Z (unde momentul Y este anterior momentului Z și posterior momentului în care protocolul de cercetare a fost stabilit)

În studiile clinice observaționale, cele în care cercetarea se bazează pe urmărirea în evoluția sa naturală a unui fenomen biologic (fără intervenția experimentatorului), culegerea datelor se realizează în funcție de modul alcătuirii grupei/grupelor de subiecți luați în studiu.

Se pot evidenția astfel trei tipuri de culegere a datelor.

÷ *Culegere de tip eșantion reprezentativ* – în care grupul de subiecți studiați este reprezentativ pentru populația țintă la care se vor extrapola rezultatele obținute. Acest tip de eșantion poate fi la fel de bine utilizat pentru un studiu descriptiv (descrierea unui fenomen de sănătate) cât și pentru unul analitic (în care este obligatorie compararea a cel puțin două grupuri de subiecți).

| | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|
| | B⁺ | B⁻ | |
| FR⁺ | a | b | ne⁺ |
| FR⁻ | c | d | ne⁻ |
| | nB⁺ | nB⁻ | N |

În acest tip de culegere a datelor se fixează de la început numărul total al subiecților luați în studiu = N (calculabil ca talie după criteriile de eșantionare, în funcție în principal de dimensiunea populației țintă și prevalența afecțiunii studiate).

Datele culese pot fi înscrise într-un tabel de contingență 2x2 (tabel ce cuprinde rezultatele a două sau mai multe seturi de observații discrete prezentate astfel în scopul analizării relației ce există între ele), diferențiind astfel patru categorii de subiecți:

- a= veritabil pozitivi – bolnavi cu prezența factorului de risc
- b= fals pozitivi – indemni de boală, dar cu factorul de risc prezent
- c= fals negativi – bolnavi fără factorul de risc
- d= veritabili negativi – indemni de boală și fără factori de risc.

Aceste categorii sunt realizate în funcție de apartenența subiectului la următoarele grupuri:

- FR^+ = factor de risc prezent
- FR^- = factor de risc absent
- B^+ = purtător al bolii
- B^- = indemn de boală (nu se utilizează termenul de sănătos pentru că acesta presupune a fi indemn de orice altă boală și în afara celei studiate).

- ne^+ = numărul total al celor expuși la factorul de risc
- ne^- = numărul total al celor neexpuși la factorul de risc
- nB^+ = numărul total al celor bolnavi (purători ai bolii luate în studiu)
- nB^- = numărul total al celor indemni de boala luată în studiu.

Asupra grupelor astfel alcătuite se vor putea efectua prelucrări statistice. Acest tip de culegere a datelor permite efectuarea celor mai multe tipuri de calcule, neexistând îngrădiri de eșantionare astfel:

- calcularea prevalenței bolii
- calcularea frecvenței expunerii
- măsurarea gradului de asociere între factorul incriminat (de risc sau prognostic) și boală.

Limitele acestui tip de culegere rezidă în dificultățile de urmărire a unui număr mare de subiecți ridicând probleme de cost, riscul scăderii interesului subiecților pentru studiu, riscul numărului mare de pierduți din vedere.

÷ *Culegerea de tip expus - nonexpus* împarte subiecții luați în studiu în două categorii respectiv cei expuși și cei neexpuși la factorul de risc (sau de prognostic) studiat. Cercetarea poate urmări frecvența evenimentelor induse de factorul de risc luat în studiu, eventual diferit în grupul expus față de nonexpuși. Se fixează factorul de expunere și se studiază apariția în timp a îmbolnăvirii (boala în acest tip de culegere fiind factorul aleator).

Numărul subiecților din cele două grupe poate fi identic sau diferit.

| | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|
| | B⁺ | B⁻ | |
| FR⁺ | a | b | ne⁺ |
| FR⁻ | c | d | ne⁻ |
| | nB⁺ | nB⁻ | N |

Prin prestabilirea celor două grupe ne^+ și ne^- pe baza unor criterii clare de incluziune și excluziune (criterii de eligibilitate pentru apartenența unui subiect la un anumit grup de studiu), calculele statistice se pot efectua între componenta de deasupra și sub linia îngroșată a tabelului de contingență. Nu sunt posibile comparații și calcule între componentele din stânga și dreapta tabelului, studiul urmărind de fapt apariția afecțiunii la cele două grupe prestabilite. Se pot compara așadar între cele două grupe proporția de subiecți bolnavi dau nu și frecvența expunerii.

Avantajului tipului de culegere expus - nonexpus rezidă în posibilitatea alcătuirii de grupe mai mici de subiecți dar ca neajuns major apar multiple probleme de urmărire a subiecților.

÷ *Culegerea de tip "caz - martor"* începe prin predefinirea a două grupe de subiecți:

- B^+ = bolnavi = grupul caz
- B^- = indemni de boală = grupul martor,

criteriul principal de incluziune sau excluziune într-un grup sau altul fiind prezența sau absența bolii.

Se fixează (se prestabilește) deci boala și se studiază în general retrospectiv (în trecut) prezența expunerii (factorului de risc sau de prognostic) în apariția (sau evoluția) îmbolnăvirii. Astfel în acest tip de culegere boala devine factorul controlat iar expunerea factorul aleator,

| | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|
| | B⁺ | B⁻ | |
| FR⁺ | a | b | ne⁺ |
| FR⁻ | c | d | ne⁻ |
| | nB⁺ | nB⁻ | N |

nB^+ putând fi identic sau diferit de nB^- .

Calculul și comparații se pot face numai între grupele prestabilite B^+ și B^- , respectiv coloana din dreapta și din stânga a tabelului de contingență, despărțite prin linia verticală îngroșată. Pot fi comparate proporțiile celor expuși dar nu poate fi calculată frecvența bolii.

Rezultate semnificative prin acest tip de culegere a datelor se pot obține și la grupe cu număr mai mic de subiecți, ca în cazul eșantionului reprezentativ; de aceea modul de culegere corespunde pentru studiul îmbolnăvirilor cu cazuri rare.

Culegerea datelor realizându-se în momentul când boala este deja cunoscută, pot să apară aprecieri subiective atât din partea celui care culege datele (experimentator) cât și a subiectului urmărit (bolnavul). Apare deci riscul mare al erorilor sistematice de culegere.

Elaborare specificații și metodologie la colectarea datelor experimentale

Ideea de design al experimentului a fost introdusă de Fisher în 1920 care a demonstrat utilitatea acestui concept în domeniul agriculturii, analizând problema de optimizare a recoltelor ținând cont de diverși factori (apă, ploaie, soare, condiții de sol). În 1986, Taguchi a continuat ideea de design experimental, prin împartirea problemelor de optimizare în două categorii:

÷ **Probleme statice** (problemele în care există câțiva factori de control ce decid valoarea dorită a funcției). Metode folosite în astfel de probleme sunt:

- Aproximarea Smaller-the-Better este folosită când:
 - Valoarea ideală pentru toate caracteristicile nedorite este zero
 - Valoarea ideală este finită și este definit maximul sau minimul ei
- Aproximarea Nominal-the-Best este folosită când este așteptată o anumită valoare și nu se dorește nici maximul nici minimul;
- Aproximarea Larger-the-Better are la bază un model care cuprinde trei etape:
 - proiectarea sistemului
 - proiectarea parametrilor
 - proiectarea toleranțelor.

În metoda Taguchi etapa de proiectare a sistemului, parametrilor și a toleranțelor se face în așa fel încât să se obțină un proiect ROBUST, adică insensibil, pe cât posibil, la variații normale ale mediului produsului.

În acest scop, Taguchi a introdus noțiunea de *parametri critici de proiectare* pentru acei parametri a căror variație influențează foarte mult calitatea / funcționalitatea produsului. În principiu, acești parametri trebuie controlați în proiect prin toleranțe foarte strânse, în timp ce restul parametrilor trebuie lăsați cu toleranțe relaxate pentru a scădea costurile.

Taguchi propune abordarea problematicii prin 2 instrumente:

- Elaborarea funcției obiectiv pe baza unuia din modelele de mai sus
- Tehnica experimentelor (*metoda matricelor ortogonale*)

Metoda șirurilor ortogonale este o metodă de optimizare a proceselor industriale, care asemenea altor tehnici de optimizare are ca și obiective:

- Profit maxim
- Configurație cât mai bună a mașinilor
- Alocare optimă a materialelor
- Alocare optimă a forțelor de muncă în timp minim

Metoda Taguchi constă în următoarele etape:

- identificarea factorilor controlabili și necontrolabili
- conceperea tabelelor de testare, separat pentru factorii controlabili și cei necontrolabili
- construirea matricei ortogonale a experimentelor prin combinarea tabelelor de testare pe orizontală și pe verticală
- efectuarea experimentelor ce corespund fiecărei căsuțe din matricea ortogonală, pe orizontală în tabloul factorilor controlabili și pe verticală în tabloul factorilor necontrolabili
- calculul raportului semnal / zgomot pentru fiecare linie a matricei ortogonale.

÷ **Probleme dinamice (există un parametru de intrare care decide rezultatul):**

- Senzitivitatea pantei: panta trebuie să aibă o valoare dată (de obicei 1) când funcția rezultat este:
 - O caracteristică nedorită (poate fi tratată ca o problemă de tip Smaller-the-Better)
 - O caracteristică dorită (poate fi tratată ca o problemă de tip Larger-the-Better)
- Linearitatea (Larger-the-Better): este folosită când caracteristicile dinamice trebuie să aibă proporționalitate între intrare și rezultat.

De fapt obiectivul optimizării procesului de prelucrare ar fi: calitate maximă-pierderi minime-productivitate maximă. Pentru determinarea celor mai bune valori ale parametrilor implicați în procesul de prelucrare, se poate utiliza analiza factorială.

În opoziție cu analiza factorială, aproximarea lui Taguchi reduce numărul de experimente la un model rezonabil în termeni de cost și timp, prin folosirea șirurilor ortogonale. Ea este utilizată în domenii diferite ca: fizică, chimie, științe agricole, statistică, management și afaceri, medicină.

Alegerea șirurilor ortogonale potrivite pentru problema studiată este principala dificultate a aproximării lui Taguchi. În literatura de specialitate sunt cunoscute multe șiruri ortogonale însă nu a fost găsită o schemă completă care să conțină toate șirurile ortogonale corespunzătoare chiar și unui număr mic de experimente.

Activitatea 7. Testare și validare model

Modelul analizei de dependențe liniare a fost implementat în PHP. În continuare se prezintă testarea și validarea acestuia pe un set de date din literatura de specialitate:

÷ Setul de date

| id | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|----|------|------|------|-------|---------|
| 1 | 0 | 0 | 4.9 | 0 | 0 |
| 2 | 0 | 0 | 5.81 | 0 | 0 |
| 3 | 1.46 | 0.81 | 4.69 | 4.1 | 114 |
| 4 | 1.69 | 0.82 | 4.31 | 4.58 | 126 |
| 5 | 1.97 | 0.82 | 4.77 | 5.11 | 139 |
| 27 | 0.72 | 0.46 | 6.6 | 1.85 | 150 |
| 28 | 0.92 | 0.61 | 3.29 | 2.58 | 157 |
| 29 | 1 | 0.65 | 1.75 | 2.8 | 164 |
| 30 | 1.02 | 0.63 | 1.75 | 2.67 | 160 |
| 31 | 1.14 | 0.68 | 1.4 | 2.74 | 201 |
| 53 | 0 | 0 | 0 | 0 | 0 |
| 54 | 0 | 0 | 1.5 | 0 | 0 |
| 55 | 1 | 0.69 | 1.89 | 2.77 | 187 |
| 56 | 1.2 | 0.69 | 1.61 | 3.24 | 133 |
| 57 | 1.27 | 0.7 | 1.75 | 3.33 | 160 |
| 79 | 0.97 | 0.61 | 5.49 | 2.57 | 160 |
| 80 | 0.97 | 0.6 | 2.81 | 2.47 | 137 |
| 81 | 0.88 | 0.58 | 2.61 | 2.39 | 146 |
| 82 | 0.82 | 0.57 | 2.27 | 2.31 | 156 |
| 83 | 0.79 | 0.55 | 1.98 | 2.23 | 167 |

÷ Analiza completă a dependențelor liniare în setul de date:

- Significant Correlation is set to: 0.707106781186548
- Significant Probability is set to: 0.01
- DESCRIPTIVE
 - number of measurements: 20
 - number of variables: 5
- QUANTITATIVE
 - Pearson's correlation matrix

| r(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|----------|---------------|---------------|---------------|----------------|
| d_IP | 1 | 0.9529 | 0.0929 | 0.9968 | 0.7273 |
| d_IR | | 1 | 0.0024 | 0.9578 | 0.8749 |
| d_Cr | | | 1 | 0.0928 | -0.1078 |
| d_RSD | | | | 1 | 0.7234 |
| d_Volum | | | | | 1 |

r BLUE: Hypothesis of linear dependence can be accepted. **r RED**: Hypothesis of linear dependence is rejected.

- F-ratio matrix

| F(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|----------|--------|-----------|---------|
| d_IP | INF | 177.5184 | 0.1568 | 2801.1466 | 20.2177 |
| d_IR | | INF | 0.0001 | 199.8466 | 58.757 |
| d_Cr | | | INF | 0.1562 | 0.2116 |
| d_RSD | | | | INF | 19.7629 |
| d_Volum | | | | | INF |

- Probability of wrong model matrix

| pF(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|----------|------------------|---------------|------------------|-----------------|
| d_IP | 0 | 9.201e-11 | 0.6968 | 3.280e-21 | 2.792e-4 |
| d_IR | | 0 | 0.992 | 3.461e-11 | 4.480e-7 |
| d_Cr | | | 0 | 0.6973 | 0.651 |
| d_RSD | | | | 0 | 3.124e-4 |
| d_Volum | | | | | 0 |

p BLUE: Hypothesis of linear dependence can be accepted with 99% of success. **p RED**: Hypothesis of linear dependence is rejected for 95% confidence.

- t-ratio matrix

| t(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|--------|--------|---------|---------|
| d_IP | INF | 9.1911 | 0.0366 | 37.3643 | 2.6448 |
| d_IR | | INF | 0 | 9.7784 | 5.0474 |
| d_Cr | | | INF | 0.0365 | 0.0493 |
| d_RSD | | | | INF | 2.6057 |
| d_Volum | | | | | INF |

- Probability of wrong model matrix

| pt(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|----------|-----------------|---------------|------------------|-----------------|
| d_IP | 0 | 3.218e-8 | 0.9712 | 1.637e-18 | 0.0165 |
| d_IR | | 0 | 1 | 1.259e-8 | 8.380e-5 |
| d_Cr | | | 0 | 0.9713 | 0.9612 |
| d_RSD | | | | 0 | 0.0179 |
| d_Volum | | | | | 0 |

p BLUE: Hypothesis of linear dependence can be accepted with 99% of success. **p RED**: Hypothesis of linear dependence is rejected for 95% confidence.

- SEMI-QUANTITATIVE

- semi-Quantitative correlation matrix

| λ (XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|----------------|------|------|------|-------|---------|
|----------------|------|------|------|-------|---------|

| | | | | | |
|---------|----------|---------------|---------------|---------------|---------------|
| d_IP | 1 | 0.9744 | 0.2626 | 0.9963 | 0.7814 |
| d_IR | | 1 | 0.0423 | 0.9779 | 0.8604 |
| d_Cr | | | 1 | 0.263 | 0.2756 |
| d_RSD | | | | 1 | 0.7805 |
| d_Volum | | | | | 1 |

r BLUE: Hypothesis of linear dependence can be accepted. **r RED**: Hypothesis of linear dependence is rejected.

- F-ratio matrix

| F(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|----------|--------|-----------|---------|
| d_IP | INF | 338.3177 | 1.3331 | 2427.7965 | 28.2213 |
| d_IR | | INF | 0.0323 | 394.0679 | 51.2985 |
| d_Cr | | | INF | 1.3371 | 1.4798 |
| d_RSD | | | | INF | 28.0531 |
| d_Volum | | | | | INF |

- Probability of wrong model matrix

| pF(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|----------|------------------|---------------|------------------|-----------------|
| d_IP | 0 | 4.067e-13 | 0.2634 | 1.179e-20 | 4.746e-5 |
| d_IR | | 0 | 0.8594 | 1.096e-13 | 1.141e-6 |
| d_Cr | | | 0 | 0.2627 | 0.2395 |
| d_RSD | | | | 0 | 4.909e-5 |
| d_Volum | | | | | 0 |

p BLUE: Hypothesis of linear dependence can be accepted with 99% of success. **p RED**: Hypothesis of linear dependence is rejected for 95% confidence.

- t-ratio matrix

| t(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|---------|--------|---------|---------|
| d_IP | INF | 12.8365 | 0.2932 | 34.7766 | 3.2709 |
| d_IR | | INF | 0.0076 | 13.8793 | 4.6713 |
| d_Cr | | | INF | 0.2941 | 0.3232 |
| d_RSD | | | | INF | 3.2588 |
| d_Volum | | | | | INF |

- Probability of wrong model matrix

| pt(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|----------|------------------|---------------|------------------|-----------------|
| d_IP | 0 | 1.693e-10 | 0.7727 | 5.857e-18 | 4.245e-3 |
| d_IR | | 0 | 0.994 | 4.692e-11 | 1.900e-4 |
| d_Cr | | | 0 | 0.7721 | 0.7502 |
| d_RSD | | | | 0 | 4.360e-3 |
| d_Volum | | | | | 0 |

p BLUE: Hypothesis of linear dependence can be accepted with 99% of success. **p RED:** Hypothesis of linear dependence is rejected for 95% confidence.

○ QUALITATIVE

- Spearman's correlation matrix

| $\rho(XX)$ | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|------------|----------|---------------|---------------|---------------|---------------|
| d_IP | 1 | 0.9965 | 0.742 | 0.9958 | 0.8395 |
| d_IR | | 1 | 0.7462 | 0.9985 | 0.8461 |
| d_Cr | | | 1 | 0.7455 | 0.7047 |
| d_RSD | | | | 1 | 0.842 |
| d_Volum | | | | | 1 |

r BLUE: Hypothesis of linear dependence can be accepted.

- F-ratio matrix

| F(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|-----------|---------|-----------|---------|
| d_IP | INF | 2528.2032 | 22.0494 | 2142.0179 | 42.9544 |
| d_IR | | INF | 22.6198 | 5818.6434 | 45.346 |
| d_Cr | | | INF | 22.5179 | 17.7616 |
| d_RSD | | | | INF | 43.8601 |
| d_Volum | | | | | INF |

- Probability of wrong model matrix

| pF(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|----------|------------------|-----------------|------------------|-----------------|
| d_IP | 0 | 8.205e-21 | 1.801e-4 | 3.608e-20 | 3.698e-6 |
| d_IR | | 0 | 1.579e-4 | 4.687e-24 | 2.598e-6 |
| d_Cr | | | 0 | 1.616e-4 | 5.212e-4 |
| d_RSD | | | | 0 | 3.230e-6 |
| d_Volum | | | | | 0 |

p BLUE: Hypothesis of linear dependence can be accepted with 99% of success.

- t-ratio matrix

| t(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|---------|--------|---------|---------|
| d_IP | INF | 35.4911 | 2.7981 | 32.6577 | 4.2139 |
| d_IR | | INF | 2.8444 | 53.8965 | 4.3495 |
| d_Cr | | | INF | 2.8362 | 2.4278 |
| d_RSD | | | | INF | 4.2657 |
| d_Volum | | | | | INF |

- Probability of wrong model matrix

| pt(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|--------|----------|------------------|--------|------------------|-----------------|
| d_IP | 0 | 4.082e-18 | 0.0119 | 1.786e-17 | 5.219e-4 |

| | | | | | |
|---------|--|----------|----------|------------------|-----------------|
| d_IR | | 0 | 0.0108 | 2.370e-21 | 3.863e-4 |
| d_Cr | | | 0 | 0.011 | 0.0259 |
| d_RSD | | | | 0 | 4.652e-4 |
| d_Volum | | | | | 0 |

p BLUE: Hypothesis of linear dependence can be accepted with 99% of success.

- Kendall's correlation matrix

| $\tau_a(XX)$ | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|--------------|----------|------------|---------------|---------------|---------------|
| d_IP | 1 | 0.9 | 0.0474 | 0.8947 | 0.2368 |
| d_IR | | 1 | 0.0105 | 0.9316 | 0.2632 |
| d_Cr | | | 1 | 0.0158 | 0.2053 |
| d_RSD | | | | 1 | 0.2368 |
| d_Volum | | | | | 1 |

r BLUE: Hypothesis of linear dependence can be accepted. **r RED:** Hypothesis of linear dependence is rejected.

- Z-ratio matrix

| Z(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|-------|---------|---------|---------|
| d_IP | INF | 5.548 | -0.292 | 5.5155 | 1.46 |
| d_IR | | INF | -0.0649 | 5.7426 | 1.6222 |
| d_Cr | | | INF | -0.0973 | -1.2653 |
| d_RSD | | | | INF | 1.46 |
| d_Volum | | | | | INF |

- Probability of wrong model matrix

| pZ(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|----------|-----------------|---------------|-----------------|---------------|
| d_IP | 0 | 2.890e-8 | 1.2297 | 3.477e-8 | 0.1443 |
| d_IR | | 0 | 1.0517 | 9.321e-9 | 0.1048 |
| d_Cr | | | 0 | 1.0775 | 1.7942 |
| d_RSD | | | | 0 | 0.1443 |
| d_Volum | | | | | 0 |

p BLUE: Hypothesis of linear dependence can be accepted with 99% of success. **p RED:** Hypothesis of linear dependence is rejected for 95% confidence.

- Kendall's correlation matrix

| $\tau_b(XX)$ | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|--------------|----------|---------------|---------------|---------------|---------------|
| d_IP | 1 | 0.9293 | 0.0481 | 0.9239 | 0.2446 |
| d_IR | | 1 | 0.0107 | 0.962 | 0.2717 |
| d_Cr | | | 1 | 0.016 | 0.2086 |
| d_RSD | | | | 1 | 0.2446 |
| d_Volum | | | | | 1 |

r BLUE: Hypothesis of linear dependence can be accepted. **r RED:** Hypothesis of linear dependence is rejected.

- Z-ratio matrix

| Z(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|--------|---------|---------|---------|
| d_IP | INF | 5.5987 | -0.2933 | 5.566 | 1.4733 |
| d_IR | | INF | -0.0652 | 5.7951 | 1.637 |
| d_Cr | | | INF | -0.0978 | -1.2711 |
| d_RSD | | | | INF | 1.4733 |
| d_Volum | | | | | INF |

- Probability of wrong model matrix

| pZ(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|----------|-----------------|---------------|-----------------|---------------|
| d_IP | 0 | 2.160e-8 | 1.2307 | 2.607e-8 | 0.1407 |
| d_IR | | 0 | 1.052 | 6.827e-9 | 0.1016 |
| d_Cr | | | 0 | 1.0779 | 1.7963 |
| d_RSD | | | | 0 | 0.1407 |
| d_Volum | | | | | 0 |

p BLUE: Hypothesis of linear dependence can be accepted with 99% of success. **p RED:** Hypothesis of linear dependence is rejected for 95% confidence.

- Kendall's correlation matrix

| $\tau_c(XX)$ | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|--------------|----------|--------------|--------------|--------------|--------------|
| d_IP | 1 | 0.855 | 0.045 | 0.85 | 0.225 |
| d_IR | | 1 | 0.01 | 0.885 | 0.25 |
| d_Cr | | | 1 | 0.015 | 0.195 |
| d_RSD | | | | 1 | 0.225 |
| d_Volum | | | | | 1 |

r BLUE: Hypothesis of linear dependence can be accepted. **r RED:** Hypothesis of linear dependence is rejected.

- Z-ratio matrix

| Z(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|--------|---------|---------|---------|
| d_IP | INF | 5.3188 | -0.2787 | 5.2877 | 1.3997 |
| d_IR | | INF | -0.0619 | 5.5054 | 1.5552 |
| d_Cr | | | INF | -0.0929 | -1.2076 |
| d_RSD | | | | INF | 1.3997 |
| d_Volum | | | | | INF |

- Probability of wrong model matrix

| pZ(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|----------|--------|----------|---------|
| d_IP | 0 | 1.045e-7 | 1.2195 | 1.239e-7 | 0.1616 |
| d_IR | | 0 | 1.0494 | 3.684e-8 | 0.1199 |
| d_Cr | | | 0 | 1.074 | 1.7728 |
| d_RSD | | | | 0 | 0.1616 |
| d_Volum | | | | | 0 |

p BLUE: Hypothesis of linear dependence can be accepted with 99% of success. **p RED:** Hypothesis of linear dependence is rejected for 95% confidence.

- Gamma correlation matrix

| $\Gamma(XX)$ | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|--------------|------|--------|--------|--------|---------|
| d_IP | 1 | 0.9553 | 0.0503 | 0.9341 | 0.2514 |
| d_IR | | 1 | 0.0112 | 0.9779 | 0.2809 |
| d_Cr | | | 1 | 0.0166 | 0.2179 |
| d_RSD | | | | 1 | 0.2486 |
| d_Volum | | | | | 1 |

r BLUE: Hypothesis of linear dependence can be accepted. **r RED:** Hypothesis of linear dependence is rejected.

- Z-ratio matrix

| Z(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|--------|--------|--------|---------|
| d_IP | INF | 5.6257 | 0.0156 | 5.3783 | 0.3896 |
| d_IR | | INF | 0.0008 | 5.895 | 0.4864 |
| d_Cr | | | INF | 0.0017 | 0.2926 |
| d_RSD | | | | INF | 0.381 |
| d_Volum | | | | | INF |

- Probability of wrong model matrix

| pZ(XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|---------|------|----------|--------|----------|---------|
| d_IP | 0 | 1.847e-8 | 0.9876 | 7.518e-8 | 0.6968 |
| d_IR | | 0 | 0.9994 | 3.748e-9 | 0.6267 |
| d_Cr | | | 0 | 0.9986 | 0.7698 |
| d_RSD | | | | 0 | 0.7032 |
| d_Volum | | | | | 0 |

p BLUE: Hypothesis of linear dependence can be accepted with 99% of success. **p RED:** Hypothesis of linear dependence is rejected for 95% confidence.

○ GLOBAL

- All correlation and probability matrices matrix

| all _{r,p} (XX) | d_IP | d_IR | d_Cr | d_RSD | d_Volum |
|-------------------------|------|------|------|-------|---------|
| d_IP | YES | YES | NO | YES | NO |
| d_IR | | YES | NO | YES | NO |
| d_Cr | | | YES | NO | NO |
| d_RSD | | | | YES | NO |
| d_Volum | | | | | YES |

Modelul designului de experiment prin șiruri ortogonale a fost testat. În continuare sunt redată șirurile ortogonale și testarea acestora folosind Microsoft Excel®:

| Tablou ortogonal | | | | | Verificare | | | |
|--------------------------------|------------------|------|------|---|------------|------|------|------|
| 4 ² ×2 ¹ | Factori (nivele) | | | | | | | |
| | A(4) | B(4) | C(2) | | | A(4) | B(4) | C(2) |
| | 1 | 0 | 2 | 0 | | 1 | | |
| | 2 | 1 | 0 | 1 | | 0 | 1 | |
| | 3 | 2 | 3 | 1 | | 0 | 0 | 1 |
| 4 | 3 | 1 | 0 | | | | | |
| 2 ³ | Factori (nivele) | | | | | | | |
| | A(2) | B(2) | C(2) | | A(2) | B(2) | C(2) | |
| | 1 | 0 | 1 | | 0 | 1 | | |
| | 2 | 1 | 1 | | 1 | 0 | 1 | |
| | 3 | 0 | 0 | | 1 | 0 | 0 | 1 |
| 4 | 1 | 0 | 0 | | | | | |
| 6 ¹ ×3 ² | Factori (nivele) | | | | | | | |
| | A(6) | B(3) | C(3) | | A(6) | B(3) | C(3) | |
| | 1 | 0 | 1 | | 0 | 1 | | |
| | 2 | 1 | 1 | | 2 | 0 | 1 | |
| | 3 | 2 | 0 | | 1 | 0 | 0 | 1 |
| | 4 | 3 | 2 | | 2 | | | |
| | 5 | 4 | 2 | | 0 | | | |
| 6 | 5 | 0 | 1 | | | | | |
| 3 ³ | Factori (nivele) | | | | | | | |
| | A(3) | B(3) | C(3) | | A(3) | B(3) | C(3) | |
| | 1 | 0 | 1 | | 1 | 1 | | |
| | 2 | 1 | 0 | | 2 | 0 | 1 | |
| | 3 | 2 | 2 | | 2 | 0 | 0 | 1 |
| | 4 | 0 | 1 | | 1 | | | |
| | 5 | 1 | 2 | | 0 | | | |
| 6 | 2 | 0 | 0 | | | | | |
| 3 ² ×2 ¹ | Factori (nivele) | | | | | | | |
| | A(3) | B(3) | C(2) | | A(3) | B(3) | C(2) | |
| | 1 | 1 | 1 | | 0 | 1 | | |
| 2 | 1 | 1 | 1 | 0 | 1 | | | |
| | | | | 0 | 0 | 1 | | |

| | | | | | | | | | | | | |
|-------------------|------------------|------|------|------|------|------|-------|------|------|------|------|------|
| | 4 | 2 | 2 | 1 | | | | | | | | |
| | 5 | 0 | 2 | 0 | | | | | | | | |
| | 6 | 0 | 0 | 1 | | | | | | | | |
| 3^5 | Factori (nivele) | | | | | | | | | | | |
| | A(3) | B(3) | C(3) | D(3) | E(3) | A(3) | B(3) | C(3) | D(3) | E(3) | | |
| | 1 | 0 | 0 | 0 | 0 | 0 | | | | | | |
| | 2 | 0 | 0 | 2 | 2 | 1 | | | | | | |
| | 3 | 0 | 2 | 0 | 2 | 2 | | | | | | |
| | 4 | 1 | 1 | 2 | 0 | 2 | | | | | | |
| | 5 | 1 | 2 | 1 | 0 | 1 | | | | | | |
| | 6 | 1 | 2 | 2 | 1 | 0 | | | | | | |
| | 7 | 2 | 0 | 1 | 1 | 2 | | | | | | |
| | 8 | 2 | 1 | 0 | 1 | 1 | | | | | | |
| 9 | 2 | 1 | 1 | 2 | 0 | | | | | | | |
| | | | | | | | | | | | | |
| 9^4 | Factor (levels) | | | | | | | | | | | |
| | A(9) | B(9) | C(9) | D(9) | A(9) | B(9) | C(9) | D(9) | | | | |
| | 1 | 0 | 0 | 7 | 5 | | | | | | | |
| | 2 | 1 | 8 | 0 | 4 | | | | | | | |
| | 3 | 2 | 1 | 1 | 1 | | | | | | | |
| | 4 | 3 | 7 | 8 | 6 | | | | | | | |
| | 5 | 4 | 6 | 6 | 0 | | | | | | | |
| | 6 | 5 | 5 | 3 | 7 | | | | | | | |
| | 7 | 6 | 2 | 2 | 8 | | | | | | | |
| | 8 | 7 | 4 | 5 | 3 | | | | | | | |
| 9 | 8 | 3 | 4 | 2 | | | | | | | | |
| | | | | | | | | | | | | |
| $9^2 \times 3^2$ | Factori (nivele) | | | | | | | | | | | |
| | A(9) | B(9) | C(3) | D(3) | A(9) | B(9) | C(3) | D(3) | | | | |
| | 1 | 0 | 1 | 1 | 0 | | | | | | | |
| | 2 | 1 | 2 | 0 | 2 | | | | | | | |
| | 3 | 2 | 4 | 2 | 1 | | | | | | | |
| | 4 | 3 | 7 | 2 | 2 | | | | | | | |
| | 5 | 4 | 8 | 1 | 1 | | | | | | | |
| | 6 | 5 | 6 | 0 | 0 | | | | | | | |
| | 7 | 6 | 5 | 1 | 0 | | | | | | | |
| | 8 | 7 | 3 | 0 | 2 | | | | | | | |
| 9 | 8 | 0 | 2 | 1 | | | | | | | | |
| | | | | | | | | | | | | |
| $9^1 \times 3^3$ | Factori (nivele) | | | | | | | | | | | |
| | A(9) | B(3) | C(3) | D(3) | A(9) | B(3) | C(3) | D(3) | | | | |
| | 1 | 0 | 1 | 1 | 1 | | | | | | | |
| | 2 | 1 | 1 | 0 | 0 | | | | | | | |
| | 3 | 2 | 0 | 2 | 2 | | | | | | | |
| | 4 | 3 | 1 | 0 | 2 | | | | | | | |
| | 5 | 4 | 2 | 2 | 1 | | | | | | | |
| | 6 | 5 | 2 | 1 | 0 | | | | | | | |
| | 7 | 6 | 0 | 2 | 0 | | | | | | | |
| | 8 | 7 | 2 | 1 | 2 | | | | | | | |
| 9 | 8 | 0 | 0 | 1 | | | | | | | | |
| | | | | | | | | | | | | |
| $10^1 \times 5^5$ | Factori (nivele) | | | | | | | | | | | |
| | A(10) | B(5) | C(5) | D(5) | E(5) | F(5) | A(10) | B(5) | C(5) | D(5) | E(5) | F(5) |
| | 1 | 0 | 0 | 0 | 0 | 2 | 2 | B(5) | 0 | 1 | | |

Obiectivul 1.1. Documentare la zi asupra metodelor de studiu relații structură-activitate

Obiectivul a fost atins prin intermediul activităților 1-4 expuse anterior:

- ÷ Identificare și colectare metode SAR
- ÷ Dobândire competențe complementare prin participări la manifestări științifice/stagii de documentare-cercetare
- ÷ Sumarizare și stocare metode SAR
- ÷ Achiziție, instalare, testare și configurare aparatură suport

Obiectivul 1.2. *Elaborare model matematic*

Obiectivul a fost atins prin intermediul activităților 5-7 expuse anterior:

- ÷ Dobândire competențe complementare prin participări la manifestări științifice/stagii de documentare-cercetare
- ÷ Elaborare specificații și metodologie model matematic
- ÷ Testare și validare model

Etapa 1. Rezultate livrate pe etapă

÷ Site web proiect disponibil online: <http://sorana.academicdirect.ro/grants/ID0458/>

÷ Lucrare CNCSIS Tip "B+":

- Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, Amino Acids Sequences Analysis on Collagen, Bulletin of University of Agricultural Sciences and Veterinary Medicine - Animal Sciences and Biotechnologies, AcademicPres, ISSN 1843-5262, eISSN 1843-536X, Cluj-Napoca, Romania, 63-64, p. 311-316, 2007.
- Revista "Bulletin of University of Agricultural Sciences and Veterinary Medicine - Animal Sciences and Biotechnologies" este clasificată de CNCSIS în categoria "B+":

http://www.cnscis.ro/cenaposs/2006/rev_ed/reviste_cat_B+.pdf

÷ Lucrare indexată în baze de date internaționale:

- Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, Structure versus Biological Role of Substituted Thiadiazole- and Thiadiazoline- Disulfonamides, Studii si Cercetari Stiintifice Universitatea Bacau Seria Biologie, Universitatea Bacau, ISSN 1224-919X, Bacau, Romania, 12(1), p. 50-56, 2007.
- Revista "Studii si Cercetari Stiintifice Universitatea Bacau Seria Biologie" este cuprinsă în "Thomson Scientific Master Journal List":

<http://www.thomsonscientific.com/cgi-bin/jrnlst/jlresults.cgi?PC=MASTER&ISSN=1224-919X>

- Revista "Studii si Cercetari Stiintifice Universitatea Bacau Seria Biologie" este indexată de produsul Thompson Scientific "Zoological Record":

<http://www.thomsonscientific.com/cgi-bin/jrnlst/jlresults.cgi?PC=B7&ISSN=1224-919X>

÷ Lucrare Thompson ISI SCI Expanded:

- Lorentz JÄNTSCHI, Sorana D. BOLBOACĂ, Mircea V. DIUDEA, Chromatographic Retention Times of Polychlorinated Biphenyls: from Structural Information to Property Characterization, International Journal of Molecular Sciences, Molecular Diversity Preservation International, ISSN 1422-0067, Basel & www, Switzerland & Internet, 8(11), p. 1125-1157, 2007.
- Revista "International Journal of Molecular Sciences" este cuprinsă în "Thomson Scientific Science Citation Index Expanded":

<http://www.thomsonscientific.com/cgi-bin/jrnlst/jlresults.cgi?PC=D&ISSN=1422-0067>

- Revista "International Journal of Molecular Sciences" are factorii de impact calculați după cum urmează: $IF^{2005} = 1.467$; $IF^{2006} = 0.679$.