

Eficacitatea antioxidantă a derivaților de 3-indol prin integrarea informațiilor complexe

Sorana BOLBOACĂ*, Claudia FILIP**, Ștefan ȚIGAN*, Lorentz JÄNTSCHI**

*Universitatea de Medicină și Farmacie „Iuliu Hațieganu”, Cluj-Napoca, Romania

**Universitatea Tehnică Cluj-Napoca, Romania

Rezumat

Relația dintre structura și activitatea antioxidantă a opt compuși conținând 3-indol a fost obținută prin aplicarea unei metodologii originale, prin integrarea informațiilor complexe a structurii compușilor, folosind Familia Descriptorilor Moleculari. Modelele obținute au fost evaluate și validate prin studiul coeficienților de corelație, a coeficienților de corelație încrucișată, analiza stabilității modelului și compararea acestuia cu cel mai performant model raportat în literatura de specialitate prin aplicarea analizei de corelare a corelațiilor. Activitatea antioxidantă a derivaților de 3-indol studiați este deopotrivă de natură geometrică și topologică și depinde de numărul de atomi de hidrogen, electronegativitatea de grup, potențialul proprietății și inversul acestuia. Aplicarea metodologiei SAR în modelarea activității antioxidante a derivaților de 3-indol oferă modele exacte care permit înțelegerea relației structură-activitate, un pas important în descoperirea și caracterizarea de compuși noi.

Cuvinte cheie: Relații structură-activitate (SAR); Derivați ai 3-indol; Integrarea informațiilor complexe; Familia Descriptorilor Moleculari;

Introducere

Beneficiile și efectele adverse ale compușilor ce conțin 3-indoli ca și agenți chemoprotectivi, cuantificate prin relația dintre activitatea antioxidantă și carcinogeneza sunt încă contradictorii [1]. Două mecanisme sunt recunoscute astăzi ca fiind implicate în activitatea protectivă a derivaților de 3-indol împotriva distrugerii celulare, respectiv inhibarea [2] sau stimularea [3] activităților unor enzime.

Eficacitatea antioxidantă a opt compuși ce conțin 3-indol a fost studiată anterior de Shertzer și colaboratorii [4] pentru a stabili relația dintre potențialul antioxidant și parametrii electronici, polari și sterici (energia de disociere a legăturilor, lungimea legăturilor, momentele dipol, densitatea de sarcină electronică, și greutatea molară). Modelele obținute de Shertzer, tipul acestora (mono-, bi-, sau tetra-variat) și coeficienții de determinare asociați sunt în tabelul 1. În tabelul 1 s-a notat cu IC_{50} concentrația compusului necesară pentru a inhiba peroxidaza lipidică cu 50%, ΔH_f variația entalpiei de formare, Z_{dl} distanța maximă a celei mai mici căsuțe periodice care poate înconjura molecula, Z_{dipol} momentul dipol, $\log P$ logaritmul în baza zece a coeficientului de partiție, iar cu MW greutatea moleculară. Activitatea măsurată (IC_{50}) este o mărime semicantitativă obținută prin determinări pe eșantioane cu distribuții binomiale. Evaluarea corectă a intervalului de încredere [5] permite aprecierea corectă a valorilor măsurate.

Tabel 1. Modele de estimare a eficacității antioxidante a 3-indolilor raportate de Shertzer

Nr.	Model	Tip model	r^2
1	$-\log IC_{50} = 3.02 \cdot \Delta H_f + 1.36$	mono-variat	0.78
2	$-\log IC_{50} = 0.21 \cdot Z_{dl} + 2.44$	mono-variat	0.71
3	$-\log IC_{50} = 2.09 \cdot \Delta H_f + 0.08 \cdot Z_{dl} + 1.54$	bi-variat	0.81
4	$-\log IC_{50} = 4.85 \cdot \Delta H_f - 0.49 \cdot Z_{dipol} + 0.04 \cdot \log P + 0.004 \cdot MW - 1.48$	tetra-variat	0.97

Pornind de la informațiile complexe oferite de structura moleculară a opt derivați de 3-indol, scopul cercetării a fost de a studia activitatea antioxidantă prin aplicarea unei metodologii originale și de a evalua abilitățile modelelor SAR obținute în predicția activității antioxidante.

Material și Metodă

Material

Un eșantion de opt compuși conținând 3-indol au fost incluși în studiu. Denumirea derivaților de 3-indol, abrevierea și activitatea antioxidantă măsurată ($\log IC_{50}$) a acestora sunt în tabelul 2.

Tabelul 2. Denumirea derivaților de 3-indol și activitatea antioxidantă măsurată

Abreviere	Compus	$\log IC_{50}$
com_01	3-(4-N,N dimetil-aminobenzil) indol	- 5.82
com_02	3-(2,4,6-trimetil-benzil) indol	- 4.96
com_03	3-(4-hidroxy-benzil) indol	- 4.92
com_04	3-(4-metoxi-benzil) indol	- 4.62
com_05	3-hidroxy-etil-indol	- 4.54
com_06	3-benzil-indol	- 4.44
com_07	3-metil-indol	- 4.00
com_08	3-hydroximetil-indol	- 3.80

Integrarea informațiilor complexe

Eficacitatea antioxidantă a celor opt derivați de 3-indol a fost modelată prin integrarea informațiilor complexe oferite de structura moleculară a compușilor. Modelarea eficacității antioxidante s-a realizat prin aplicarea unei metodologii proprii prin folosirea Familiei Descriptorilor Moleculari.

Metodologia aplicată pentru modelarea relației structură-activitate (SAR) [6] a cuprins următoarele etape:

- Etapa I: **Reprezentarea tridimensională** a structurilor moleculelor prin folosirea programului HyperChem [7];
- Etapa II: **Crearea** fișierului care conține activitatea antioxidantă măsurată a derivaților 3-indol;
- Etapa III: **Generarea, calcularea și filtrarea** membrilor familiei de descriptori moleculari [8,9] pentru derivații de 3-indol. Cei opt derivați au fost folosiți pentru a genera familia de descriptori moleculari corespunzătoare setului. Algoritmul de generare a listei descriptorilor se bazează strict pe structura compușilor. Procesul de generare a descriptorilor a fost urmat de un proces de filtrare care a șters din baza de date descriptorii moleculari identici prin impunerea unui selector de semnificație egal cu 10^{-9} pentru îndepărtarea informației redundante. Denumirea fiecărui descriptor molecular obținut include șapte litere care exprimă modul său de calcul. Astfel, litera a șaptea descrie caracteristica care consideră cunoștințele relativ la geometria și topologia moleculei, litera a șasea caracterizează proprietatea atomică (masa, sarcina electrică parțială, cardinalitatea, electronegativitatea, electronegativitatea de grup, numărul de atomi de hidrogen legați direct), iar restul literelor descriu: descriptorul de interacțiune (litera a cincea), modelul de suprapunere a interacțiunii descriptorilor (litera a patra), metoda de fragmentare moleculară (litera a treia), metoda de cumulare a proprietăților de fragmentare (litera a doua) și procedura de linearizare aplicată în generarea descriptorului global molecular (prima literă).
- Etapa IV: **Căutarea și identificarea** celor mai semnificative modele SAR. S-a impus ca și criteriu de căutare obținerea unui coeficient de corelație cât mai apropiate de valoarea 1.
- Etapa V: **Validarea** modelelor găsite s-a realizează prin calcularea coeficientului de validare încrucișată (r^2_{cv}) [10]. Acesta s-a calculat prin excluderea a câte unui compus din eșantion, recalcularea coeficienților ecuației de estimare a activității antioxidante și estimarea activității antioxidante a compusului exclus pe baza ecuație de regresie obținută.
- Etapa VI: **Analiza** modelelor SAR selectate s-a realizează prin: analiza stabilității modelului (diferența dintre coeficientul de determinare al modelului SAR și coeficientul de determinare obținut prin validarea modelului – cu cât diferența este mai mică cu atât modelul este mai stabil); și compararea modelului SAR cu cel mai performant model raportat anterior prin aplicarea analizei de corelare a coeficienților de corelație (testul Steiger [11]).

Rezultate

În urma integrării complexe a cunoștințelor, au fost identificate și înregistrate cele mai performante modele SAR, unul mono-variat și unul bi-variat. Cele două modele SAR sunt:

- Mono-variat: $\hat{Y} = -3.76 - 1.34 \cdot 10^{-5} \cdot asMmtQg$
- Bi-variat: $\hat{Y} = 7.18 - 1.1 \cdot lbPMkHg - 33.24 \cdot iAPrVGt$

Statisticile asociate celor două modele sunt redată în tabelul 3.

Tabelul 3. Statistici asociate modelelor SAR

Caracteristica	Notația	Mono-variat	Bi-variat
Numărul de descriptorii folosiți în model	n	1	2
Coeficientul de corelație	r	0.9508	0.9999
Coeficientul de determinare	r ²	0.9040	0.9998
Coeficientul de determinare ajustat	r ² _{adj}	0.8880	0.9997
Eroarea standard	S _{est}	0.2093	0.0104
Parametrul Fisher	F _{est}	56.49	12651
Probabilitatea unui model de regresie greșit	p _{est} (%)	2.9 · 10 ⁻²	5.5 · 10 ⁻⁸
Coeficientul de validare încrucișată	r ² _{cv(100)}	0.8316	0.9994
Parametrul Fisher asociat analizei de validare încrucișată	F _{pred}	29.32	3877
Probabilitatea unui model de validare încrucișată greșit	p _{pred} (%)	1.6 · 10 ⁻¹	1.05 · 10 ⁻⁶
Eroarea standard a analizei de validare încrucișată	S ₁₀₀	0.2784	0.0188
Diferența dintre coeficientul de determinare și coeficientul de validare încrucișată	r ² - r ² _{cv(100)}	0.0724	0.0004

Denumirea descriptorilor moleculari folosiți în modele, valorile acestora și activitatea antioxidantă estimată (\hat{Y}) sunt în tabelul 4.

Tabelul 4. Denumirea și valorile descriptorilor moleculari și activitatea antioxidantă estimată cu ajutorul acestora (cele mai bune modele obținute)

Compus	Mono-variat		Bi-variat		
	asMmtQg	\hat{Y}_{mono-v}	iAPrVGt	lbPMkHg	\hat{Y}_{bi-v}
com_01	1.43 · 10 ⁵	-5.67	2.84 · 10 ⁻¹	3.23 · 10 ⁰	-5.81
com_02	8.85 · 10 ⁴	-4.94	2.54 · 10 ⁻¹	3.36 · 10 ⁰	-4.96
com_03	8.26 · 10 ⁴	-4.86	2.61 · 10 ⁻¹	3.12 · 10 ⁰	-4.92
com_04	4.64 · 10 ⁴	-4.38	2.72 · 10 ⁻¹	2.52 · 10 ⁰	-4.63
com_05	8.86 · 10 ⁴	-4.94	2.47 · 10 ⁻¹	3.22 · 10 ⁰	-4.55
com_06	5.65 · 10 ⁴	-4.51	2.50 · 10 ⁻¹	3.02 · 10 ⁰	-4.43
com_07	1.30 · 10 ⁴	-3.93	2.37 · 10 ⁻¹	3.00 · 10 ⁰	-3.99
com_08	8.14 · 10 ³	-3.86	2.39 · 10 ⁻¹	2.77 · 10 ⁰	-3.80

Analiza corelației dintre fiecare descriptor molecular și eficacitatea antioxidantă măsurată este redată în tabelul 5. Sunt incluse coeficientul de determinare (r²(desc, IC₅₀)), valorile coeficienților drepte de regresie, eroarea standard a coeficienților (ErStd), parametrul Student (t), probabilitatea asociată testului Student (p_t) și marginea inferioară (CI_I 95%) și superioară (CI_S 95%) a intervalului de încredere de 95%.

Tabelul 5. Statistici de regresie asociate modelului bi-variat

	r ² (descr, IC ₅₀)	Coeficienți	ErStd	t	p _t (%)	CI _I 95%	CI _S 95%
Intercept	-	7.18	0.07	96.05	2.32 · 10 ⁻⁷	6.99	7.37
lbPMkHg	0.2479	-1.10	0.01	-76.45	7.26 · 10 ⁻⁷	-1.13	-1.06
iAPrVGt	0.7689	-33.24	0.24	-137.95	3.80 · 10 ⁻⁸	-33.86	-32.62

Reprezentarea grafică a dependenței eficacității antioxidante a derivaților de 3-indol de structura acestora exprimată prin modelul bi-variat este în figura 1.

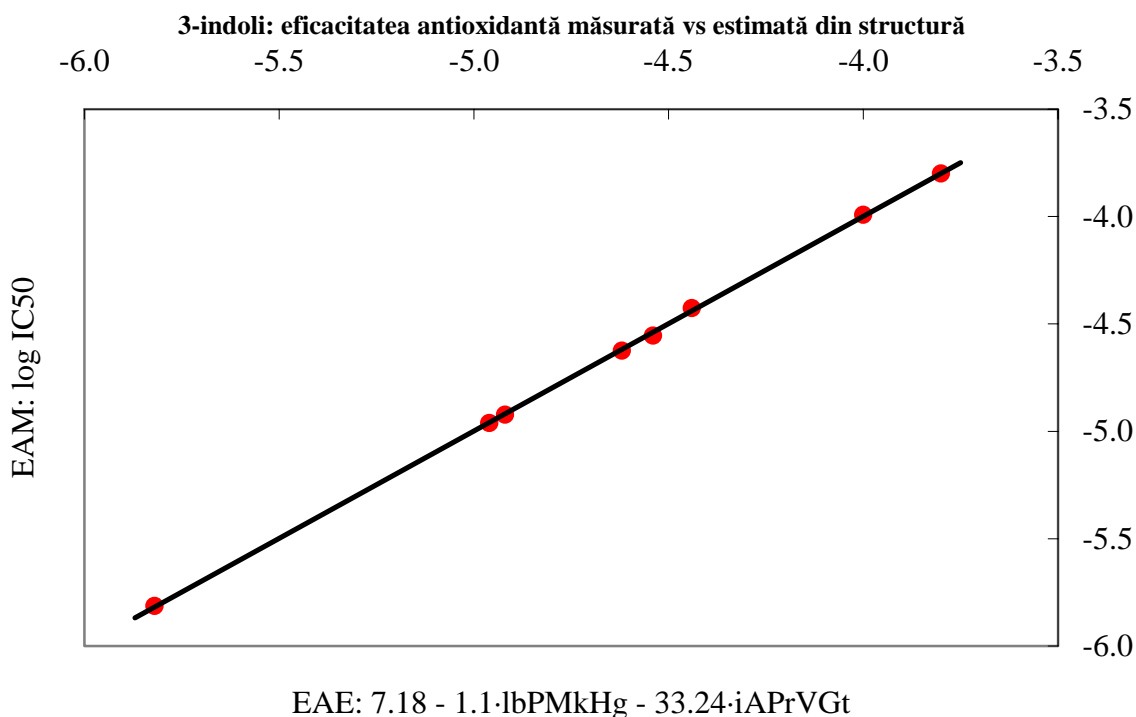


Figura 1. Eficacitatea antioxidantă măsurată (EAM) vs estimată (EAE) cu modelul bi-variat

Rezultatele comparării modelelor SAR obținute prin integrarea informațiilor complexe ale structurii compușilor cu cel mai performant model raportat anterior în literatura de specialitate (modelul în patru variabile, tabelul 1) sunt în tabelul 6.

Tabelul 6. Rezultatele comparării modelelor SAR obținute cu cel mai performant model raportat

Testul Steiger	Model SAR	
Caracteristica	Mono-variat	Bi-variat
Număr de descriptori	1	2
$r(\text{IC}_{50}, \hat{Y})$	0.9508	0.9999
$r(\text{IC}_{50}, \hat{Y}_{\text{Model anterior}})$	0.9810	0.9810
$R(\hat{Y}, \hat{Y}_{\text{Model anterior}})$	0.9131	0.9788
Z (parametrul lui Steiger)	-1.0726	5.8628
p_Z (probabilitatea de coincidență, %)	85.82	$2.28 \cdot 10^{-7}$

Discuții

Integrarea informațiilor oferite de structura celor opt congeneri 3-indoli permite estimarea activității antioxidante pe baza analizei structurii compușilor. Modelele obținute pentru exprimarea relației dintre structura și activitatea antioxidantă a celor opt compuși 3-indol conțin trei descriptori moleculari. Doi descriptori consideră geometria moleculei (asMmtQg, lbPMkHg) și unul topologia acesteia (iAPrVGt). Un descriptor ia în considerare sarcina electrică parțială (asMmtQg), alt descriptor (lbPMkHg) numărul de atomi de hidrogen legați direct, iar altul (iAPrVGt) electronegativitatea de grup. În ceea ce privește descriptorul de interacțiune, descriptorul asMmtQg ia în considerare forța nucleară tare, iAPrVGt potențialul proprietății iar lbPMkHg inversul potențialului proprietății.

Modelul SAR mono-variat este semnificativ statistic având o probabilitate de model greșit egală cu $2.9 \cdot 10^{-2}$ %. Inspectând modelul SAR mono-variat se poate afirma că activitatea antioxidantă a derivaților de 3-indol este de natură geometrică, și este dependentă de sarcina parțială pe atomi și de interacțiunile date de forța nucleară tare între atomi.

Coeficientul de corelație al modelului SAR bi-variat ($r = 0.9999$) vine să susțină rolul descriptorilor moleculari în estimarea activității antioxidante a derivaților de 3-indol. Peste 99% din

variația activității antioxidante a derivaților de 3-indol poate fi explicată de relația lineară dintre aceasta și descriptorii moleculari lbPMkHg și iAPrVGt. Modelul bi-variat este semnificativ statistic, probabilitatea de model greșit fiind egală cu $5.5 \cdot 10^{-8}$ %. Rezultatele regresiei modelului bi-variat arată faptul că coeficienții dreptei de regresie sunt semnificativ statistic diferiți de zero ($p < 0.001$), semnificație susținută și de valorile intervalelor de încredere asociate fiecărui coeficient (tabelul 5). Coeficientul de determinare dintre descriptorii moleculari ai modelului bi-variat și valoarea măsurată a activității antioxidante, arată că nu există corelație semnificativă statistic între descriptorul lbPMkHg și activitatea măsurată ($r = 0.49$) dar există o corelație semnificativă statistic între iAPrVGt și activitatea măsurată ($r = 0.87$). Validitatea modelului este susținută de probabilitatea extrem de mică de model greșit obținută în urma analizei încrucișate ($p_{\text{pred}}(\%) = 1.05 \cdot 10^{-6}$) și de valoarea coeficientului de validare încrucișată ($r^2_{\text{cv}(100)} = 0.9994$). Stabilitatea modelului SAR bi-variat este susținută de valoarea diferenței dintre coeficientul de determinare al modelului și coeficientul de validare încrucișată ($r^2 - r^2_{\text{cv}(100)} = 0.0004$). Modelul bi-variat aduce ca și informație suplimentară față de modelul mono-variat că activitatea antioxidante a derivaților de 3-indol este deopotrivă de natură geometrică și topologică și depinde de numărul de atomi de hidrogen legați direct, electronegativitatea de grup și de potențialul proprietății și a inversului proprietății.

Compararea coeficienților de corelație obținuți cu cel mai performant model raportat în literatura de specialitate (modelul în patru variabile, tabelul 1, $r^2 = 0.97$) arată că coeficientul de corelație obținut de modelul SAR mono-variat (model semnificativ statistic, tabelul 3, $r^2 = 0.9$), nu diferă semnificativ statistic de coeficientul de corelație obținut de modelul în patru variabile raportat anterior (probabilitatea de coincidență cu acesta $p_z = 82.5\%$, tabelul 6); este de notat aici reducerea ordinului de varianță de la patru la unu. Valoarea coeficientului de corelație obținut de modelul SAR bi-variat este semnificativ statistic mai mare în comparație cu cel obținut de modelul în patru variabile raportat anterior (probabilitatea de coincidență cu acesta $p_z = 2.28 \cdot 10^{-7}$ %, tabelul 6). Se pot obține astfel rezultate semnificativ statistic mai bune în estimarea activității antioxidante a derivaților de 3-indol folosind perechi de doi descriptori de structură, și mai mult, s-a crescut stabilitatea modelului prin reducerea numărului de variabile necesare pentru estimarea activității de interes (de la patru variabile la două variabile).

Plecând de la rezultatele obținute de prezentul studiu se poate prezice activitatea antioxidante a altor congeneri ai 3-indol folosind resursa online [12], care necesită ca etapă prealabilă reprezentarea tridimensională a derivaților de 3-indol de interes sub formă de fișiere *.hin. Astfel, se pot obține informații cu privire la activitatea antioxidante a unui nou compus plecând de la structura acestuia, fără a mai realiza nici un experiment suplimentar. Se deschide astfel calea spre identificarea de noi compuși, prin experimente virtuale, mai ieftine, mai puțin consumatoare de timp, și care nu implică experimente pe animale de laborator sau celule.

Concluzii

Modelarea activității antioxidante a derivaților de 3-indol prin integrarea informațiilor complexe aduce informații utile în înțelegerea relației dintre structura compușilor și activitatea antioxidante a acestora, informații care pot fi folosite în sinteza de noi compuși biologic activi.

Activitatea antioxidante a compușilor 3-indol este de natură geometrică și topologică și depinde de numărul de atomi de hidrogen legați direct, electronegativitatea de grup și de potențialul proprietății și a inversului proprietății.

Aplicarea metodologiei SAR în modelarea activității antioxidante a derivaților de 3-indol oferă modele exacte care permit reducerea numărului de variabile necesare pentru înțelegerea relației-structură activitate, un pas important în descoperirea și caracterizarea de compuși noi.

Notă

Cercetarea a fost suportată parțial de către Ministerul Educației și Cercetării prin proiectul ET36/UEFISCSU/2005.

Referințe

- [1] LEE B.M., PARK K.K.: Beneficial and adverse effects of chemopreventive agents, *Mutat Res-Fund Mol M*, 2001, 523-524, 265-78.
- [2] VETRANO A.M., HECK D.E., MARIANO T.M., MISHIN V., LASKIN D.L., LASKIN J.D.: Characterization of the oxidase activity in mammalian catalase, *J Biol Chem*, 2005, 280(42), 35372-81.
- [3] FONG A.T., SWANSON H.I., DASHWOOD R.H., WILLIAMS D.E., HENDRICKS J.D., BAILEY G.S.: Mechanisms of anti-carcinogenesis by indole-3-carbinol. Studies of enzyme induction, electrophile-scavenging, and inhibition of aflatoxin B1 activation, *Biochem Pharmacol*, 1990, 39(1), 19-26.
- [4] SHERTZER G.H., TABOR M.W., HOGAN I.T.D., BROWN J.S., SAINSBURY M.: Molecular modeling parameters predict antioxidant efficacy of 3-indolyl compounds, *Arch Toxicol*, 1996, 70, 830-4.
- [5] BOLBOACĂ S., JÄNTSCHI L.: Binomial Distribution Sample Confidence Intervals Estimation for Positive and Negative Likelihood Ratio Medical Key Parameters, *American Informatics Medical Association Annual Symposium on Biomedical and Health Informatics from Foundations to Applications to Policy*, Washington D.C., U.S.A., October 22-26, Conference Proceedings, 2005, 66-70.
- [6] JÄNTSCHI L.: Molecular Descriptors Family on Structure Activity Relationships 1. Review of the Methodology, *Leonardo Electronic Journal of Practices and Technologies*, 2005, 6, 76-98.
- [7] ***, HyperChem , Molecular Modelling System [Internet page]; ©2003, Hypercube, Inc. [cited 2005 Nov]. Available from: URL: <http://hyper.com/products/>
- [8] DIUDEA M., GUTMAN I., JÄNTSCHI L.: Molecular Topology, Nova Science, Huntington, New York 2001, 332.
- [9] JÄNTSCHI L., KATONA G., DIUDEA M.: Modeling Molecular Properties by Cluj Indices, *Commun Math Comput Chem (MATCH)*, Bayreuth, Germany, 2000, 41, 151-88.
- [10] ***, Leave-one-out Analysis. ©2005, Virtual Library of Free Software [cited 2005 Nov]. Available from: URL: http://vl.academicdirect.org/molecular_topology/mdf_findings/loo/
- [11] STEIGER J.H.: Tests for comparing elements of a correlation matrix, *Psychol Bull*, 1980, 87, 245-51.
- [12] ***, MDF SAR Predictor. ©2005, Virtual Library of Free Software [cited 2005 Nov]. Available from: URL: http://vl.academicdirect.org/molecular_topology/mdf_findings/sar/

Antioxidant Efficacy of 3-Indolyl Derivates by Complex Information Integration

Sorana BOLBOACĂ*, Claudia FILIP**, Ștefan ȚIGAN*, Lorentz JÄNTSCHI**

*"Iuliu Hațieganu" University of Medicine and Pharmacy, Cluj-Napoca, Romania

**Technical University of Cluj-Napoca, Romania, Cluj-Napoca, Romania

Abstract

The relationship between structure and antioxidant activity of eight 3-indolyl containing compounds was obtained by applying an original methodology, through complex information integration of compounds structure, by the use of Molecular Descriptors Family. The obtained models were evaluated and validated by means of correlation coefficients, cross-validation correlation coefficients, analysis of model stability, and comparison of the model with the best performing previous reported model by correlated correlation analysis. Antioxidant activity of 3-indolyl containing compounds is alike geometrical and topological and depend on the directly bonded hydrogen's, the group electronegativity, the property potential and its inverse. Application of SAR methodology in modeling of antioxidant activity of 3-indolyl containing compounds offer exact models that allow understanding of the structure-activity relationship, an important step in discovering and characterization of new 3-indolyl compounds.

Keywords: Structure-Activity Relationships (SAR); 3-Indolyl Containing Compounds; Complex Information Integration; Molecular Descriptors Family

Pentru corespondență:

Sorana Daniela Bolboacă

Str. Louis Pasteur nr. 6, 400349 Cluj-Napoca, Cluj, Romania

Tel: +40 264 431697

e-mail: sbolboaca@umfcluj.ro