



MOLECULAR DESCRIPTORS FAMILY ON CHROMATOGRAPHY

Lorentz JÄNTSCHI ^a, Sorana Daniela BOLBOACĂ ^b

^a Technical University of Cluj-Napoca, 400020 Cluj-Napoca, Romania, lori@academicdirect.org & ^b "Iuliu Hațeganu" Medicine and Pharmacy University of Cluj-Napoca, 400023, Cluj-Napoca, Romania, sbolboaca@umfcluj.ro

Molecular Descriptors Family is a family of structure-based molecular descriptors. It implies a molecular design on molecules which are subject of investigation. Are applied on molecules sets on which a given property are known. It provides a structure-property model for given set of molecules. The model contains a number of seven parameters for every involved descriptor. These parameters allow: ► Structure-property analysis of the given property; ► Validation of experimental measurements in relationship with compounds properties; ► Prediction of property of interest for not included compounds which are structural and/or property related with given set; ► Further developments will allow virtual synthesis of new compounds (by including a combinatorial algorithm which will generate structures using a library)

Address http://l.Chemistry/SARs/MDF_SARs/j_mdf_demo.php

Up MDF Demo Calculator -> You must select one option from every list.

Molecule filename: 01_mr1001.hin

Distance operator: Topological distance, t Geometrical distance, g

Atomic property:

- Cardinality, C
- Count of directly bounded hydrogen's, H
- Relative atomic mass, M
- Atomic electronegativity, E
- Group electronegativity, G
- Partial charge, Q

Descriptor (of interaction) formula:

- Distance, 'D' = d
- Inverted distance, 'd' = 1/d
- First atom's property, 'P' = p1
- Inverted O, 'o' = 1/p1
- Product of atomic properties, 'P' = p1p2
- Inverted P, 'p' = 1/p1p2
- Squared P, 'Q' = p1p2^2/2
- Inverted Q, 'q' = 1/p1p2^2/2
- First atom's Property multiplied by distance, 'J' = p1d
- Inverted J, 'j' = 1/p1d
- Product of atomic properties and distance, 'K' = p1p2d
- Inverted K, 'k' = 1/p1p2d
- Product of distance and squared atomic properties, 'L' = d(p1p2)^1/2
- Inverted L, 'l' = 1/p1p2d
- First atom's property potential, 'V' = p1/d
- First atom's property field, 'E' = p1/d^2
- First atom's property work, 'W' = p1^2/d
- Properties work, 'w' = p1p2/d
- First atom's property force, 'F' = p1^2/d^2
- Properties force, 'f' = p1p2/d^2
- First atom's property weak nuclear force, 'S' = p1^2/d^3
- Properties weak nuclear force, 's' = p1p2/d^3
- First atom's property strong nuclear force, 'T' = p1^2/d^4
- Properties strong nuclear force, 't' = p1p2/d^4

Molecular overall superposing formula:

- Cond., smallest, m
- Cond., highest M
- Cond., smallest absolute, n
- Cond., highest absolute, N
- Avg., sum, S
- Avg., average, A
- Avg., S/count(fragments), a
- Avg., Avg.(Avg./atom)/count(atoms), B
- Avg., S/count(bonds), b
- Geom., product P
- Geom., mean, G
- Geom., P^1/count(fragments), g
- Geom., Geom.(Geom./atom)/count(atoms), F
- Geom., P^1/count(bonds), f
- Harm., sum, s
- Harm., mean, H
- Harm., s/count(fragments), h
- Harm., Harm.(Harm./atom)/count(atoms), I
- Harm., s/count(bonds), i

Fragmentation criteria:

- Minimal fragments, m
- Maximal fragments, M
- Szeged distance based fragments, D
- Cluj path based fragments, P

Interaction model:

- Rare model and resultant relative to fragment's head, R
- Rare model and resultant relative to conventional origin, r
- Medium model and resultant relative to fragment's head, M
- Medium model and resultant relative to conventional origin, m
- Dense model and resultant relative to fragment's head, D
- Dense model and resultant relative to conventional origin, d

Linearization operator:

- Identity (no change), I
- Inversed I, i
- Absolute I, A
- Inversed A, a
- Logarithm of A, L
- Logarithm of I, l

Submit Query

Training Test L(Training) L(Test)

$r^2_{\text{training}} = 0.9971$ $r^2_{\text{test}} = 0.9974$

RR: estimated by Eq5

Results:

Following results were obtained using MDF Methodology on estimation of chromatographic parameters:

Set	Compounds	Number	Property	Equation	r^2	Ref
IChr_10	organophosphorus herbicides	10	retention index	$I_{\text{CHR}} = -3.4 + 0.32 \cdot IBPdqHg$	0.94	[7]
PCB_rrf	polychlorinated biphenyls	209	relative response factor	$R_{\text{RF}} = 6.417 \cdot iMrfFHt + 2.3 \cdot iHddFHg + 1.83 \cdot iMMmjQg - 2.510^3 \cdot iAMrVQg$	0.74	[6]
PCB_rrt	polychlorinated biphenyls	209	relative retention time	$R_{\text{Rt}} = -0.17 + 0.08745 \cdot iDRwHg$	0.98	[8]

Following results were obtained using MDF Methodology on estimation of chromatographic-like parameters:

Set	Compounds	Number	Property	Equation	r^2	Ref
MR_10	cyclic organophosphorus compounds	10	molar refraction	$MR = 17 + 28 \cdot iGdmSMt - 84 \cdot iAmrfEt$	0.94	[9]
PCB_lkow	polychlorinated biphenyls	209	Octanol-Water Partition Coefficient	$\log K_{\text{ow}} = 3 - 0.4 \cdot iDDKGg + 0.04 \cdot iHDKEG + 0.07 \cdot aHMmjQt - 37.5 \cdot aSMMjQg$	0.74	[4]
PCB_rrt	para-substituted phenols	30	idem	$\log K_{\text{ow}} = 1 + 3.4 \cdot iSDkGg - 0.4 \cdot iMMrKQg$	0.95	[2]

MDF Methodology: ► Constructing of 3D model of the given set of molecules (structure are known); we use HyperChem software to do this. ► Optimizing of the molecular geometry if is necessary (for a "in vivo" model we can use a periodic box of water molecules surrounding the molecule subject to geometry optimization). ► Generating of Molecular Descriptors Family (a family with 131328 relatives). ► Biasing of MDF and simple linear regression with measured property (this task apply 6 types of linearization operators on MDF members rising up to 787968 their number and in same time reduces it's number through bias procedure to a number of about 100000). ► If simple linear regression between measured property and MDF members does not produce a satisfactory result, then procedure it continue with multiple linear regression (with two, three, or more MDF members as independent variables)

MDF Assessment: ► A series of assessing procedures were developed and are available (online). ► Leave-One-Out Analysis let each compound out of the set and reconstruct MDF-SPR model without it, and using this new model their property are predicted; using all predicted properties a new coefficient, called leave-one-out squared correlation coefficient. ► Training vs. Test experiment split the data set in two sets, training one and test one; use the training set for obtaining MDF-SPR model and apply the model to the test set (size of the training and test set can be choused from the interface; compounds are randomly selected into the training and test sets). ► Correlated correlations analysis apply the Steiger's Z test between two or mode models in order to see if the models express same thing (being significantly correlated or not).

MDF Features: ► Our investigated sets are available online. ► Our MDF-SPR models are available online. ► MDF Investigator application allow to select a MDF-SPR model from the database, to submit a compound (a HyperChem file containing compound 3D structure) and to predict the property based on the selected model.

Selected references:

- [1] Sorana D. BOLBOACĂ, 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.*
- [2] Lorentz JÄNTSCHI, Sorana-Daniela BOLBOACĂ, 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.*
- [3] Lorentz JÄNTSCHI, Sorana BOLBOACĂ, 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.*
- [4] Lorentz JÄNTSCHI, Sorana BOLBOACĂ, 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.*
- [5] 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.*
- [6] Lorentz JÄNTSCHI, QSPR on Estimating of Polychlorinated Biphenyls Relative Response Factor using Molecular Descriptors Family, *Leonardo Electronic Journal of Practices and Technologies, AcademicDirect, ISSN 1583-1078, www, Internet, 3(5), p. 67-84, 2004.*
- [7] Lorentz JÄNTSCHI, MDF - A New QSAR/QSPR Molecular Descriptors Family, *Leonardo Journal of Sciences, AcademicDirect, ISSN 1583-0233, www, Internet, 3(4), p. 68-85, 2004.*
- [8] Lorentz JÄNTSCHI, Sorana D. BOLBOACĂ, Mircea V. DIUDEA, Relative Retention Times of Polychlorinated Biphenyls: from Structural Information to Property Characterization, *International Journal of Molecular Sciences, Molecular Diversity Preservation International, ISSN 1422-0067, www & Basel, Internet & Switzerland, submitted, 2007.*
- [9] Lorentz JÄNTSCHI, Sorana Daniela BOLBOACĂ, 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.*

Address http://l.academicdirect.org/Chemistry/SARs/MDF_SARs/

Up MDF

- DC Predictor (DC: demo calculator)
- SARs (SAR: structure-activity relationship)
- LOO Analysis (LOO: leave one out)
- Investigator
- TvT Experiment (TvT: training vs. test)
- BorQ SARs by sets (BorQ: browse or query)
- Statistics
- ADB (articles database) - still in work

Powered by UNIX

February 2007 Lorentz JÄNTSCHI