Available online at www.scholarsresearchlibrary.com

Scholars Research Library

Der Pharmacia Lettre, 2013, 5 (6):88-98
(http://scholarsresearchlibrary.com/archive.html)

Scholars Research Library
ISSN 0975-5071
USA CODEN: DPLEB4

Prediction of anti-inflammatory activity of anthranylic acids using Structural Molecular Fragment and topochemical models

Cyrille Menye[1*], Claude M. Ngabireng[2], Simeon F. Kouam[3]

[1*]*Faculty of sciences/Department of Physics, University of Yaounde I, Yaounde, P.O.Box 812, Cameroon*
[2] *National Advanced School of Engineering/Department of Mathematics and Physical Science University of Yaounde I, Yaounde, P.O.Box 8390, Cameroon*
[3] *Higher Teachers' Training College/Department of Chemistry, University of Yaounde I Yaounde, P.O.Box 47, Cameroon*

ABSTRACT

The method of Substructural Molecular Fragments based on the representation of the molecular graph by ensembles of fragments and involving calculations of those contributions to a given property. We also use the relationship between the topochemical indices, Wiener's index : defined as the sum of all distance between unordered pairs of vertices, Zagreb group parameter M1 and M2: defined as the summation of the squares of chemical degrees over all the vertices an adjacency and eccentric connectivity index : defined as the summation of the product of chemical eccentricity and the chemical degree of each vertex with anthranylic acids has been investigated. A data set comprising of 100 analogues of anthranylic acids was selected for the present study. The values of the Wiener's index, Zagreb group parameter, and eccentric connectivity index were computed for each of the 100 analogues using an in-house computer program and suitable models were developed after identification of the active ranges. For the first model, the predicted values for the biological activity of the structures in the prediction set are pertinent: the plot of A_{cal} vs. A_{obs} showed a correlation R^2= 0.9175. Subsequently for the second model, each compound was assigned a biological activity using these models, which was then compared with the reported anti-flammatory activity. Accuracy of prediction was found to be, ≈86% using models based upon topochemical descriptors.

Keywords: Substructural Molecular Fragments, anthranylic acids, topochemical descriptors, ISIDA/QSPR

INTRODUCTION

Anthranilic acids belong to the category of non-steroidal anti-inflammatory drugs [1, 2]. They are amino isosteres of salicylates and are also known as fenamates. Important molecules of this class include mefenamic acid, flufenamic acid and meclofenamic acid. As an analgesic agent, mefenamic acid has been used to relieve pain arising from rheumatic conditions, soft tissue injuries, other painful musculoskeletal conditions and dysmenorrhea. Fenamates act by blocking the metabolism of arachidonic acid by the enzyme cyclooxygenase (COX), one of the key enzymes in the arachidonic acid cascade [3]. This enzyme, also known as prostaglandin H Synthase (PGH synthase /PGHS/PHS) is a prominent and well-studied protein which catalyzes the conversion of arachidonic acid to prostaglandin H_2 (PGH_2), the committed step in prostaglandin (PG) biosynthesis. There are two isoforms of this enzyme: cyclooxygenase 1 (COX1) and cyclooxygenase 2 (COX2). COX1 are responsible for the maintenance and the protection of the gastrointestinal tract, COX2 is responsible for inflammation and pain [4]. The existing non-steroidal and anti-inflammatory drugs (NSAIDs) differ in their relative specificities for COX-1 and COX-2; while aspirin is equipotent at inhibiting COX-2 and COX-1 enzymes in vitro [5]. The finding of the structure of a molecule had an important role to play in its biological activity coupled with the need for safer potent drugs to be

developed with minimum expenditure, animal sacrifice and time loss led to the quantity of structure-activity relationship (QSAR) studies.

Molecular structure is the central theme of chemistry. According to the principle of molecular structure, properties, and behavior of molecules follow from their structures. If one considers nonmetric properties of a molecule, then the molecule can be represented by a (fragment) graph, which is essentially a nonnumeric mathematical object. Measurable properties of a molecule are usually expressed by means of numbers. Hence, to correlate property or activity of a molecule with its topology, one must first convert by an algorithm the information contained in the graph to a numerical characteristic and then one can establish relationships between structure of chemical compounds and their properties. [6]

In the present study, relationship of structural molecular fragment, Wiener's topochemical index, eccentric connectivity topochemical index and Zagreb's topochemical index with Anthranilic acids has been investigated.

MATERIALS AND METHODS

Substructural Molecular Fragments model

Substructural Molecular Fragments (SMF) is the method developed in ISIDA/QSPR [7]; the latest is based on the splitting of a molecular graph on fragments (subgraphs), and on the calculation of their contributions to a given property Y. Two classes of fragments are used: "sequences" (I) and "augmented atoms" (II). Three sub-types AB, A and B are defined for each class. For the fragments I, they represent sequences of atoms and bonds (AB), of atoms only (A), or of bonds only (B). Shortest or all paths from one atom to the other are used. For each type of sequences, the minimal (n_{min}) and maximal (n_{max}) number of constituted atoms must be defined. Thus, for the partitioning I(AB, n_{min} - n_{max}), I(A, n_{min} - n_{max}) and I(B, n_{min}- n_{max}), the program generates "intermediate" sequences involving n atoms ($n_{min} \leq n \leq n_{max}$). In the current version of ISIDA/QSPR, $n_{min} \geq 2$ and $n_{max} \leq 15$. The number of sequences' types of different length corresponding to $n_{min} = 2$ and $n_{max} = 15$ is equal to 105 for each of three subtypes AB, A and B, totally 315 types of sequences. QSPR modeling was performed using Multiple Linear Regression Analysis (MLR) of the ISIDA/QSPR program[8] with combined forward and backward stepwise variable selection techniques.[9] MLR is applied to build linear relationships between independent variables (SMF descriptors: N_i i =1, 2,…) and a dependent variable (here target property Y = A): $Y = a0 + \Sigma aiNi$ (1), where every descriptor value is associated with observed property value (Y), a_i is descriptor contribution, and a0 is the independent term which is omitted in a part of models. The Singular Value Decomposition method is used to fit contributions a_i and to minimize the sum of squared residuals which are squared differences between the property values calculated by the model (y_{calc}) and observed values (y_{exp}) in the training set. The program can generate more than 25,000 MLR models; each of them corresponds to particular type of the SMF descriptors and MLR equation (a0= 0 or a0≠0) and applied variable selection technique. In order to validate consensus model, the external 5-fold cross validation (5-CV) was applied. [11,12] ISIDA, implicitly keeps every 5th compound in the test set, the initial set was randomly split into 5 subsets, each of which was iteratively ignored at the training stage, in order to serve as internal validation set while the four others formed, together, the learning set. For each of these 5 splitting schemes, models were built followed by prediction calculations on the corresponding validation set. Finally, all values calculated for five test sets are merged into one file to analyze overall linear correlations between experimental and predicted property. One can use Determination Coefficient (R^2), Root Mean Squared Error (RMSE) or Mean Average Error (MAE), to estimate the quality of the linear correlation between predicted (Y_{pred}) and experimental (Y_{exp}) data for n compounds. Formulas for the statistical parameters are formulated below.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_{pred,i} - y_{exp,i})^2}{\sum_{i=1}^{n}(y_{exp,i} - \bar{y}_{exp})^2} \qquad (2)$$

Root –mean square error

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=n}^{n}(y_{pred,i} - y_{exp,i})^2} \qquad (3)$$

Mean average error

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left| y_{pred,i} - y_{exp,i} \right| \qquad (4)$$

ISIDA calculates a Consensus Model (CM) combining the information issued from several models. At the first step, hundreds of models are built using different initial pools of descriptors corresponding to different fragmentation

types. Then predictive performance (R^2LOO) is estimated using Leave One Out (LOO) procedure and the best models (R^2LOO>0.7) are combined into a consensus model. In the "leave one out" method, each compound is predicted in turn, based on a model learned from all other compounds. Predicted values are compared to experimental value, to compute leave one out cross-validation determination coefficient. For each compound from the test set, the program computes the property as an arithmetic mean of values obtained with these best models; those leading to outlying values were excluded according to Grubbs's statistics [13]. Generally, some 30 individual MLR models were used in consensus calculations.

Topochemical models

In general, a topological index, sometimes also known as a graph-theoretic index, is a numerical invariant of a graph. There are several topological indices having been defined such as Wiener index, Zagreb index. Recently, a lot of results on the eccentric connectivity index have been obtained and some of them have been applied as means for modeling chemical, pharmaceutical and other properties of molecules, [14, 15, 16].

Throughout this paper, all graphs we considered are simple and connected. Let G = (V (G), E(G)) be a simple connected graph with n vertices and m edges. For a vertex v ∈ V (G), d_G (v) (or just d (v) briefly) denotes the degree of v . δ (G), Δ(G) represent the minimum and maximum degree of G, respectively. For vertices u, v ∈ V (G), the distance d (u, v) is defined as the length of the shortest path between u and v in G. The eccentricity ε (v) of a vertex v is the maximum distance from v to any other vertex.

Wiener's topochemical index (Wc): It is a topochemical version of oldest and most widely used distance based topological index – Wiener's index[17] and this modified index takes into consideration the presence as well as relative position of heteroatom in a hydrogen suppressed molecular structure. Wiener's topochemical index is defined as the sum of the chemical distances between all the pairs of vertices in hydrogen suppressed molecular graph. [18]

$$W(G) = \sum_{\{u,v\} \subset V(G)} d(u,v) = \frac{1}{2} \sum_{v \in V(G)} D(G) \qquad (5)$$

The first and second Zagreb indices were first introduced by Gutman and Trinajstic [19]. It is reported that these indices are useful in the study of anti-inflammatory activities of certain chemical instances.

$$M_1(G) = \sum_{u \in V(G)} d(u)^2 \qquad (6)$$

$$M_2 = \sum_{uv \in E(G)} d(u)d(v) \qquad (7)$$

Eccentric connectivity topochemical index (ξ^c(G)) is defined as the summation of the product of chemical eccentricity and the chemical degree of each vertex in the hydrogen suppressed molecular graph having n vertices [26], that is

$$\xi^C(G) = \sum_{v \in V(G)} e(v)d(v) \qquad (8)$$

RESULTS AND DISCUSSION

Substructural Molecular Fragments model

A dataset comprising of 100 anthranilic acids (Figure 1) was selected**.** [20] Structural Molecular Fragment developped in ISIDA/QSPR. The modeled physical or chemical property Y can be quantitatively calculated accounting for contributions of fragments using linear equation (1), as we told it before, ai are fragment contributions, Ni is the number of fragments of i type. The ao term is fragment independent. An extra term Γ = ΣcmDm can be used to describe any specific feature of the compound using external descriptors Dm (e.g., topological, electronic, etc.); by default Γ= 0. The equation (1) represents calculation of property Y by using additive contributions of fragments.

The contributions of ai are calculated by minimizing a functional

$$U(a_i) = \sum_{i=1}^{n} W_i (Y_{exp,i} - Y_{cal,i})^2 \Rightarrow min \qquad (9)$$

where n is the number of the compounds in the training set, wi the weight accounting for the accuracy of the experimental data, Y_{exp} and Y_{calc} are, respectively, experimental and calculated according to (1) property values (table 1) and the program plot: calculated vs experimental property for compound set (figure 2), graphical analysis of residuals (figure 3), LOO predicted vs experimental property for training set (figure 4) and LMO predicted vs experimental property for training set (figure 5).

In this work, our model took 582 descriptors (fragments), of which only 33 descriptors contributed in the determination of the calculation of the property.

Figure.1 chemical structure of anthranilic acid

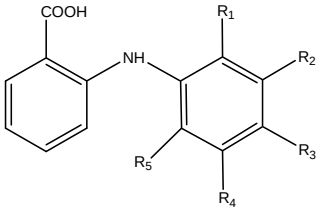


Table1: A dataset of 100 anthranilic acids with anti-inlammatory activity

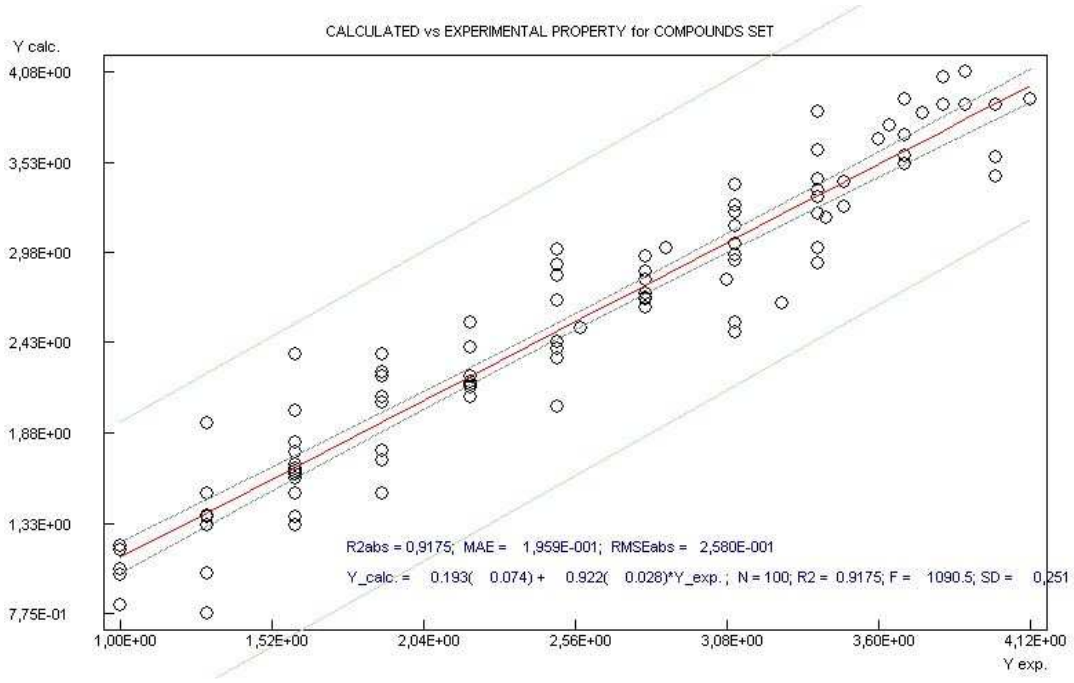
mol	R_1	R_2	R_3	R_4	R_5	MED[a]	A_{exp}	A_{cal}	A_{exp}- A_{cal}
1	H	H	H	H	H	200	1.300000	1.503339	-0.203339
2	H	CF_3	H	H	H	3.3	3.080000	2.808559	0.271441
3	H	CH_3	H	H	H	100	1.600000	1.652845	-0.052845
4	H	Cl	H	H	H	25	2.200000	2.543298	-0.343298
5	H	NH_2	H	H	H	400	1.000000	1.010030	-0.010030
6	H	OCH_3	H	H	H	50	1.900000	1.503339	0.396661
7	H	$SO_2N(CH_3)_2$	H	H	H	50	1.900000	1.707636	0.192364
8	H	$COCH_3$	H	H	H	200	1.300000	1.375310	-0.075310
9	H	$N(CH_3)_2$	H	H	H	100	1.600000	1.622921	-0.022921
10	H	H	Cl	H	H	200	1.300000	0.774721	0.525279
11	H	C_4H_9	H	H	H	200	1.300000	1.363197	-0.063197
12	H	CN	H	H	H	25	2.200000	2.400071	-0.200071
13	H	C_3H_7	H	H	H	50	1.900000	1.767528	0.132472
14	H	SCH_3	H	H	H	100	1.600000	1.503339	0.096661
15	H	NO_2	H	H	H	100	1.600000	1.685576	-0.085576
16	H	OC_2H_5	H	H	H	100	1.600000	1.503339	0.096661
17	H	Br	H	H	H	50	1.900000	2.351296	-0.451296
18	H	C_2H_5	H	H	H	25	2.200000	2.171859	0.028141
19	Cl	H	H	H	H	50	1.900000	2.094197	-0.194197
20	CH3	H	H	H	H	200	1.300000	1.503339	-0.203339
21	H	H	CH3	H	H	400	1.000000	1.164721	-0.164721
22	Cl	H	Cl	H	H	100	1.600000	1.365578	0.234422
23	H	Cl	Cl	H	H	100	1.600000	1.814680	-0.214680
24	CH3	CH_3	H	H	H	10.4	2.580000	2.510477	0.069523
25	CH3	CF_3	H	H	H	1	3.600000	3.666191	-0.066191
26	CH3	$SO_2N(CH_3)_2$	H	H	H	6.2	2.800000	2.722831	0.077169
27	CH3	NH2	H	H	H	50	1.900000	2.246686	-0.346686
28	CH3	$N(CH_3)_2$	H	H	H	6.2	2.800000	2.859577	-0.059577
29	CH3	Cl	H	H	H	5.3	2.870000	2.996924	-0.126924
30	CH3	OCH_3	H	H	H	6.2	2.800000	2.950619	-0.150619
31	H	CF_3	H	CF_3	H	100	1.600000	1.600000	-0.000000
32	Br	CF_3	H	H	H	1.6	3.390000	3.595250	-0.205250
33	Br	Br	H	H	H	3.1	3.110000	3.137988	-0.027988
34	H	CH_3	H	CH_3	H	100	1.600000	1.314226	0.285774
35	Cl	H	H	H	CH_3	12.5	2.500000	2.992235	-0.492235
36	Br	CN	H	H	H	1.5	3.420000	3.186762	0.233238
37	F	Cl	H	H	H	3.1	3.110000	2.543298	0.566702
38	H	Cl	H	Cl	H	50	1.900000	2.217521	-0.317521
39	Cl	Cl	H	H	H	2.1	3.270000	2.663507	0.606493
40	CH_3	NO_2	H	H	H	3.1	3.110000	2.922232	0.187768
41	CH_3	CN	H	H	H	3.1	3.110000	3.257703	-0.147703

42	CH_3	C_2H_5	H	H	H	3.1	3.110000	3.029491	0.080509
43	Cl	H	H	H	Cl	3.1	3.110000	2.486483	0.623517
44	Cl	CH_3	H	H	H	6.2	2.800000	2.697328	0.102672
45	Cl	H	H	Cl	H	12.5	2.500000	2.389331	0.110669
46	CH_3	H	H	H	CH_3	50	1.900000	2.062759	-0.162759
47	CH_3	H	H	CH_3	H	200	1.300000	1.503339	-0.203339
48	H	CH_3	CH_3	H	H	200	1.300000	1.314226	-0.014226
49	CH_3	H	CH_3	H	H	400	1.000000	0.826103	0.173897
50	CH_3	$SO_2N(CH_3)_2$	H	H	Cl	0.7	3.750000	3.827169	-0.077169
51	Cl	Cl	H	Cl	H	3.1	3.110000	2.958642	0.151358
52	H	Cl	Cl	Cl	H	200	1.300000	1.018253	0.281747
53	CH_3	CH_3	H	CH_3	H	25	2.200000	2.171859	0.028141
54	CH_3	H	CH_3	CH_3	H	100	1.600000	1.683735	-0.083735
55	H	Cl	CH_3	Cl	H	100	1.600000	2.006750	-0.406750
56	CH_3	H	CH_3	H	CH_3	400	1.000000	1.046904	-0.046904
57	Cl	$SO_2N(CH_3)_2$	H	H	Cl	1.3	3.480000	3.402831	0.077169
58	Cl	OCH_3	H	H	Cl	0.3	4.120000	3.905000	0.215000
59	CH_3	Br	H	H	CH_3	1.6	3.390000	2.910716	0.479284
60	Cl	CN	H	H	Cl	1.6	3.390000	3.836841	-0.446841
61	CH_3	Cl	H	H	Cl	3.1	3.110000	3.389210	-0.279210
62	CH_3	Cl	H	H	CH_3	0.4	4.000000	3.556344	0.443656
63	Cl	OC_2H_5	H	H	Cl	0.8	3.690000	3.905000	-0.215000
64	CH_3	$COCH_3$	H	H	CH_3	0.9	3.640000	3.744712	-0.104712
65	CH_3	$N(CH_3)_2$	H	H	CH_3	1.6	3.390000	3.418997	-0.028997
66	C_2H_5	NO2	H	H	C_2H_5	12.5	2.500000	2.677674	-0.177674
67	NH_2	Cl	H	H	CH_3	25	2.200000	2.216149	-0.016149
68	CH_3	CH_3	H	Cl	H	25	2.200000	2.184699	0.015301
69	CH_3	CN	H	H	CH_3	0.4	4.000000	3.438622	0.561378
70	CH_3	SCH_3	H	H	CH_3	0.4	4.000000	3.872166	0.127834
71	CH_3	NO_2	H	H	Cl	1.6	3.390000	3.314518	0.075482
72	CH_3	C_3H_7	H	H	CH_3	6.2	2.800000	2.806079	-0.006079
73	C_2H_5	$SO_2N(CH_3)_2$	H	H	C_2H_5	12.5	2.500000	2.427038	0.072962
74	C_2H_5	$COCH_3$	H	H	C_2H_5	25	2.200000	2.095288	0.104712
75	Cl	H	CF_3	H	Cl	0.8	3.690000	3.690000	-0.000000
76	CH_3	$SO_2N(CH_3)_2$	H	H	CH_3	0.5	3.900000	4.076463	-0.176463
77	CH_3	NH_2	H	H	Cl	6.2	2.800000	2.638972	0.161028
78	CH_3	CH_3	H	H	Cl	12.5	2.500000	2.902764	-0.402764
79	Cl	Cl	H	H	CH_3	0.8	3.690000	3.561545	0.128455
80	Cl	H	C_2H_5	H	Cl	0.8	3.690000	3.512326	0.177674
81	Cl	H	Cl	Cl	H	400	1.000000	1.190064	-0.190064
82	Cl	Cl	Cl	H	H	200	1.300000	1.934889	-0.634889
83	Cl	H	Cl	H	Cl	100	1.600000	1.757864	-0.157864
84	NH_2	CH_3	H	H	CH_3	25	2.200000	2.183851	0.016149
85	CH_3	CH_3	H	H	CH_3	6.2	2.800000	2.691396	0.108604
86	Cl	CH_3	H	H	CH_3	3.1	3.110000	3.216865	-0.106865
87	CH_3	Cl	H	CH_3	H	1.6	3.390000	2.996924	0.393076
88	CH_3	C_2H_5	H	H	CH_3	1.6	3.390000	3.210410	0.179590
89	CH_3	NH_2	H	H	Cl	1.3	3.480000	3.251863	0.228137
90	CH_3	SO_2CH_3	H	H	CH_3	0.6	3.820000	3.872166	-0.052166
91	Cl	$N(CH_3)_2$	H	H	Cl	0.6	3.820000	4.042193	-0.222193
92	CH_3	$SOCH_3$	H	H	CH_3	0.5	3.900000	3.872166	0.027834
93	Cl	Cl	Cl	H	CH_3	12.5	2.500000	2.832926	-0.332926
94	CH_3	CH_3	H	CH_3	CH_3	100	1.600000	2.352778	-0.752778
95	Cl	Cl	Cl	H	Cl	12.5	2.500000	2.327175	0.172825
96	Cl	CH_3	Cl	H	Cl	12.5	2.500000	2.035218	0.464782
97	Cl	Cl	Cl	Cl	H	100	1.600000	1.759374	-0.159374
98	Cl	Cl	H	Cl	Cl	1.6	3.390000	3.350928	0.039072
99	Cl	Cl	Cl	Cl	Cl	25	2.200000	2.151661	0.048339
100	CH_3	CH_3	Cl	CH_3	Cl	100	1.600000	1.637597	-0.037597

a: the biological activity A was calculated from the minimal effective dose (MED mg/kgbody) by formula: A= log(4000/MED)

We are not going to represent the matrix of contribution (33*100), because it is big enough. But, we are going to give the two better calculated property equations (10, 11), because their residual are equal to 0.

Figure 2: calculated vs. experimental property for compound set



$$Y_{cal,31} = -0.398275(0.16) \times N_{C-C=O} + 0.204297(0.0705) \times N_{C-N-C} + 1.236656(0.125) \times N_{C-C-C-N} - 0.49331(0.158) \times N_{C=C-C-N} + 1.44728(0.241) \times N_{C-C-C-O} - 0.731878(0.289) \times N_{C-C-C=C-N} - 0.338618(0.0577) \times N_{C-C-C=C-C} - 0.887086(0.281) \times N_{C-C=C-C-F} + 0.597106(0.177) \times N_{C-C=C-C-C-F} + 1.22(0.348) \times N_{C-C-C=C-N-C-C-C-O} + 0.675218(0.173) \times N_{C-C=C-C-N-C=C-C-C-F} \quad (10)$$

With (According the descriptor matrix).

$$N_{C-C=O} = 1, N_{C-N-C} = 1, N_{C-C-C-N} = 1, N_{C=C-C-N} = 2, N_{C-C-C-O} = 1, N_{C-C-C=C-N} = 1, N_{C-C-C=C-C} = 2, N_{C-C=C-C-F} = 6, N_{C-C=C-C-C-F} = 6, N_{C-C-C=C-N-C-C-C-O} = 1, N_{C-C=C-C-N-C=C-C-C-F} = 3$$

Figure3: Graphical analysis of residuals

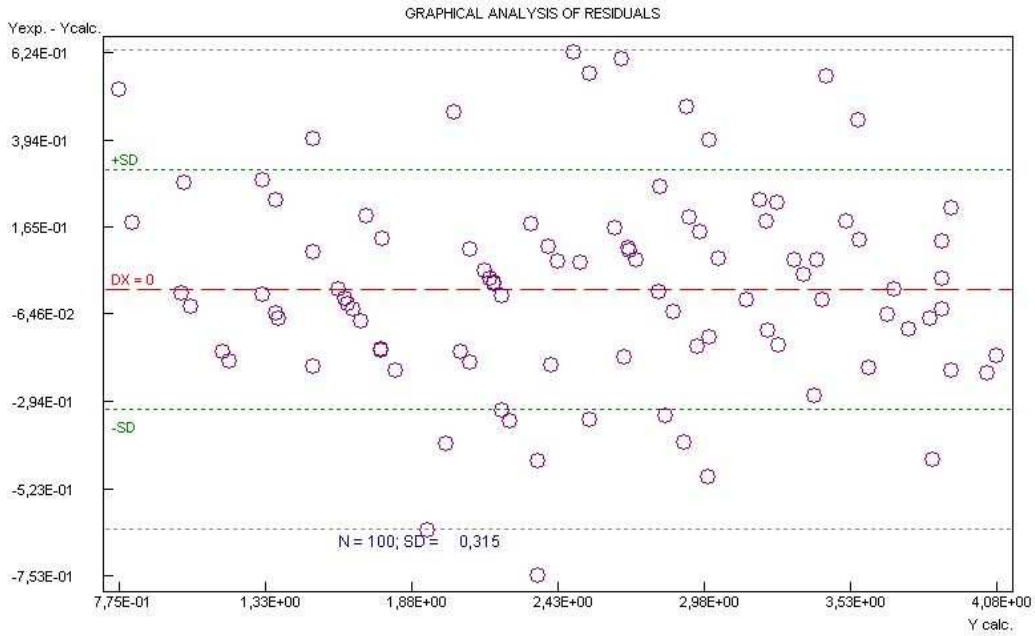


Figure 4: LOO predicted vs. experimental property for training set

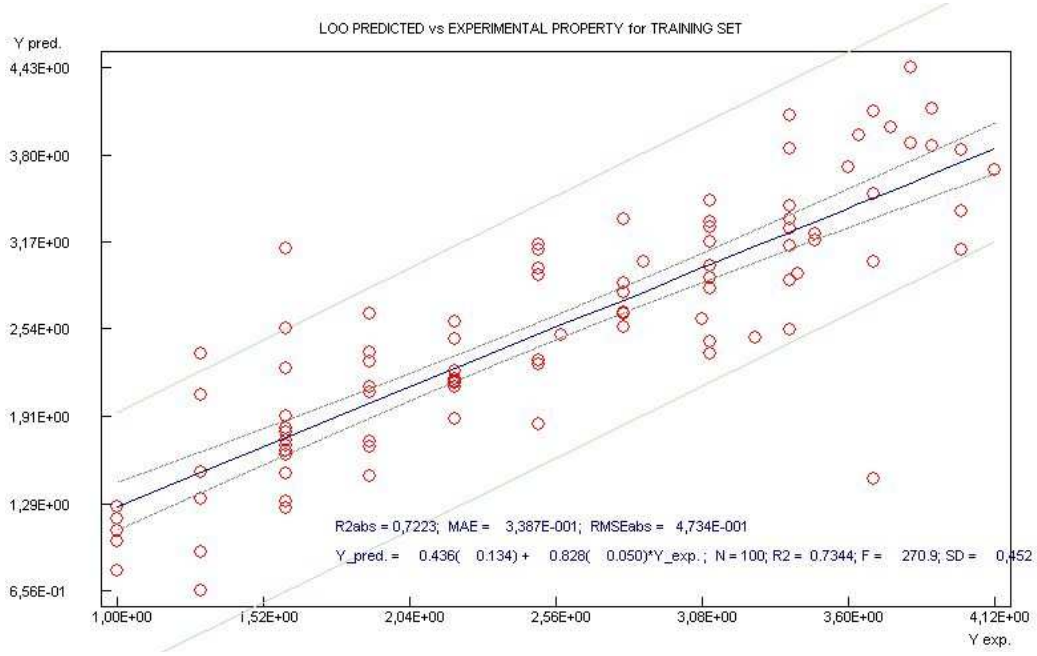
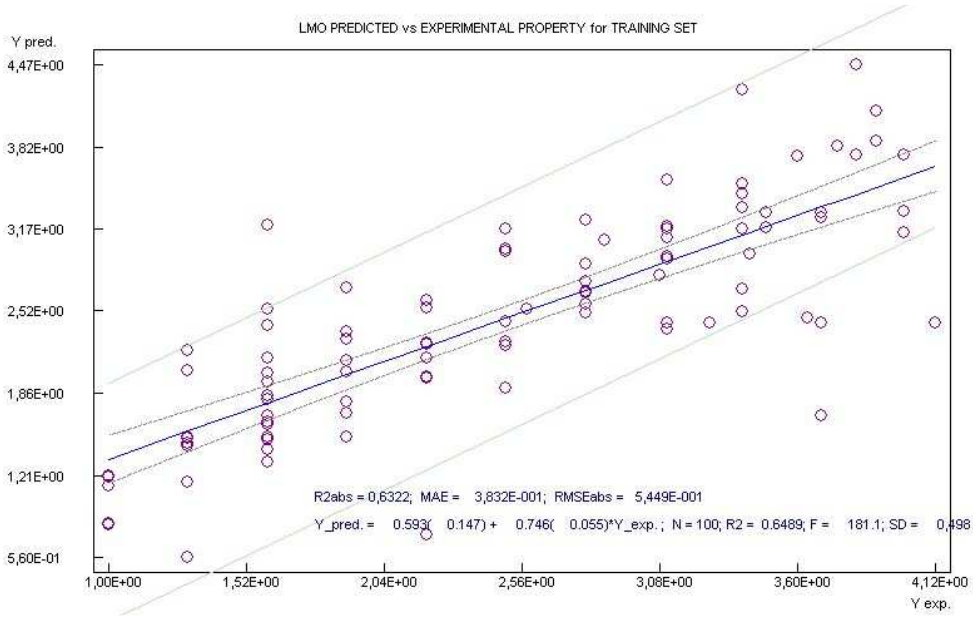


Figure 5: LMO predicted vs. experimental property for training set



$$Y_{cal,75} = -0.398275(0.16) \times N_{C-C=O} + 0.204297(0.0705) \times N_{C-N-C} + 1.236656(0.125) \times N_{C-C-C-N} - 0.49331(0.158) \times N_{C=C-C-N} + 1.44728(0.241) \times N_{C-C-C-O} - 0.338618(0.0577) \times N_{C-C-C=C-C} - 0.887086(0.281) \times N_{C-C=C-C-F} + 0.597106(0.177) \times N_{C-C=C-C-C-F} - 0.325778(0.103) \times N_{Cl-C=C-C} + 0.916635(0.138) \times N_{Cl-C=C-N-C-C-C-O} + 0.453626(0.151) \times N_{Cl-C-C-C} \quad (11)$$

With (According the descriptor matrix)

$$N_{C-C=O} = 1, N_{C-N-C} = 1, N_{C-C-C-N} = 2, N_{C=C-C-N} = 2, N_{C-C-C-O} = 1,$$
$$N_{C-C-C=C-C} = 3, N_{C-C=C-C-F} = 3, N_{C-C=C-C-C-F} = 6,$$
$$N_{Cl-C=C-C} = 1, N_{Cl-C=C-N-C-C-C-O} = 1, N_{Cl-C-C-C} = 1$$

A total number of 100 models, sharing 33 descriptors among them, were obtained through MLR. All these 33 descriptors along with their brief meaning, average regression coefficients and total incidence , which will

serve as a measure of their estimate across these models. These models have accounted for up to 91.75 (R^2=0.9175) percent variance in the observed activities.
The results of the QSAR study give rise to QSAR models with good predictive ability for anti-inflammatory activity. Linear regression for the total data set of 100 anthranilic derivative in the present study with the anti-inflammatory activity demonstrated that the fragment descriptors (C-C=O , C-N-C , C-C-C-N, C=C-C-N, C-C-C-O) appears to be the governing factors for the anti-inflammatory potency for synthesized anthranilic derivatives.

- For residual: $\left|A_{exp} - A_{cal}\right| \leq 2.10^{-1}$, 69 out of 100 equations were predicted correctly.
- For residual: $\left|A_{exp} - A_{cal}\right| < 10^{-1}$, 35 out of 100 equations were predicted correctly.
- For residual: $\left|A_{exp} - A_{cal}\right| < 0.05$, 18 out of 100 equations were predicted correctly.
- For residual: $\left|A_{exp} - A_{cal}\right| = 0$, 2 out of 100 equations were predicted correctly.

Topochemical models
The values of topochemical indices were calculated using an in-house computer program. Resulting data was analyzed and suitable models were developed after identification of the active ranges by maximization of moving average with respect to active compounds (<35 % = inactive, 35–6 5 % = transitional, >65% = active).[21] Subsequently, each compound was assigned a biological activity using these models, which was then compared with the reported anti-inflammatory activity (table 2).

This is the proposed model for anti-inflammatory activity of anthranilic acids:

- for Wiener's index, lower inactive range is > 775 , transitional range is $775 \rightarrow < 900$, active range is $900 \rightarrow < 1074$, upper inactive range is ≥ 1074
- for Zagreb's index M1, lower inactive range is > 100, transitional range is $100 \rightarrow < 104$, active range is $104 \rightarrow < 112$, upper inactive range is ≥ 112
- for Zagreb's index M2, lower inactive range is > 112, transitional range is $112 \rightarrow < 121$, active range is $121 \rightarrow < 129$, upper inactive range is ≥ 129
- for eccentric connectivity's index, lower inactive range is > 296 , transitional range is $296 \rightarrow < 313$, active range is $313 \rightarrow < 336$, upper inactive range is ≥ 336

Table 2: Relationships between topochemical indices and activity of derivates anthranilic acids

N0	Activity	W	M1	M2	ξ^c	A	W	M1	M2	ECI
1	1,301	447	78	88	217	-	-	-	-	-
2	3,083	828	102	116	304	-	-+	-+	-+	-+
3	1,602	528	84	95	232	-	-	-	-	-
4	2,204	528	84	95	232	-	-	-	-	-
5	1	528	84	95	232	-	-	-	-	-
6	1,903	626	88	100	270	-	-	-	-	-
7	1,903	1074	112	130	363	-	-	-	-	-
8	1,301	726	94	107	287	-	-	-	-	-
9	1,602	726	94	107	287	-	-	-	-	-
10	1,301	538	84	96	255	-	-	-	-	-
11	1,301	877	96	108	357	-	-+	-	-	-
12	2,204	626	88	100	270	-	-	-	-	-
13	1,903	742	92	104	310	-	-	-	-	-+
14	1,602	626	88	100	270	-	-	-	-	-
15	1,602	726	94	107	287	-	-	-	-	-
16	1,602	742	92	104	310	-	-	-	-	-+
17	1,903	528	84	95	232	-	-	-	-	-
18	2,204	626	88	100	270	-	-	-	-	-
19	1,903	518	84	96	230	-	-	-	-	-
20	1,301	518	84	96	230	-	-	-	-	-
21	1	538	84	95	255	-	-	-	-	-
22	1,602	613	90	103	268	-	-	-	-	-
23	1,602	622	90	103	270	-	-	-	-	-
24	2,585	602	90	104	245	-	-	-	-	-
25	3,602	914	108	125	317	+	+	+	+	+
26	2,809	1170	118	139	376	-	-	-	-	-
27	1,903	602	90	104	245	-	-	-	-	-
28	2,809	808	100	116	300	-	-+	-+	-+	-+
29	2,877	602	90	104	245	-	-	-	-	-
30	2,809	704	94	109	283	-	-	-	-	-

31	1,602	1297	126	144	370	-	-	-	-	-
32	3,397	914	108	125	317	+	+	+	+	+
33	3,11	602	90	104	245	-	-	-	-	-
34	1,602	613	90	102	247	-	-	-	-	-
35	2,505	593	90	104	243	-	-	-	-	-
36	3,426	704	94	109	283	+	-	-		-
37	3,11	602	90	104	245	-	-	-	-	-
38	1,903	613	90	102	247	-	-	-	-	-
39	3,279	602	90	104	245	+	-	-	-	-
40	3,11	808	100	116	300	-	-+	-+	-+	-+
41	3,11	704	94	109	283	-	-	-	-	-
42	3,11	704	94	109	283	-	-	-	-	-
43	3,11	593	90	104	243	-	-	-	-	-
44	2,809	602	90	104	245	-	-	-	-	-
45	2,505	604	90	103	245	-	-	-	-	-
46	1,903	593	90	104	243	-	-	-	-	-
47	1,301	900	106	124	313	-	+	+	+	+
48	1,301	622	90	103	270	-	-	-	-	-
49	1	613	90	103	268	-	-	-	-	-
50	3,757	1282	124	147	389	+	-	-	-	-
51	3,11	692	96	111	260	-	-	-	-	-
52	1,301	710	96	111	285	-	-	-	-	-
53	2,204	692	96	111	260	-	-	-	-	-
54	1,602	702	96	111	283	-	-	-	-	-
55	1,602	710	96	111	285	-	-	-	-	-
56	1	682	96	112	258	-	-	-	-+	-
57	3,488	1282	124	147	389	+	-	-	-	-
58	4,124	790	100	117	296	+	-+	-+	-+	-+
59	3,397	682	96	112	258	+	-	-	-+	-
60	3,397	790	100	117	296	+	-+	-+	-+	-+
61	3,11	682	96	112	258	-	-	-	-+	-
62	4	682	96	112	258	+	-	-	-+	-
63	3,699	918	104	121	336	+	+	+	+	-
64	3,647	900	106	124	313	+	+	+	+	+
65	3,397	900	106	112	313	+	+	+	-+	+
66	2,505	1126	114	134	343	-	-	-	-	-
67	2,204	682	96	112	258	-	-	-	-+	-
68	2,204	692	96	111	260	-	-	-	-	-
69	4	790	100	117	296	+	-+	-+	-+	-+
70	4	790	100	117	296	+	-+	-+	-+	-+
71	3,397	900	106	124	313	+	+	+	+	+
72	2,809	918	104	121	336	-	+	+	+	-
73	2,505	1548	132	157	419	-	-	-	-	-
74	2,204	1126	114	134	343	-	-	-	-	-
75	3,699	1168	114	133	330	+	-	-	+	+
76	3,903	1282	124	147	389	+	-	-	-	-
77	2,809	682	96	112	258	-	-	-	-+	-
78	2,505	682	96	112	258	-	-	-	-+	-
79	3,699	682	96	112	258	+	-	-	-+	-
80	3,699	790	100	117	296	+	-+	-+	-+	-+
81	1	702	96	111	283	-	-	-	-	-
82	1,301	700	96	112	283	-	-	-	-+	-
83	1,602	692	96	111	281	-	-	-	-	-
84	2,204	682	96	112	258	-	-	-	-+	-
85	2,809	682	96	112	258	-	-	-	-+	-
86	3,11	682	96	112	258	-	-	-	-+	-
87	3,397	692	96	111	260	+	-	-	-	-
88	3,397	790	100	117	296	+	-+	-+	-+	-+
89	3,488	900	106	124	313	+	+	+	+	+
90	3,823	1012	114	133	330	+	+	-	+	+
91	3,823	900	106	124	313	+	+	+	+	+
92	3,903	900	106	124	313	+	+	+	+	+
93	2,505	784	102	120	296	-	-+	-+	-+	-+
94	1,602	775	102	120	273	-	-+	-+	-+	-
95	2,505	784	102	120	296	-	-+	-+	-+	-+
96	2,505	784	102	120	296	-	-+	-+	-+	-+
97	1,602	793	102	120	298	-	-+	-+	-+	-+
98	3,397	775	102	120	273	+	-+	-+	-+	-+
99	2,204	880	108	129	311	-	-+	+	-	-+
100	1,602	880	108	129	311	-	-+	+	-	-+

-: Inactive compound (compounds having A less than 3.204),+: active compound, +-: transitional, W —Wiener's index, ξ^C—eccentric connectivity index, M1—Zagreb index M1and M2—Zagreb index M2 and A-- reported activity.

The methodology used in the present studies aims at the development of suitable models for providing lead molecules through exploitation of the active ranges in the proposed models based on topochemical indices. Proposed models are unique and differ widely from conventional QSAR models. Both systems of modeling have their own advantages and limitations. In the instant case, the modeling system adopted has distinct advantage of identification of narrow active range(s), which may be erroneously skipped during routine regression analysis in conventional QSAR modeling. Since the ultimate goal of modeling is to provide lead structures, therefore, these active ranges can play vital role in lead identification [22].

Retrofit analysis of the data in table 2 reveals following information with regard to Wiener's topochemical index:

- 54 out 60 compounds in the lower inactive range were predicted correctly (90%).
- A transitional range with index values of 775 to <900 was observed. Existence of a transitional range is ideal because it simply indicates gradual change in biological activity.
- 10 out of 12 compounds in the active range were predicted correctly (83.33%).
- 6 out 10 compounds in the upper inactive range were predicted correctly (60%)
- The overall predictability of the model based upon the wiener's index was 85.36 %.

Retrofit analysis of the data in table 2 reveals following information with regard to Zagreb's topochemical index_M1:

- 55 out 61 compounds in the lower inactive range were predicted correctly (90.16%).
- A transitional range with index values of 100 to <104 was observed. Existence of a transitional range is ideal because it simply indicates gradual change in biological activity.
- 9 out of 13 compounds in the active range were predicted correctly (69.23%).
- 6 out 11 compounds in the upper inactive range were predicted correctly (54.54%)
- The overall predictability of the model based upon the Zagreb's index_M1 was 84.7%.

Retrofit analysis of the data in table 2 reveals following information with regard to Zagreb's topochemical index_M2:

- 46 out 49 compounds in the lower inactive range were predicted correctly. (93.8%)
- A transitional range with index values of 112 to <121 was observed. Existence of a transitional range is ideal because it simply indicates gradual change in biological activity.
- 10 out of 8 compounds in the active range were predicted correctly. (80%)
- 13 out 8 compounds in the upper inactive range were predicted correctly (61.53%)
- The overall predictability of the model based upon the Zagreb's index_M2 was 87.5 %.

Retrofit analysis of the data in table 2 reveals following information with regard to eccentric connectivity's topochemical index:

- 57 out 60 compounds in the lower inactive range were predicted correctly (95%).
- A transitional range with index values of 296 to <313 was observed. Existence of a transitional range is ideal because it simply indicates gradual change in biological activity.
- 10 out of 11 compounds in the active range were predicted correctly (90.9%).
- 8 out 12 compounds in the upper inactive range were predicted correctly (66.66%).
- The overall predictability of the model based upon the eccentric connectivity's index was 86, 58 %.

CONCLUSION

The results and discussion made above lead to the conclusion that the anti-inflammatory activity of series of anthranilic acids can be successfully modeled using structural molecular fragment and topological indices. The study using structural molecular fragment revealed that for anti-inflammatory activity, 33 out of 582 descriptors (fragments) were contributed for a good prediction of our model. Linear correlation between experimental and predicted property is very good (R^2= 0.9175). The studies using topological models are unique and differ widely from conventional QSAR models, the model based upon Zagreb's topochemical index_M2: has also demonstrated good predictability. Amongst the Zagreb indices, M2 has proven to be better in this study with higher predictability than M1.

REFERENCES

[1] J.C. Dearden, J.C. Duffy, *J.Pharm. pharmacol*., **1993**, 45, 1142-1146.
[2] J.C. Dearden, J.C. Duffy, C. J .C.,Roston, *J.Pharm. pharmacol*., **1996**, 48(9), 883–886.
[3] J. R. Vane, *Nature*., **1971**, 231, 232-235.
[4] D.S. Goodsell ; *Oncologist* 5, **2000**, 169-171.
[5] S. Debyani, R.K. Sahu, *Researcher*., **2010**. 2(10), 17-23.

[6] V. P. Solov'ev, A. Varnek, G. Wipff, *J. Chem. Inf. Comput. Sci.*, **2000**, 40, 847-858.
[7] ISIDA (In Silico Design and Data Analysis) Sofware, http://infochim.ustrasbg.fr/recherche/isida/index.php**: 2007**.
[8] A.Varnek, D. Fourches, F. Hoonaker, V. Solov'ev *J. Comput. Aided Mol*. Des**., 2005**, 19(9-10), 693-703.
[9] D. Horvath, F. Bonachera, V .Solov'ev, C.Gaudin, A .Varnek, *J.Chem.Inf.Model*., **2007,** 47(3), 927-939.
[10] V.P. Solov'ev, A. Varnek, *Rus. Chem. Bull*., **2004**, 53(7), 1434-1445.
[11] A.Varnek, N. Kireeva, I. V. Tetko, I. Baskin, V. Solov'ev, *J. Chem. Inf. Model.* **2007,** 47(3), 1111–1122.
[12] I. V Tetko, V.P. Solov'ev, A.V. Antonov, X.Yao, J.P. Doucet, B.Fan, F. Hoonakker, D. Fourches, P. Jost, *J. Chem. Inf. Model*. **2006**, 46(2), 808–819.
[13] F.E Grubbs, *Technometrics*. **1969**, 11, 1-21.
[14] S. Gupta, M. Singh, A. K. Madan, *J. Math. Anal. Appl*. **2002**, 266, 259-268.
[15] A. Ilic, I. Gutman, B. Furtula, *Math. Chem. Monogr*, **2010**, 9, 139-168.
[16] V. Sharma, R. Goswami, A. K. Madan, *J. Chem. Inform. Model.*, **1997**, 37, 273-282.
[17] H. Wiener, J. Am. Chem. Soc., **1947**, 69, 2636-2638.
[18] S. Bajaj, S.S. Sambi, A. K. Madan, *J. Mol. Struct. (THEOCHEM).*, **2004**, 684, 197-203.
[19] I. Gutman, N. Trinajstic, *Chem. Phys. Lett*., **1972**, 17, 535–538.
[20] U. Oleg, K. Gabriel, V.D. Mircea, *Rev. Roum. Chim*. **2004**, 3400, 1-12.
[21] S.Gupta, M. Singh, A. K. Madan, *J. Comput. Aid. Mol. Des*., **2001**, 15, 671-678.
[22] H. Dureja, A. K. Madan, *J. Mol. Graph. Mod.*, **2006**, 25, p. 373-379.