STUDIA UNIVERSITATIS BABEŞ-BOLYAI, CHEMIA, XLVIII, 2, 2003

MODELING BIOLOGICAL ACTIVITY OF (3-PYRIDYLMETHYL)BENZOQUINONE DERIVATIVES

CRINA V. VERES, MIHAELA CAPRIOARA, CSABA L. NAGY and MIRCEA V. DIUDEA

Faculty of Chemistry and Chemical Engineering, Babes-Bolyai University, 400028 Cluj, Romania

ABSTRACT. (3-Pyridylmethyl)benzoquinone derivatives inhibit thromboxane A_2 synthase and leukotriene biosynthesis enzymes. For such chemicals, a QSAR study has been performed in order to stress the relation between the molecular structures and their biological activity. The models are built up on the ground of hypermolecule concept. Partial charges, autocorrelated by a multivariate regression, are used for molecular description. The hypermolecule has the meaning of a complement of the investigated biological receptor space. A general procedure for developing and validating QSAR models is given.

INTRODUCTION

Lipoxygenases (LOX) are a family of cytosolic enzymes. They are monomeric proteins that contain a "non-heme" iron per molecule in the active site as high-spin Fe(II) in the native state, and high-spin Fe(III) in the activated state.[1-3] Lipoxygenases as dioxygenases recognize the 1,4-pentadiene structure of polyunsaturated fatty acids and catalyze their oxygenation to corresponding lipid hydroperoxides. LOX have a high degree of homology in the proposed iron binding region and the primary reactions they catalyze are essentially the same. Lipoxygenases are responsible for the biosynthesis of leukotrienes, eicosanoids, monohydroperoxy- and dihydroperoxy-polyenoic fatty acids, lipoxins. An overproduction of these products can cause disturbances in the metabolic reactions, and are involved in some metabolic diseases and pathologies. These effects have been linked to immunological and radiation disorders, tumors, toxicoses, hypodinamy, coronary and angiological pathologies (vasospasm, thrombosis and arteriosclerosis).

The major products of 5-LOX, leukotrienes (LT), are a family of important biologically active molecules. LTB_4 is a potent chemotactic agent and inflammatory mediator[4] and the peptidoleukotrienes LTC_4 and LTD_4 are powerful spasmogens in vascular and bronchial tissues[5]. Elevated levels of LT are associated with a number of inflammatory conditions including asthma, psoriasis, ulcerative colitis, and rheumatoid arthritis, and indeed LT have been recovered from various pathological tissues. Therefore, potent inhibitors of this enzyme are candidate drugs for the treatment of these diseases.[6,7] These inhibitors can be broadly classified into two main categories: first, competitive lipid substrate inhibitors and, second, redox-type inhibitors, which act by intermediate produced during the catalytic step.[8]

(3-Pyridylmethyl)benzoquinone derivatives[9] inhibit thromboxane A_2 synthase and leukotriene biosynthesis enzymes. For the inhibition of 5-LOX the experiments were made in human blood (human whole blood assay).

Many thousands of compounds have been screened as LOX inhibitors in industrial laboratories and a large number of active compounds with novel structures are undergoing clinical trials. This evaluation provides data sets suitable for quantitative structure-activity relationships (QSAR). The laboratory tests utilized in identifying lipoxygenase inhibitors are: human granulocytes, rat basophilic cells (RBL-1) and human whole blood assay (HWBL).

For drugs acting as LOX inhibitors, the hydrophobicity[10-14] is an important property, also significant in their susceptibility to the attack of the P450 enzymes.[15]

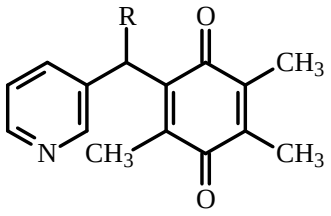
METHODS

A QSAR study was made for the (3-pyridylmethyl)benzoquinone derivatives.[9] The negative logarithm of IC_{50} was used in the correlation analyses. Thus a higher log $1/IC_{50}$ value represents a more potent compound. The following QSAR model is obtained with the data reported in Structure 1 and Table 1:

$$\log 1/IC_{50} = 0.49(\pm 0.234)MR - 1.914(\pm 0.421)I_{COOH} - \\ -1.641(\pm 0.535)I_{Ph} + 2.174(\pm 2.077) \quad (1)$$

n = 19, r = 0.957, r^2 = 0.916, q^2 = 0.858, s = 0.359, $F_{3.15}$ = 6.949

In the following QSAR model, *MR* is the overall calculated molecular refractivity. Since *MR* is a primarily measure of the bulk of the substituent, the positive coefficient for this term indicates that molecules are contacting polar space in the enzyme[16], not hydrophobic space. A positive coefficient might suggest an interaction depending on the polarizability of the substituents although there is a little evidence for the importance of such an effect. On the other hand I_{COOH} and I_{Ph} (indicator variables having a value of 1 when R has a carboxylic or phenyl group) have negative signs, which mean that the presence of these groups decreases the inhibition of LOX.



Structure 1

However, if the experimental measured lipophilicity log P_E is used in place of *MR*, equation (2) is obtained:

$$\log 1/IC_{50} = 0.498(\pm 0.149)\log P_E - 1.439(\pm 0.562)I_{Ph} + 4.461(\pm 0.533) \quad (2)$$

n = 19, r = 0.921, r^2 = 0.849, q^2 = 0.810, s = 0.467, $F_{2.16}$ = 33.118

Equation 2 shows that there is a linear relationship between the biological activity (BA) log $1/IC_{50}$ and log P_E (experimental lipophilicity).

Table 1

R	CD_1	CD_2	CD_4	CD_5	CD_6	CD_8	CD_{13}	CD_{15}	CD_{16}	CD_{19}	Y
ci	2758.2	2313.8	224.5	8758.1	1395	-15326	-101.05	9339.8	10725	-20586	
Me	0.034	0.031	-0.232	0.063	0.118	0.062	0.060	0.118	0.018	0.018	5.022
CH_2Me	0.033	0.031	-0.235	0.064	0.117	0.062	0.060	0.117	0.017	0.017	6.018
$(CH_2)_3Me$	0.033	0.031	-0.239	0.064	0.117	0.062	0.060	0.116	0.016	0.016	6.081
$(CH_2)_5Me$	0.033	0.031	-0.241	0.064	0.116	0.062	0.060	0.115	0.016	0.015	7.137
$=CH(CH_2)_4Me(Z)$	0.037	0.033	-0.240	0.065	0.118	0.063	-0.004	0.119	0.018	0.019	6.585
$=CH(CH_2)_4Me(E)$	0.037	0.033	-0.240	0.065	0.118	0.063	-0.004	0.119	0.018	0.019	6.658
$(CH_2)_4OH$	0.035	0.032	-0.233	0.065	0.117	0.063	0.063	0.117	0.017	0.017	7.000
$(CH_2)_5OH$	0.034	0.032	-0.241	0.065	0.117	0.063	0.062	0.116	0.017	0.016	7.000
$(CH_2)_4OAc$	0.038	0.035	-0.239	0.068	0.120	0.065	0.069	0.120	0.019	0.020	6.602
$(CH_2)_5OAc$	0.037	0.034	-0.240	0.067	0.119	0.065	0.066	0.119	0.019	0.018	7.060
$(CH_2)_3COOH$	0.038	0.034	-0.238	0.067	0.120	0.065	0.069	0.120	0.020	0.020	5.000
$(CH_2)_4COOH$	0.034	0.037	-0.240	0.067	0.119	0.064	0.066	0.119	0.019	0.019	4.222
$(CH_2)_5COOH$	0.036	0.033	-0.241	0.066	0.118	0.064	0.064	0.118	0.018	0.018	5.032
$(CH_2)_6COOH$	0.035	0.033	-0.242	0.066	0.118	0.064	0.063	0.117	0.017	0.017	5.155
$=CH(CH_2)_4COOH$	0.039	0.035	-0.239	0.068	0.119	0.065	0.000	0.121	0.020	0.021	5.143
Ph	0.036	0.032	-0.243	0.066	0.117	0.063	0.079	0.117	0.017	0.017	5.119
PhCH=C(Me)COOH	0.038	0.035	-0.246	0.068	0.119	0.065	0.083	0.119	0.018	0.019	4.000
PhCH=CHCOOH	0.038	0.034	-0.245	0.068	0.119	0.065	0.083	0.119	0.018	0.019	4.060
$Ph(CH_2)_2COOH$	0.038	0.034	-0.245	0.068	0.118	0.065	0.082	0.119	0.018	0.018	4.000

CALCULATION OF MOLECULAR DESCRIPTORS

In order to achieve the QSAR, the structure is encoded in a numerical form. The arrangement of substituent groups on the (3-Pyridylmethyl)benzoquinones can be accounted for by the *hypermolecule HM* concept,[17] viewed as a complement of the investigated biological receptor space. In the construction of the hypermolecule, a *row-vector* P_i of dimension N_{HM} is attached to each molecule M_i :

$$P_i = \{P_{ij};\ j = 1,2,..., N_{HM}\} \tag{3}$$

where N_{HM} is the number of vertices in the hypermolecule. The molecules of the set are superimposed according to their maximal common substructures. In the associated vector, the matching positions take P_{ij} = 1 while for the non-matching ones P_{ij} = 0. The description of the j^{th} position in *HM* (*e.g.*, the chemical and/or topological nature of the j^{th} atoms) is given by X_{ij}.

The hypermolecule and the molecules under study can be numerically described by using some molecular *topological descriptors* TD_I .[18,19]

$$TD_i = \sum_j TD_{ij} = \sum_j a_j P_{ij} X_{ij} \tag{4}$$

where P_{ij} and X_{ij} have the meaning above mentioned while a_j represent the regression coefficients as given by the multivariate regression log*BA* = *f*(X_i). The above *TDs* are "ad-hoc" descriptors[20,21] depending on the set of molecules considered and the selected molecular property. Therefore, all the (3-pyridylmethyl)benzoquinones can be described using particular local properties that characterize both the substituted/ unsubstituted positions.

As local property for calculating *TDs*, we used X_{ij} = partial charges related to each position in (3-pyridylmethyl)benzoquinones.

STATISTICS

For developing the QSAR models, the set of (3-Pyridylmethyl) benzoquinones was submitted to REGLINEWIN (a home made statistical software package). The correlating equation is of the form:

$$Y_i = b_0 + \sum_{j=1}^{m} b_j \cdot TD_{ij} \tag{5}$$

where Y_i is the dependent variable, TD_{ij} are the predictor variables, $m < n$, n being the number of structures in the set.

The correlating algorithm follows the steps:

1. generate the *hypermolecule*
2. perform Leave-one-out LOO analysis on the set of local descriptors
3. calculate the global descriptors by using partial charges
4. find the best regression equation
5. test the predictive capability of the model

DESCRIPTOR REDUCTION

The LOO procedure was applied on the set of descriptors in view of finding the relevant descriptors (*i.e.*, the relevant positions in the hypermolecule). The results are listed in Table 2.

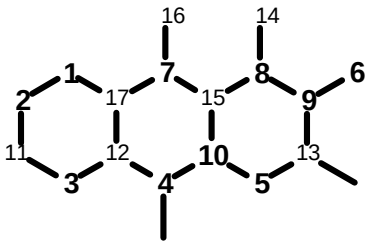
Table 2

No of descriptors	Model (MLR)	Irrelevant positions	r*
k-1	BA = f(C_i), *i* =1, 2,… *k*-1	C_{11}	0.999
k-2	BA = f(C_i), *i* =1, 2,… *k*-2	C_{11}, C_{12}	0.998
k-3	BA = f(C_i), *i* = 1, 2,… *k*-3	C_{11}, C_{12}, C_{13}	0.996
k-4	BA = f(C_i), *i* = 1, 2,… *k*-4	C_{11}, C_{12}, C_{13}, C_{14}	0.995
k-5	BA = f(C_i), *i* = 1, 2,… *k*-5	C_{11}, C_{12}, C_{13}, C_{14}, C_{15}	0.992
k-6	BA = f(C_i), *i* = 1, 2,… *k*-6	C_{11}, C_{12}, C_{13}, C_{14}, C_{15}, C_{16}	0.989
k-7	BA = f(C_i), *i* = 1, 2,… *k*-7	C_{11}, C_{12}, C_{13}, C_{14}, C_{15}, C_{16}, C_{17}	0.988

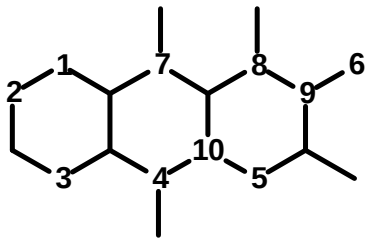
*after descriptor elimination

From Table 2, it appears that a satisfactory statistics is ensured by a map comprising *k*-7 (C_{11}, C_{12}, C_{13}, C_{14}, C_{15}, C_{16}, C_{17}) relevant positions, for which the partial charge descriptors CD_i are given in Table 1.

The numbering of the hypermolecule is given in Structures 2 (initial map) and 3 (final map).



Structure 2



Structure 3

Table 3 lists the global *CD* descriptor (i.e., the sum of CD_i descriptors over all vertices), experimental and calculated biological activity values and the residuals (i.e., the difference between the experimental and calculated BAs) as well.

Table 3

	Molecule	CD	$BA_{obs.}$	$BA_{calc.}$	Residual
1	Me	806.94	5.022	4.928	0.094
2	CH_2Me	807.835	6.018	5.809	0.209
3	$(CH_2)_3Me$	807.993	6.081	5.965	0.116
4	$(CH_2)_5Me$	809.211	7.137	7.165	-0.028
5	$=CH(CH_2)_4Me(Z)$	808.680	6.585	6.641	-0.056
6	$=CH(CH_2)_4Me(E)$	808.680	6.658	6.641	0.017
7	$(CH_2)_4OH$	809.041	7	6.997	0.003
8	$(CH_2)_5OH$	809.073	7	7.029	-0.028

	Molecule	CD	$BA_{obs.}$	$BA_{calc.}$	Residual
9	$(CH_2)_4OAc$	808.372	6.602	6.338	0.264
10	$(CH_2)_5OAc$	808.790	7.06	6.749	0.311
11	$(CH_2)_3COOH$	807.230	5	5.213	-0.213
12	$(CH_2)_4COOH$	806.223	4.222	4.221	0.001
13	$(CH_2)_5COOH$	807.285	5.032	5.267	-0.235
14	$(CH_2)_6COOH$	807.498	5.155	5.477	-0.322
15	$=CH(CH_2)_4COOH(E+Z)$	807.117	5.143	5.102	0.042
16	Ph	807.354	5.119	5.335	-0.216
17	PhCH=C(Me)COOH	805.762	4	3.766	0.234
18	PhCH=CHCOOH	806.099	4.06	4.098	-0.038
19	$Ph(CH_2)_2COOH$	806.159	4	4.157	-0.157

The best QSAR model found is:

$$\log 1/IC_{50} = -810.46 + 1.01 \cdot CD \quad (6)$$

n = 19; r = 0.988; s = 0.181; cv% = 3.22; F = 683.71

Random correlation: $r_1 = 0.662$; $r_2 = 0.691$;

A random test showed a marked drop in the correlation coefficient, with the meaning that no chance correlation occurred.

The model was cross-validated by LOO procedure applied on the set of molecules. Figure 1 presents the plot cross-validated BA vs. experimental BA, (i.e., log $1/IC_{50}$). The excellent result in prediction (q = 0.985; $q^2 = 0.970$) proves the quality of the model.

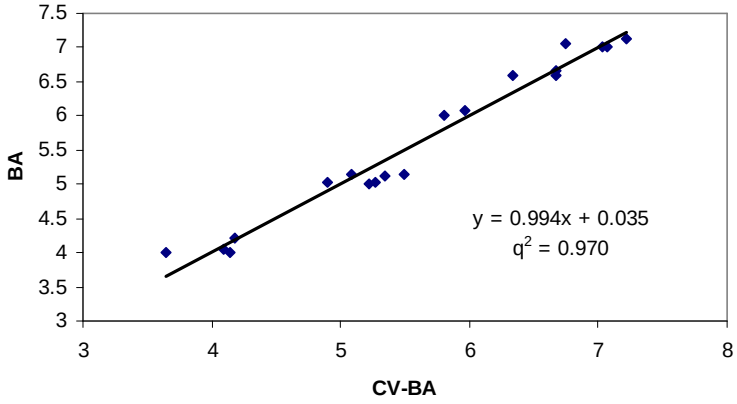


Figure 1. *The plot of LOO calculated vs. experimental biological activity*.

CONCLUSIONS

Partial charges were used as local descriptors for encoding the *hypermolecule*, that mimics the biological receptor space.

The descriptors have been fitted by means of multivariate regression and represent “ad-hoc” (or “autocorrelated”) descriptors, which change with the set of molecules and selected property.

Within the hypermolecule model, data reduction (*i.e*., selection of relevant descriptors) was achieved. The statistics of the model derived on the final map (the relevant substituted positions) are excellent both in estimation and prediction, surpassing the literature reported results.

REFERENCES

1. P. M. Hilbers, A. Rossi, A. Finazzi-Agro, G. A. Veldink and F. G. J. Vliegenthart, *Biochim*. *Biphys*. *Acta*, **1994**, *1211,* 239-242.
2. C. W. Stallings, B. A. Kroa, R. T. Carroll, A. L. Mertzer and M. O. Funk, *J. Mol. Biol.*, **1990**, *211*, 685-687.
3. J. C. Boyington and L. M. Amzel, *J. Mol. Biol.* **1990**, *165*, 12771-12773.
4. A. G. Gilman, T.W. Rall, A. S. Nies, P. Taylor (Ed.) The Pharmacological Basis of Therapeutics, McGraw-Hill International Editions, Medical Series, **1992**, pp. 600-617.
5. M. O’Donnell and A. Welton, *Therapeutic Approaches in Inflammatory Disease*, Eds. A. J. Lewis and N. S. Doherty, Elsevier, New York, **1989**, p. 169.
6. P. J. Piper, L. G. Letts, M. N. Samhoun, J. R. Tippins and M. A. Palmer, *Advances in Prostaglandin,Thromboxane and Leukotriene Research*, Vol 9, Raven Press, New York, **1982**, p. 169.
7. K. Muller, *Arch. Pharm*., **1994**, *327*, 3-19.
8. R. D. Abeysinghe, P. J. Roberts, Ch. E. Cooper, K. H. MacLean, R. C. Hider, J. B. Porter, *J. Biol. Chem.*, **1996**, *271*, 7965-7972.
9. S. Ohkawa, S. Terao, Z. Terashita, Y. Shibouta, and K. Nishikawa, *J*. *Med*. *Chem*., **1991**, *34*, 267–276.
10. K. H. Kim, Y. C. Martin, D. W. Brooks, D. Dyer and G. W. Carter, *J*. *Pharm*. *Sci*., **1994**, *83*, 433–438.
11. K. H. Kim, Y. C. Martin and D. W. Brooks, *Quant*. *Struct*. *Act*. *Relat*., **1996**, *15*, 491–497.
12. Y. Naito, M. Sugiura, Y. Yamaura, C. Fukaya, K. Yokoyama, Y. Nakagawa, T. Ikeda, M. Senda and T. Fujita, *Chem*. *Pharm*. *Bull*., **1991**, *39*, 1736–1745.
13. J. B. Summers, B. P. Gunn, J. G. Martin, M. B. Martin, H. Mazdiyasni, A. O. Stewart, P. R. Young, J. B. Bouska, A. M. Goetze, R. D. Dyer, D. W. Brooks and G. W. Carter, *J*. *Med*. *Chem*., **1988**, *31*, 1960–1964.
14. J. B. Summers, K. H. Kim, H. Mazdiyasni, J. H. Holms, J. D. Ratajczyk, A. O. Stewart, R. D. Dyer and G. W. Carter, *J*. *Med*. *Chem*., **1990**, *33*, 992–998.
15. C. Hansch and L. Zhang, *Drug Meta*b. *Re*v., **1993**, *25*, 1–48.
16. C. Hansch and J. Caldwell, *J*. *Comput*.*–Aided Mol*. *De*s., **1991**, *5*, 441–453.
17. J. A. Platts, M. H. Abraham, D. Butina, and A. Hersey, *J*. *Chem*. *Inf*. *Comput*. *Sci*., **2000**, *40*, 71-80.

18. R. P. Schvarzenbach, P. M. Gschwend and D. M. Imboden, *Environmental Organic Chemistry*, John Wiley&Sons, New York, **1992**.
19. O. M. Minailiuc and M. V. Diudea, TI-MTD Model. Applications in molecular Design, in: *QSPR/QSAR Studies by Molecular Descriptors*, Nova Science Publishers, Inc., **2001**, pp. 363- 388.
20. R. P. Sheridan, *J*. *Chem*. *Inf*. *Comput*. *Sci*., **2000**, *40*, 1456-1469.
21. I. Motoc, S. Holban, R. Vancea and Z. Simon, *Studia Biophys.* (*Berlin*) **1977**, *66*, 75-78.