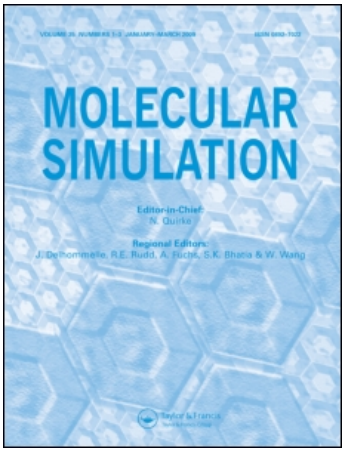

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Molecular Simulation of Chemical Structure and Biological Activity of Thromboxane Synthase Inhibitors by Multivariate Neural-network Analysis

Peter P. Mager[a]
[a] Research Group of Pharmacochemistry, Institute of Pharmacology and Toxicology of the University, Leipzig, Saxony, Germany

Online publication date: 26 October 2010

To cite this Article Mager, Peter P.(2003) 'Molecular Simulation of Chemical Structure and Biological Activity of Thromboxane Synthase Inhibitors by Multivariate Neural-network Analysis', Molecular Simulation, 29: 1, 71 — 75
To link to this Article: DOI: 10.1080/0892702031000065746
URL: http://dx.doi.org/10.1080/0892702031000065746

Molecular Simulation, 2003 Vol. 29 (1), pp. 71–75

Molecular Simulation of Chemical Structure and Biological Activity of Thromboxane Synthase Inhibitors by Multivariate Neural-network Analysis

PETER P. MAGER*

Research Group of Pharmacochemistry, Institute of Pharmacology and Toxicology of the University, D-04107 Leipzig, Härtelstr. 16–18, Saxony, Germany

(Received May 2001; In final form June 2001)

The effect of [2-substituted-4-(3-pyridyl)-1,3-dioxan-5-yl]alkenoic acids of thromboxane synthase was examined using a multivariate quantitative structure-activity relationship (QSAR) model. As biological parameters, the dual inhibition of thromboxane receptors (inhibition of the *in vitro* aggregation of human platelets) and of thromboxane synthase were used. As chemical parameters, the inductive constant, van der Waals radius, the STERIMOL parameters *L* and *B*1 of the substituents, and the dihedral angle energy of the geometry-optimized conformations of the antagonists (hybrid Gasteiger molecular-mechanics approach) were chosen. The back-propagation neural network approach has the ability to approximate multivariate structure-activity functions with satisfactory accuracy.

Keywords: Multivariate quantitative-structure activity relationships, QSAR; Backpropagation neural network analysis; Thromboxane synthase; [2-substituted-4-(3-pyridyl)-1,3-dioxan-5-yl]alkenoic acids

1 INTRODUCTION

Thromboxane (TxA_2) is a potent vasoconstrictor and platelet aggregating agent. TxA_2 is synthesized by the action of thromboxane synthase. It has been suggested that TxA_2 antagonists may be implicated in therapy of various diseases of the heart-circulation system. The dual-acting thromboxane receptor/synthase inhibiting potency of [2-substituted-4-(3-pyridyl)-1,3-dioxan-5-yl]alkenoic acids was shown elsewhere [1]. To develop quantitative structure-activity relationships (QSARs), these dual response must be analyzed by multivariate regression approaches [2] or neural network methods that are able to treat simultaneously with two and more output variables [3]. The advantage of the first technique is well documented (review: [2], some other examples see [4,5]) while there are only a few examples that deal with multivariate neural network designs applied simultaneously to two batteries consisting of biological and physicochemical variables [6].

The aim of this study is to demonstrate the power of neural network analysis applied to a profile of biological variables.

2 METHOD

2.1 Synthesis

The synthetic route of the compounds (Fig. 1, Table I) was described previously [1], together with all data that verify the structures.

2.2 Pharmacological Data

The *in vitro* antagonist activity was determined at human platelet TXA_2 receptor site aggregated by the 16(*S*)-hydroxy-11$_\alpha$-9$_\alpha$-(epoxymethano)prosta-5(*Z*), 13(*E*)-dienoic acid in presence and absence of different concentrations of the antagonists to estimate the apparent pA_2 constants. The TXA_2 synthase inhibition was measured with human blood platelet microsomes incubated with [1-^{14}C]arachidonic acid in presence and absence of the antagonists to estimate the 50% inhibition concentration IC_{50} (μM). The pA_2 and IC_{50} data were taken

*E-mail: magp@medizin.uni-leipzig.de

ISSN 0892-7022 print/ISSN 1029-0435 online © 2003 Taylor & Francis Ltd
DOI: 10.1080/0892702031000065746

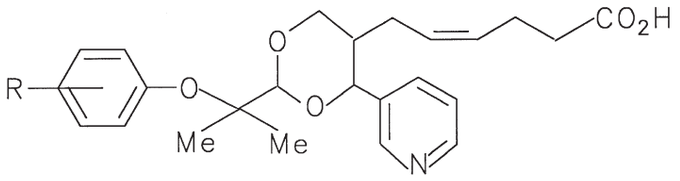

FIGURE 1 Lead structure of [2-substituted-4-(3-pyridyl)-1,3-dioxan-5-yl]alkenoic acids.

from [1]. The IC_{50} scores were transformed by natural logarithms (8 digits were used in calculation) to get normalized and variance-stabilized data.

2.3 Physicochemical Data

The inductive constant σ_I, van der Waals radius r_v, and the STERIMOL parameter L (length) and $B1$(minimum width) of the substituents were taken from literature [7–9] or from our compilation of linear-free enthalpy energy (LFER) parameters, which is downloadable from the Internet (http://www.uni-leipzig.de/ ~ pharma/ppm2.htm).

The conformations of the antagonists were energy minimized using the MM+ force field without an electrostatic term. The MM+ empirical potential (force field) is an improved MM2/MM3 version [10,11]. The whole MM+ procedure was repeated with electrostatic parameters of the connectivity-based iterative partial equalization of orbital electronegativity [12]. Correlation-gradient geometry optimization [13] was then achieved. The structures were refined using a conjugate gradient minimizer (Fletcher–Reeves modification of the Polak–Ribière method). Convergence was obtained when the gradient root mean square RMS was $RMS < 0.05$ kcal/Å mol. The various energies of the geometry-optimized conformations were determined, and the dihedral angle energy (kcal/mol) was chosen in order to model the three-dimensional properties of the antagonists.

2.4 Neural Network Approach

The algorithm of the generalized backpropagation neural network was applied [3,6,14,15]. The sigmoidal backpropagation functions were solved by the nonliner Levenberg–Marquardt algorithm using 100,000 iterations. Optimization was achieved by (i) estimating the global error vector prior to adjusting weights, and (ii) updating successively the weights until convergence was reached. The multivariate and squared multiple correlation coefficients were used as goodness-of-fit criteria and satistically tested according to the MASCA model ([2]; the free

TABLE I Substituents of [2-substituted-4-(3-pyridyl)-1, 3-dioxan-5-yl]alkenoic acids (Fig. 1), physicochemical LFER descriptors $X_1 = \sigma_I$ (inductive substituent constant), $X_2 = r_v$ (van der Waals radius), the STERIMOL parameter $X_3 = L$ (length) and $X_4 = B1$ (minimum width), and X_5 = torsional angle energy (kcal/mol) of Gasteiger-MM+ geometry-optimized conformations of the antagonists

Compound	*Substituent*	X_1	X_2	X_3	X_4	X_5
1	H	0.00	1.20	2.06	1.00	−5.12
2	4-Br	0.44	1.85	3.82	1.95	−5.09
3	4-OMe	0.25	1.56	3.98	1.35	−5.25
4	4-*tert*-Bu	−0.07	2.44	4.11	2.60	−5.56
5	4-CN	0.55	1.60	4.23	1.60	−5.11
6	4-SMe	0.18	1.84	4.30	1.70	−5.18
7	4-SO_2Me	0.59	2.08	4.11	2.03	−3.88
8	4-F	0.52	1.47	2.65	1.35	−5.11
9	4-Me	−0.06	1.87	3.00	1.52	−5.95
10	3-F	0.52	1.47	2.65	1.35	−5.05
11	2-F	0.52	1.47	2.65	1.35	−5.05
12	2-OMe	0.25	1.56	3.98	1.35	−6.72
13	2-SMe	0.18	1.84	4.30	1.70	−4.64
14	2-Br	0.44	1.85	3.82	1.95	−4.27
15	4-Br	0.44	1.85	3.82	1.95	−5.09
16	2-CN	0.55	1.60	4.23	1.60	−4.99
17	2-SO_2Me	0.59	2.08	4.11	2.03	−6.84
18	2-CO_2Me	0.31	2.21	4.73	1.64	−3.04
19	2-NO_2, 4-OMe	0.89	3.74	7.42	3.05	−1.39
20	2-NO_2, 4- Me	0.58	4.05	6.44	3.22	−2.07
21	2-NO_2, 4- F	1.16	3.65	6.09	3.05	−1.57
22	2-NO_2, 4-SMe	0.82	4.02	7.74	3.40	−1.59
23	2-CN, 4-Me	0.49	3.47	7.23	3.12	−5.79
24	2-CN, 5-Me	0.49	3.47	7.23	3.12	−5.69
25	2-CN, 6-Me	0.49	3.47	7.23	3.12	−6.39
26	2-CN, 4-F	1.07	3.07	6.88	2.95	−5.00

software can be downloaded from the Internet (http://www.uni-leipzig.de/~pharma/ppm2.htm).

3 RESULTS AND DISCUSSION

The lead structure and substituents of the compounds are given in Fig. 1 and Table I, together with the design matrix of the chemical parameters. The experimentally obtained profile of the biological activities ($Y_1 = pA_2$, $Y_2 = \ln[IC_{50}]$) is listed in Table II.

The simple linear correlation (r) between the dual responses is not overwhelming ($r = -0.57$) due to a deviation from linearity (cluster consisting of compounds 1–3, 6–8, 10–15, and 17; Fig. 2). Table III shows the product-moment correlation coefficients between the activity profile and physicochemical descriptors. The inductivity of the substituents determines mainly the inhibitory action against thromboxane synthase. As expected, the *in vitro* inhibition of platelet aggregation depends also on steric effects. Employing the traditional Hansch design [7–9], the squared multivariate correlation coefficient becomes $R^2 = 0.88$, and the squared multiple correlation coefficients are $R_1^2 = 0.77$ and $R_2^2 = 0.62$ (critical quantiles at the 5% significance level: maximum likelihood criterion = 0.584, Roy's largest root criterion = 0.498). There is a considerable multicollinearity among the physicochemical descriptors, however. Alternative approach must be applied for subsequent analysis, such as nonleast-squares and ridge regression, or neural networks methods.

The following layers of the optimized backpropagation neural network were used: one input layer with linear transfer functions and five nodes (physicochemical descriptors), one hidden layer with sigmoid transfer function and three nodes, one output layer with sigmoid transfer function and two nodes (biological activities). Full connections between the nodes were assumed. The momentum was equal to 0.8, and the learning rates were 0.001 (minimum) and 0.3 (maximum). The so-called generalized delta was used as learning rule.

To improve the interpretation of the results of network model, the weights should be visualized [16]. The weight matrix is listed in Table IV. The number of physicochemical descriptors is $c = 5$, and the number of biological variables is $p = 2$. There are only $p + c = 7$ weights per node so that the possible effect of overfitting can surely be excluded. Table V shows the goodness-of-fit results. The theoretically calculated biological activity profile is collected in Table II.

Within the experimental error rate and the approximate validity of molecular descriptors, the agreement between experimentally found and theoretically calculated biological responses is quite tolerable. The results support the hypothesis that

TABLE II Experimentally obtained (obtd) and theoretically calculated (calcd) inhibition of aggregation of human platelets, $Y_1 = pA_2$, and inhibition of thromboxane synthase, $Y_2 = \ln(IC_{50})$

Compound	Y_1		Y_2	
	Obtd	*Calcd*	*Obtd*	*Calcd*
1	7.12	6.66	−2.96	−2.99
2	6.42	6.78	−2.90	−3.22
3	6.52	6.76	−3.19	−3.19
4	5.72	5.98	−1.27	−1.18
5	6.89	7.24	−3.08	−3.37
6	6.43	6.72	−3.24	−3.12
7	7.04	7.14	−3.38	−3.35
8	6.92	6.77	−2.98	−3.23
9	6.20	5.95	−1.02	−1.14
10	6.90	6.77	−3.65	−3.23
11	6.78	6.77	−3.10	−3.23
12	6.65	6.62	−2.81	−2.87
13	7.05	6.76	−3.32	−3.19
14	7.20	6.82	−3.54	−3.26
15	6.48	6.78	−3.82	−3.22
16	7.88	7.29	−3.10	−3.39
17	6.23	6.70	−3.10	−3.01
18	7.19	7.06	−2.72	−3.32
19	8.21	8.60	−3.30	−3.73
20	8.34	8.30	−3.61	−3.61
21	8.60	8.60	−3.91	−3.73
22	8.35	8.60	−3.82	−3.73
23	8.10	7.98	−3.00	−3.01
24	7.80	8.04	−3.00	−3.09
25	7.60	7.52	−2.41	−2.35
26	9.15	8.60	−4.20	−3.72

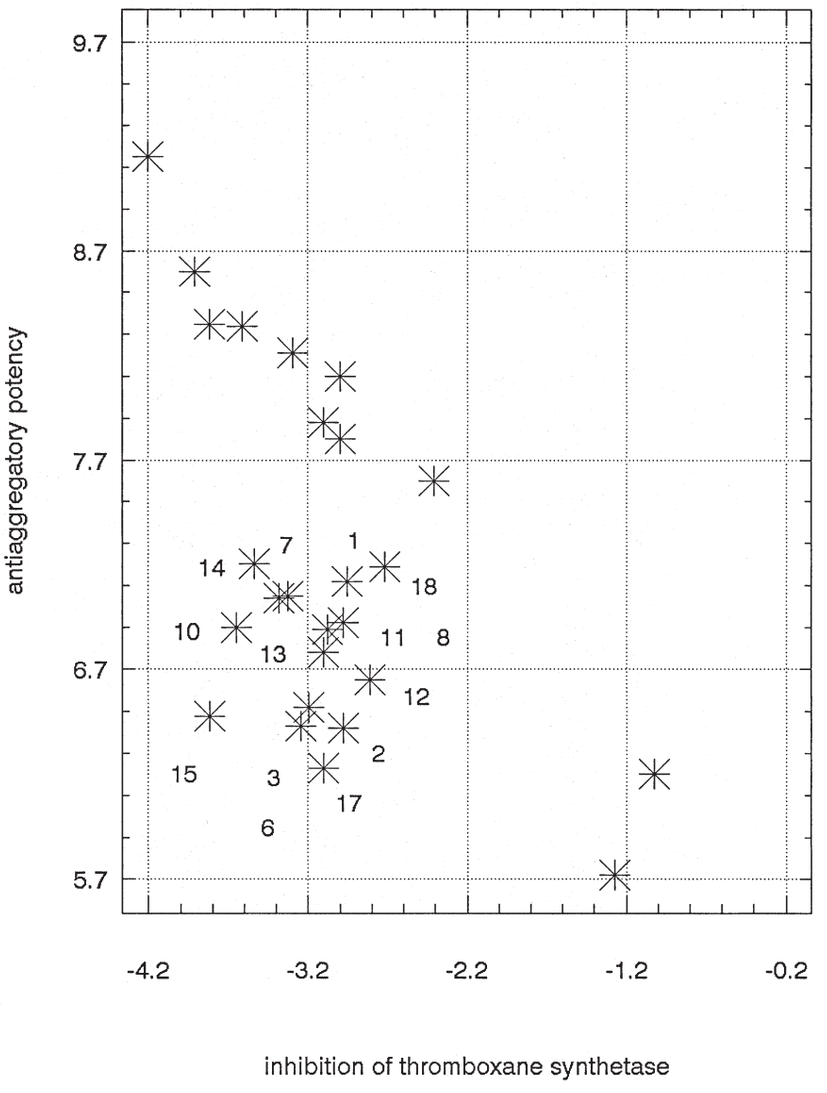

FIGURE 2 Linear relationships between the dual responses.

TABLE III Linear correlation coefficients between activity profile and physicochemical descriptors. Each estimated correlation coefficient that is absolutely equal to or larger than the critical value $C_0 = 0.557$ is significantly different from zero at the 5% level or less (Studentized maximum modulus test)

Descriptor	*Activity profile*	
	Y_1	Y_2
X_1	0.76	−0.71
X_2	0.72	−0.20
X_3	0.74	−0.26
X_4	0.66	−0.22
X_5	0.57	−0.44

TABLE IV Network weights and current adjustment deltas (node 1: response variable Y_1, node 2: response variable Y_2)

Layer	*Node*	*Connection*	*Weight*	*Weight delta*
2	1	1	16.96197	0.000030
2	1	2	−11.18712	−0.000048
2	1	3	24.82221	0.000060
2	1	4	−0.99330	0.000010
2	1	5	7.12548	0.000021
2	2	1	8.73060	0.000000
2	2	2	−13.35675	0.000000
2	2	3	5.10746	−0.000002
2	2	4	2.54491	0.000002
2	2	5	9.35706	−0.000004
3	1	1	1.58809	0.000000
3	1	2	2.66424	−0.000001
3	2	1	−0.49518	0.000000
3	2	2	−6.72591	0.000003

TABLE V Results of neural network analysis (node 1: response variable Y_1, node 2: response variable Y_2; SD = standard deviation, R^2 = squared multiple correlation coefficient)

NODE 1	*SD*	*Bias*	*Maximum error*	R^2
1	0.297	0.0014	0.594	0.875
2	0.265	0.0001	0.604	0.853

small substituents with electrophilic inductive effects, but with low resonance effects, may improve the biological response. Regarding the example, the optimized backpropagation neural network has therefore the ability to simulate multivariate structure-activity functions with satisfactory accuracy.

References

[1] Faull, A.W., Brewster, A.G., Brown, G.R., Smithers, M.J. and Jackson, R. (1995) "Dual-acting thromboxane receptor antagonists/synthase inhibitors: synthesis and biological properties of [2-substituted -4-(3-pyridyl)-1,3-dioxan-5-yl]alkenoic acids", *J. Med. Chem.* **38**, 686–694.

[2] Mager, P.P. (1991) Design Statistics in Pharmacochemistry (Wiley, New York).

[3] Schneider, G. (2000) "Neural networks are useful tools for drug design", *Neural Netw.* **13**, 15–60.

[4] Mager, P.P., De Clercq, E., Takashima, H., Ubasawa, M., Sekiya, K., Baba, M. and Walther, H. (1996) "Molecular simulation of 5,6-substituted 1-[2-hydroxyethoxy)methyl]uracils with anti-HIV-1 activity", *Eur. J. Med. Chem.* **31**, 701–712.

[5] Mager, P.P. (1994) "A_2 agonists: structure-activity relationships of 2-(cycloalkylalkynyl)adenosine derivatives", *Eur. J. Med. Chem.* **29**, 369–380.

[6] Mager, P.P. (1998) "Neural network approaches applied to selective A_{2a} adenosine receptor agonists", *Med. Chem. Res.* **8**, 277–290.

[7] Hansch, C. and Leo, A. (1995) Exploring QSAR: Fundamentals and Applications in Chemistry and Biology (American Chemical Society, Washington, DC).

[8] Fujita, T. (1990) "The extrathermodynamic approach to drug design", *Compr. Med. Chem.* **4**, 497–560.

[9] Hansch, C. and Leo, A. (1979) Substituent Constants for Correlation Analysis in Chemistry and Biology (Wiley, New York).

[10] Allinger, N.L. and Yan, L. (1993) "Molecular mechanics (MM3). Calculation of furan, vinyl, ethers, and related compounds", *J. Am. Chem. Soc.* **115**, 11918–11925.

[11] Tai, J.C., Yang, L. and Allinger, N.L. (1993) "Molecular mechanics (MM3). Calculation on nitrogen-containing aromatic heterocyclics", *J. Am. Chem. Soc.* **115**, 11906–11917.

[12] Gasteiger, J. and Marsili, M. (1980) "Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges", *Tetrahedron* **36**, 3219–3288.

[13] Frey, R.F., Coffin, J., Newton, S.Q., Ramek, M., Cheng, V.K.W., Momany, F.A. and Schäfer, L. (1992) "Importance of correlation-gradient geometry optimization for molecular conformational analyses", *J. Am. Chem. Soc.* **114**, 5369–5377.

[14] Mager, P.P. and Reinhardt, R. (1999) "A comparison of backpropagation and generalized-regression genetic-neural network models", *Drug Des. Discovery* **16**, 49–53.

[15] Jiang, J.-H., Wang, J.-H., Liang, Y.-Z. and Yu, R.-Q. (1996) "A non-linear mapping-based generalized backpropagation network for unsupervised learning", *J. Chemometr.* **10**, 241–252.

[16] Manallack, D.T. and Livingstone, D.J. (1999) "Neural network in drug discovery: have they lived up their promise?", *Eur. J. Med. Chem.* **34**, 195–208.