Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 7, No. 7, pp. 877-880, 1997
© 1997 Elsevier Science Ltd
All rights reserved. Printed in Great Britain
0960-894X/97 $17.00 + 0.00

PII: S0960-894X(97)00128-5

A QSAR STUDY OF ANTIPLATELET AGENTS USING ARTIFICIAL NEURAL NETWORK - CORRELATION WITH MICELLE-WATER PARTITION COEFFICIENT

Nanda Ghoshal*[a], Basudeb Achari[a] and Tapan K. Ghoshal[b]

(a) Indian Institute of Chemical Biology, Jadavpur, Calcutta - 700032, India

(b) Centre for Knowledge Based Systems, Jadavpur University, Calcutta - 700032, India

Abstract: Antiplatelet activity *ex vivo* data reported for 2-substituted phenyl- and benzimidazolyl-5-methyl-4-(3-pyridyl) imidazoles have been analysed using BP type ANN. Using micelle-water partition coefficient as an independent descriptor, a network system (1-3-1) produced very good duplication of observed activities (r=0.860, SD=0.183, n=21) in the training cycle. The results provide an improved model for prediction of antiplatelet activity. © 1997 Elsevier Science Ltd.

In a recent QSAR study[1], the antiplatelet activity of 2-substituted phenyl- and benzimidazolyl-5-methyl-4-(3-pyridyl) imidazoles had been shown to be non-linearly dependent on micelle-water partition coefficient (as $\log P_{mw}$), this is a parameter advocated[1] to be a descriptor for hydrophobicity and which is simpler to measure than the 1-octanol-water partition coefficient, the usual measure for hydrophobicity. This postulates a significant impact in the field of QSAR studies. However, about 20% of the experimental data were declared as outliers, neither the fit was very good nor the nature of the curve justified the data pattern.

This prompted us to carry out a detailed study using a back propagation (BP) type artificial neural network (ANN) method with the datasets provided by Tanaka *et al*[1] which yielded distinct improvements in the results. Earlier reports[2-6] on neural network fit indicated that ANN is at least as good as if not marginally better than the statistical method. In this paper it is shown that ANN can produce significantly better results even with a small dataset[7].

In the ANN investigation, the training was carried out with one input ($\log P_{mw}$). A very simple network configuration (1-3-1) was chosen to avoid overfitting[6]. Other three-layer networks (1-4-1, 1-6-1) and a four-layer configuration (1-2-2-1) were also tried. As four compounds (**12,16,17,22**)[1] were reported[1] as outliers, training were carried out with 22,21,19 and 18 members of the same dataset for systematic and detailed analysis. In another training cycle, the network was retrained using a 22-member dataset after an initial training

[*] To whom correspondence should be addressed. Fax: 91-33 - 4730284 Email: ckbsjuin@giascl01.vsnl.net.in

(initial weights) using the statistically predicted data in order to investigate whether a local minimum exists in the neighbourhood of the statistically fitted curve.

For testing the prediction ability, the leave-one-out method[4,8] was followed. The BP program[9] of McClelland and Rumelhart was used as the basic module for exploration. A learning rate of 0.05 and a momentum term of 0.09 were used. The programs were run on a PC/AT 486 machine. A total of 17,000 iterations were needed for convergence. The data were scaled between 0.1-1.0 (inputs and outputs) using the transformation equation[10] : X_i (scaled) = $(X_i-X_{min}+0.1)/(X_{max}-X_{min}+0.1)$.
A pre-processing program developed by the authors was used for this purpose.

Table 1. QSAR results (goodness of fit) by ANN and Statistical method

Serial No.	Number of Compounds in a dataset[a]	ANN results (1-3-1 network confign)		Result with Statistical method[b]	
1	22	r = 0.723	SD = 0.255	-	
2	21	r = 0.860	SD = 0.183	-	
3	19	r = 0.709	SD = 0.271	r = 0.632	SD = 0.325
4	18	r = 0.869	SD = 0.183	r = 0.772	SD = 0.257

(a) Data taken from Table 1, Reference 1. (b) Values taken from Reference 1.

The results of training cycles of various datasets are shown in Table 1. Training with 19 compounds (excluding **16,17** and **22**) gave poorer r and SD values while with 21 compounds (excluding only **12**) a significant improvement in correlation was obtained compared to the 22-compound dataset. The training of the dataset excluding the four compounds **12,16,17** and **22** was also carried out for systematic study. The result is almost identical to the 21 dataset case. Therefore, out of the four compounds (**12,16,17,22**) termed as outliers in the statistical method[1], only **12** may justifiably be called an outlier. Although the other three are more hydrophilic ($\log P_{mw}$ <1) than the rest, the observed values could be duplicated quite well by ANN (vide Table 2). Though **12** proved to be an outlier, its inclusion did not lead to any drastic change in the individual values of the other compounds which showed the robustness of the ANN method.

The results of the prediction cycle, shown in Table 2 as "per unit deviation" from the original 21 dataset fitting (training cycle) using 1-3-1 network configuration, indicate that the prediction ability of ANN is very good except for compound **8**. This was not altogether unexpected because any curve fitting method would produce large errors in extrapolation at an endpoint with large change of slope. Further, in ANN analysis with optimum scaling, the new data should lie within the range of the data studied, i.e. between the two extrema to avoid saturation and for obtaining good prediction results. With **8**, the prediction was very good during training cycle with all compounds including **8** as one of the extrema. But in the leave-one-out method, as **8** was removed from the training set , the range of data trained is shortened (excluding the logP$_{mw}$ value for **8**), leading to poorer prediction .

Table 2 : Antiplatelet activity (logAPA) observed, statistically predicted and predicted by ANN (using 1-3-1 network configuration) taking all the compounds during training and deviation (per unit) from those results of the data predicted using leave-one-out method

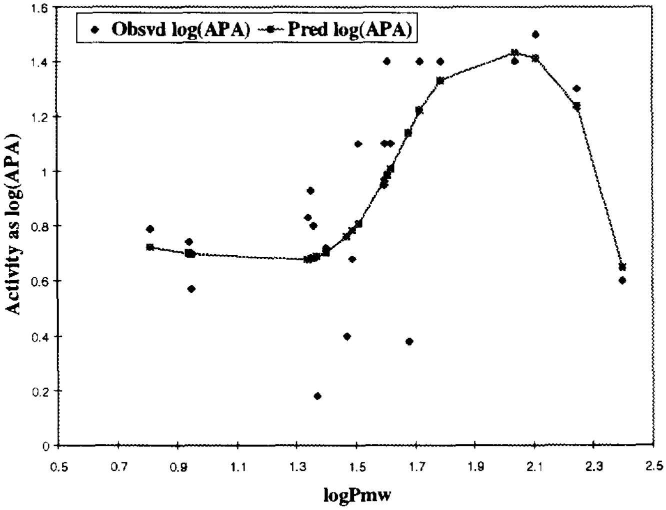
Compound No.[a]	Observed logAPA[b]	ANN method (1-3-1 network) Predicted logAPA				Statistical method Predicted[d] logAPA
		Training cycle		Leave-One-Out method	Per unit deviation[c]	
		22 dataset[e]	21 dataset[e] (excld.**12**)	21 data[e] (excld. **12**)		18 data[e] (excld. **12,16, 17, 22)**
1	1.4	1.43	1.48	1.48	-0.002	1.40
2	1.4	1.33	1.48	1.48	-0.002	1.39
3	0.80	0.68	0.65	0.62	0.024	0.60
4	1.5	1.41	1.46	1.44	0.014	1.35
5	0.18	0.69	0.65	0.73	-0.058	0.63
6	1.1	0.81	0.81	0.63	0.133	0.98
7	1.1	0.97	1.11	1.11	-0.002	1.16
8	0.60	0.65	0.61	1.19	-0.443	0.85
9	0.95	0.97	1.11	1.15	-0.037	1.16
10	1.1	1.01	1.18	1.21	-0.025	1.19
11	1.4	1.22	1.43	1.43	0.003	1.32
12	0.38	1.14	-	-	-	-
13	0.93	0.68	0.65	0.59	0.046	0.57
14	0.83	0.68	0.65	0.60	0.033	0.54
15	1.4	0.99	1.14	1.07	0.058	1.18
16	0.74	0.69	0.71	0.70	0.007	-
17	0.79	0.72	0.73	0.70	0.025	-
18	0.68	0.78	0.77	0.78	-0.013	0.94
19	0.72	0.70	0.67	0.66	0.008	0.71
20	1.3	1.23	1.28	1.24	0.034	1.16
21	0.40	0.76	0.73	0.80	-0.051	0.89
22	0.57	0.69	0.71	0.77	-0.044	-

(a) Compound Nos. are same as in Table 1, reference 1. (b) Data taken from Table 1, reference 1.
(c) Per unit deviation = deviation of 21 dataset Leave-One-out prediction from those predicted in training cycle/span of original data. (d) Results calculated by using eq. 5, reference 1.
(e) Predicted data rounded off to two decimal places are given.

Almost identical curves (with numerical tolerance) were obtained in trainings with random initial weights and also with initial weights to match the statistically fitted curve[1]. This minimises the possibility of a local functional minimum around the statistical fitting. Whether the ANN fitted curve could be the global minimum was explored further by an extensive study using a large number of cases with random initial weights. No new minima were found. The ANN fitted curve (Figure 1), showing a weak minimum between 1.3 - 1.4 logP$_{mw}$ values, followed the experimental data more closely compared to the statistically fitted curve[1].

Our investigation shows that ANN gives more statistically significant result with no *a priori* assumption of functional form. The nature of fit is qualitatively different from that of the regression method. The ANN fitment shows that the antiplatelet activity is strongly correlated with $\log P_{mw}$ for 95% of the experimental data points (21 out of 22). This suggests that investigations with $\log P_{mw}$ as a physicochemical descriptor should be carried out with more sets of compounds and different classes of biological activities to evaluate its usefulness as input parameter for deriving a pharmacophore model[11]

Figure 1. QSAR for investigated compounds by ANN with 1-3-1 network configuration



References and Notes:

1. Tanaka, A.; Nakamura, K.; Nakanishi, I.; Fuziwara, H. *J. Med. Chem.* **1994**, *37*, 4563.
2. Hirst, J.D.; King, R.D.; Sternberg, M.J.E. *J. Comput.-aided Molecular Design* **1994**, *8*, 405.
3. Hirst, J.D.; King, R.D.; Sternberg, M.J.E. *J. Comput.-aided Molecular Design* **1994**, *8*, 421.
4. Manallack, D.T.; Ellis, D.D.; Livingstone, D.J. *J. Med. Chem.* **1994**, *37*, 3758.
5. Ghoshal, N.; Mukhopadhyay, S.N.; Ghoshal, T.K.; Achari, B. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 329.
6. Ghoshal, N.; Mukhopadhyay, S.N.; Ghoshal, T.K.; Achari, B. *Indian J. Chem.* **1993**, *32B*, 1045.
7. ANN works best when the population of data is large. For small dataset, care should be taken to avoid overfitting and arriving at wrong conclusion.
8. The training cycle (with 21 compounds, leaving the compound **12**, an outlier) was repeated 21 times, each time leaving out one compound data. Each training cycle was followed by a test cycle for predicting the activity of the compound which had been left out in that training cycle. The 1-3-1 network configuration was used. "Per unit deviations" (i.e. normalised as deviation/span of original data} of these predicted data from those originally predicted (taking all 21 compounds during training) were calculated (Table 2).
9. McClelland, J.L.; Rumelhart, D.E. *Exploration in parallel distributed processing*; MIT Press; Cambridge, 1988.
10. Aoyama, T.; Suzuki, Y.; Ichikawa, H. *J. Med. Chem.* **1990**, *33*, 905.
11. Maddalena, D.J.; Johnston, G.A.R. *J. Med. Chem.* **1995**, *38*, 715.

(*Received in Belgium* 18 *October* 1996; *accepted* 3 *March* 1997)