Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 12 (2004) 4009–4015
Bioorganic & Medicinal Chemistry

A QSAR study on inhibitory activities of 1-phenylbenzimidazoles against the platelet-derived growth factor receptor

Chongli Zhong,* Jingtao He, Chunyu Xue and Yajun Li

Department of Chem. Eng., Key Lab for Nanomaterials of Ministry of Education, PO Box 100, Beijing University of Chemical Technology, Beijing 100029, China

Received 6 May 2004; revised 31 May 2004; accepted 1 June 2004

Abstract—A quantitative structure–activity relationship (QSAR) study on inhibitory activities of 1-phenylbenzimidazoles against the platelet-derived growth factor receptor (PDGFR) was carried out in this work, and a QSAR model was developed. It gives an r^2 of 0.78 for the training set of 55 active compounds, and an r^2 of 0.75 for the test set of 24 active compounds. The new model was further applied to predict inhibitory activities of additional 44 inactive compounds, and very good agreement with experimental observations was obtained. The new model requires only variable connectivity indices and two position indices as input parameters, which is simple and easy to apply. The new model is useful for developing new anticancer drugs, which also demonstrates that the recently developed variable connectivity indices are very useful structural descriptors in the QSAR studies in the fields of pharmaceutics and biochemistry.

1. Introduction

The platelet-derived growth factor (PDGF) plays a vital role as a regulator of cell growth,[1,2] therefore, studies on inhibitors of PDGF receptor (PDGFR) have received great interest. Such inhibitors are also of interest as potential anticancer drugs.[3,4] Among the various kinds of compounds reported as selective inhibitors of PDGFR,[3–6] 1-phenylbenzimidazoles seem to be a new class of promising PDGFR inhibitors.[3,4] Although a number of structure–activity relationship studies involving 1-phenylbenzimidazoles have been published,[7–11] quantitative structure–activity relationship (QSAR) models are less than satisfying for this system.[12] Recently, Shen et al.[12] proposed several QSAR models for estimation of inhibitory activities of 1-phenylbenzimidazoles as PDGFR inhibitors using electronic parameters as structural descriptors, however, they only considered the 1-phenylbenzimidazoles with substituents only on benzimidazole ring. The other two groups of 1-phenylbenzimidazoles were not included, that is, those with substituents only on phenyl ring and those with substituents on both phenyl and benzimidazole rings. Therefore, their models may have limited applications, and thus a general model applicable to all the 1-phenylbenzimidazoles is necessary.

Although various structural descriptors have been developed, connectivity indices, a kind of topological indices, have been widely used in the QSAR studies due to their simplicity and the richness in information about the molecules.[13,14] Some predictive correlations based on them have been developed successfully in our previous work,[15–20] however, we also found that they lack flexibility and cannot work well in the QSAR studies for some properties.[21,22] On the other hand, a kind of modified connectivity indices, the so-called variable connectivity indices, were developed recently,[23] which, by introducing weights into the original connectivity indices, make the connectivity indices flexible and can distinguish the local environment of nonhydrogen atoms in different environment. Our previous work shows that they work quite well in the QSAR studies,[21,22] particularly in the fields of pharmaceutics and biochemistry, encouraging us to try to develop a QSAR model for inhibitory activities of 1-phenylbenzimidazoles as inhibitors of PDGFR based on them. Therefore, in this work we will not only propose a general QSAR model for estimation of inhibitory activities 1-phenylbenzimidazoles as inhibitors of PDGFR, but will also

Keywords: QSAR; Connectivity index; The platelet-derived growth factor receptor.
* Corresponding author. Tel./fax: +86-10-64419862; e-mail addresses: zhongcl@mail.buct.edu.cn; zhongcl88@yahoo.com

0968-0896/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2004.06.002

demonstrate the usefulness of variable connectivity indices in the QSAR studies in the fields of pharmaceutics and biochemistry, which has not been well recognized yet.

2. Variable connectivity index

The general expression for the *m*th-order connectivity index is as follows:

$$^{m}\chi_k = \sum_{j=1}^{n_m} \prod_{i=1}^{m+1} (\delta_i)_j^{-0.5} \quad (1)$$

where m is the order of the connectivity index, k denotes a contiguous path type of fragment, which is divided into paths (P), clusters (C), path/clusters (PC), and chains (cycles) (CH). n_m is the number of the relevant paths, and δ_i is the simple atomic connectivity index, equal to the number of nonhydrogen atoms to which the *i*th nonhydrogen atom is bonded.

Obviously, the *m*th-order connectivity index, $^{m}\chi_k$, can be calculated easily once the molecular structure of a compound is known, however, it lacks the capability to distinguish the local environment of a nonhydrogen atom in details. Although the valence connectivity index, $^{m}\chi_k^{v}$, works better in this aspect, it still cannot reflect the different role of a nonhydrogen atom in different environment. As a result, a new kind of connectivity index was introduced,[23] that is the so-called variable connectivity index, which can be calculated by the following expression:

$$^{m}\chi_k^f = \sum_{j=1}^{n_m} \prod_{i=1}^{m+1} (\delta_i^f)_j^{-0.5} \quad (2)$$

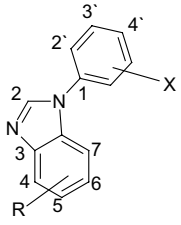
where the variable atomic connectivity index, δ_i^f, is related to the simple atomic connectivity index, δ_i, by

$$\delta_i^f = \delta_i + x_i \quad (3)$$

where x_i is a variable for a nonhydrogen atom, and the numerical value for the variable needs to be selected so to minimize the standard error for a regression. Therefore, x_i is a weight variable, which can increase or decrease the contribution of the atom to the property concerned. As a result, the numerical value of x_i for an atom may vary largely for different properties. The main advantage of the variable connectivity index is that it gives flexibility to connectivity index as structural descriptor, and thus can increase correlative accuracy with a simple expression. The numerical values of x_i should be regressed by fitting experimental data when the QSAR model is developed, and they may change largely when another property is concerned. On the other hand, for a certain property once the numerical values of the weights are determined, the variable connectivity indices can be calculated as simple as the traditional connectivity indices. Therefore, the correlations based on the variable connectivity index have the same predictive capability as those based the traditional connectivity index.

Our previous work[21,22] has shown that variable connectivity indices work quite well when the traditional connectivity indices fail to give satisfactory results. In this work, we will us them to develop a QSAR model for inhibitory activities of 1-phenylbenzimidazoles as

Table 1. Variable connectivity indices and position indices of the compounds used in this work



No	R	X	$^{1}\chi^{f}$	$^{3}\chi_c^{f}$	$B_{2,2',4,7}$	B_5
1	H	H	3.796	0.215	0	0
2	4-OMe	H	4.736	0.285	1	0
3	4-OH	H	6.123	0.643	1	0
4	5-Me	H	3.998	0.263	0	1
5	5-OMe	H	4.734	0.296	0	1
6	5-OH	H	6.122	0.690	0	1
7	5-Cl	H	4.077	0.279	0	1
8	5-COOH	H	6.692	1.200	0	1
9	5-COOMe	H	5.305	0.411	0	1
10	5-$CONH_2$	H	6.692	1.200	0	1
11	5-NO_2	H	7.879	1.599	0	1
12	5-COMe	H	4.568	0.344	0	1
13	5-CHO	H	4.386	0.252	0	1
14	5-OC_3H_7	H	5.199	0.296	0	1
15	5-OC_2H_5	H	4.948	0.296	0	1
16	5-$OCH(Me)_2$	H	5.124	0.444	0	1

Table 1 (*continued*)

No	R	X	$^{1}\chi^{f}$	$^{3}\chi_{c}^{f}$	$B_{2,2',4,7}$	B_5
17	5-OC_4H_9	H	5.450	0.296	0	1
18	5-$OCH_2CH{=}CH_2$	H	5.107	0.296	0	1
19	5-$O(CH_2)_4OH$	H	8.305	0.296	0	1
20	5-OCH_2(oxiranyl)	H	6.042	0.407	0	1
21	5-$OCH_2CH(OH)CH_2OH$	H	10.588	0.946	0	1
22	5-$O(CH_2)_2OH$	H	7.803	0.296	0	1
23	5-$O(CH_2)_2N(Me)_2$	H	6.074	0.414	0	1
24	5-$O(CH_2)_3N(Me)_2$	H	6.325	0.414	0	1
25	5-$O(CH_2)_4N(Me)_2$	H	6.576	0.414	0	1
26	5-$O(CH_2)_2$Nmorph	H	7.455	0.385	0	1
27	5-$O(CH_2)_3$Nmorph	H	7.706	0.385	0	1
28	5-$O(CH_2)_4$Nmorph	H	7.957	0.385	0	1
29	5-SH	H	9.247	1.317	0	1
30	5-SMe	H	6.654	0.456	0	1
31	5-$OCSN(Me)_2$	H	6.864	1.221	0	1
32	6-Me	H	3.998	0.263	0	0
33	6-OMe	H	4.734	0.296	0	0
34	6-OH	H	6.122	0.690	0	0
35	6-Cl	H	4.077	0.279	0	0
36	6-COOH	H	6.692	1.200	0	0
37	6-COOMe	H	5.305	0.411	0	0
38	6-$CONH_2$	H	6.692	1.200	0	0
39	6-NO_2	H	7.879	1.599	0	0
40	6-NH_2	H	6.122	0.690	0	0
41	7-OMe	H	4.736	0.285	1	0
42	4,5-DiOH	H	8.451	1.038	1	1
43	4-OH,5-OMe	H	7.063	0.679	1	1
44	4-$CH_2CH(Me)O$-5	H	5.262	0.447	1	1
45	5,6-DiOH	H	8.449	1.081	0	1
46	5,6-DiMe	H	4.201	0.302	0	1
47	5,6-OCH_2O	H	5.521	0.377	0	1
48	5-OMe,6-Me	H	4.938	0.333	0	1
49	5-OH,6-Me	H	6.325	0.691	0	1
50	5-OMe,6-COOH	H	7.632	1.271	0	1
51	5-OH,6-COOH	H	9.019	1.630	0	1
52	5-OMe,6-COOMe	H	6.245	0.482	0	1
53	5-OMe,6-CH_2OH	H	7.800	0.327	0	1
54	5-OMe,6-CHO	H	5.326	0.323	0	1
55	5-NH_2	H	6.122	0.690	0	1
56	5-Aza	H	4.023	0.215	0	0
57	7-Aza	H	4.013	0.245	0	0
58	H	3′-Me	3.998	0.263	0	0
59	H	3′-OMe	4.734	0.296	0	0
60	H	3′-OH	6.122	0.690	0	0
61	H	3′-Cl	4.077	0.279	0	0
62	H	3′-NO_2	7.879	1.599	0	0
63	H	3′-NH_2	6.122	0.690	0	0
64	H	3′-COMe	4.568	0.344	0	0
65	H	3′-CHO	4.386	0.252	0	0
66	H	4′-OMe	4.734	0.296	0	0
67	H	4′-OH	6.122	0.690	0	0
68	H	4′-Cl	4.077	0.279	0	0
69	H	4′-COOMe	5.305	0.411	0	0
70	H	4′-$CONH_2$	6.692	1.200	0	0
71	H	4′-NO_2	7.879	1.599	0	0
72	H	4′-NH_2	6.122	0.690	0	0
73	H	4′-COMe	4.568	0.344	0	0
74	H	4′-CHO	4.386	0.252	0	0
75	H	4′-CN	4.215	0.249	0	0
76	H	4′-Aza	4.023	0.215	0	0
77	5-OMe	2′-Thienyl	7.910	0.565	1	1
78	5-OMe	3′-Thienyl	7.908	0.570	0	1
79	5-OMe	4′-NH_2	7.060	0.770	0	1
80	4-COOH	H	6.693	1.193	1	0
81	4-COOMe	H	5.306	0.404	1	0

(*continued on next page*)

Table 1 (*continued*)

No	R	X	$^1\chi^f$	$^3\chi_c^f$	$B_{2,2',4,7}$	B_5
82	4-$CONH_2$	H	6.693	1.193	1	0
83	4-NO_2	H	7.880	1.590	1	0
84	4-NH_2	H	6.123	0.643	1	0
85	7-Me	H	3.999	0.254	1	0
86	7-OH	H	6.123	0.643	1	0
87	7-Cl	H	4.079	0.269	1	0
88	7-COOH	H	6.693	1.193	1	0
89	7-COOMe	H	5.306	0.404	1	0
90	7-$CONH_2$	H	6.693	1.193	1	0
91	7-NO_2	H	7.880	1.590	1	0
92	7-NH_2	H	6.123	0.643	1	0
93	4-OMe,5-OH	H	7.063	0.711	1	1
94	4,5-DiOMe	H	5.676	0.352	1	1
95	4-Br,5-OH	H	10.339	1.354	1	1
96	4-Br,5-$OCH_2CH{=}CH_2$	H	9.325	0.995	1	1
97	4-$CH_2CH{=}CH_2$,5-OH	H	6.744	0.678	1	1
98	5-$S(CH_2)_3$Nmorph	H	9.474	0.544	0	1
99	4-Me	H	3.999	0.254	1	0
100	4-Cl	H	4.079	0.269	1	0
101	2-Me	H	3.978	0.328	1	0
102	2-OH	H	6.102	1.374	1	0
103	2-NH_2	H	6.102	1.374	1	0
104	H	2′-Me	3.999	0.254	1	0
105	H	2′-OMe	4.736	0.284	1	0
106	H	2′-OH	6.123	0.643	1	0
107	H	2′-Cl	4.079	0.268	1	0
108	H	2′-COOH	6.693	1.192	1	0
109	H	2′-COOEt	5.520	0.403	1	0
110	H	2′-$CONH_2$	6.693	1.192	1	0
111	H	2′-NO_2	7.880	1.590	1	0
112	H	2′-NH_2	6.123	0.643	1	0
113	H	2′-COMe	4.569	0.336	1	0
114	H	2′-CHO	4.387	0.244	1	0
115	H	2′-CN	4.217	0.241	1	0
116	H	3′-COOH	6.692	1.200	0	0
117	H	3′-COOEt	5.518	0.411	0	0
118	H	3′-$CONH_2$	6.692	1.200	0	0
119	H	3′-CN	4.215	0.249	0	0
120	H	4′-Me	3.998	0.263	0	0
121	H	4′-COOH	6.692	1.200	0	0
122	H	2′-Aza	4.013	0.248	0	0
123	H	3′-Aza	4.023	0.215	0	0

inhibitors of PDGFR. The variable connectivity indices used in this work are shown in Table 1. A review on the development of the connectivity index was recently published by Randić.[24]

3. Results and discussion

Palmer and co-workers[3,4] performed structure–activity relationship (SAR) studies on inhibitory activities of more than 100 1-phenylbenzimidazole derivatives as inhibitors of PDGFR, which can serve as a database to develop a QSAR model. The 1-phenylbenzimidazoles studied by them are shown in Table 1, and their inhibitory activity data are listed in Tables 2 and 3. The compound with $IC_{50} \leqslant 50\,\mu M$ (definition of IC_{50} is given in the footnote of Table 2) is thought to be an active compound, otherwise, it is thought as an inactive compound. The first 79 compounds in Table 1 are active ones, while the others are inactive compounds. Obviously, there are three groups of 1-phenylbenzimidazole derivatives, that is, those with substituents only on either benzimidazole ring or phenyl ring and those with substituents on both phenyl and benzimidazole rings.

The 79 active compounds were randomly divided into a training set of 55 compounds and a test set of 24 compounds. The training set compounds were used to develop a QSAR model, and the test set compounds were used to validate the reliability and the predictive ability of the model. Based on the analysis of the experimental data for the training set compounds, we got the conclusion that the traditional connectivity indices are not suitable structural descriptors for this property, and thus the variable connectivity indices were adopted and the following model was proposed:

$$\begin{aligned}\mathrm{Log}(1/\mathrm{IC}_{50}) &= 7.0050 - 1.3877 \cdot {}^3\chi_c^f - 11.3749/{}^1\chi^f \\ &\quad + \exp(0.4970 \cdot B_5 - 1.98 \cdot B_{2,2',4,7}) \\ n = 55 \quad s = 0.363 \quad F = 43.03 \quad r^2 &= 0.78 \end{aligned} \tag{4}$$

Table 2. Results of the training set compounds

No	Log(1/IC$_{50}$)[a]		
	Exp.	Shen et al.[b]	This work
1	5.030	4.933	4.709
3	4.854	4.616	4.392
5	6.367	5.789	5.834
7	5.398	5.471	5.471
8	5.030	5.477	5.283
9	6.081	5.700	5.933
10	4.796	5.302	5.283
12	6.066	5.548	5.680
14	6.602	6.162	6.049
16	5.509	5.987	5.812
17	5.886	6.343	6.150
18	6.215	6.149	6.010
20	6.497	6.112	6.200
21	6.509	6.180	6.261
22	6.187	5.821	6.779
24	6.824	6.511	6.275
25	6.796	6.695	6.343
27	6.770	6.211	6.638
28	6.569	6.397	6.685
29	5.482	5.411	5.590
30	6.131	5.716	6.306
31	5.337	5.861	5.297
33	5.194	5.207	5.191
34	5.678	4.967	5.190
35	5.268	4.829	4.828
37	4.886	5.037	5.290
38	4.602	4.615	4.640
39	4.301	4.513	4.342
41	4.432	4.454	4.346
43	5.149	5.129	4.679
44	4.538	5.140	4.450
46	5.921	5.701	5.521
47	5.658	5.137	6.064
49	5.602	5.194	5.890
50	4.678	5.343	5.394
52	6.061	5.603	6.157
53	6.432	5.507	6.736
54	6.000	5.375	6.064
55	5.569	—	5.833
57	4.553	—	4.830
59	4.602	—	5.191
60	5.420	—	5.190
62	4.796	—	4.342
63	5.444	—	5.190
65	5.168	—	5.061
68	4.300	—	4.828
69	5.143	—	5.290
70	4.638	—	4.640
71	4.523	—	4.342
72	5.252	—	5.190
73	4.620	—	5.037
74	4.886	—	5.061
76	4.921	—	4.879
77	5.602	—	5.009
79	6.553	—	5.968
Average error		0.356	0.290

[a] IC$_{50}$: concentration of drug (M) to inhibit the phosphorylation of a random glutamate/tyrosine (4:1) copolymer by PDGFR proteins.
[b] Taken from Ref. 12.

where $^{3}\chi_{c}^{f}$ is third-order variable cluster connectivity index, $^{1}\chi^{f}$ is first-order variable connectivity index. B_5 is a position index, which is assigned the value of 1

Table 3. Results of the test set compounds

No	Log(1/IC$_{50}$)		
	Exp.	Shen et al.[a]	This work
Active compounds			
2	4.301	4.422	4.346
4	5.357	5.551	5.438
6	6.357	5.553	5.833
11	4.796	5.135	4.985
13	6.367	5.523	5.704
15	6.620	5.978	5.938
19	6.347	6.442	6.867
23	5.824	6.332	6.200
26	6.137	6.029	6.588
32	4.398	4.928	4.795
36	4.301	4.810	4.640
40	4.638	4.746	5.190
42	4.602	4.873	4.445
45	5.638	5.239	5.802
48	6.000	5.436	5.883
51	5.367	5.103	5.125
56	5.000	—	4.879
58	4.553	—	4.795
61	4.328	—	4.828
64	4.721	—	5.037
66	4.886	—	5.191
67	5.745	—	5.190
75	4.796	—	4.961
78	6.155	—	6.418
Average error		0.394	0.332
Inactive compounds			
80	<4.3	4.405	3.788
81	<4.3	4.210	4.439
82	<4.3	4.004	3.788
83	<4.3	3.912	3.493
84	<4.3	4.390	4.392
85	<4.3	4.568	3.946
86	<4.3	4.625	4.392
87	<4.3	4.668	3.981
88	<4.3	4.181	3.788
89	<4.3	3.892	4.439
90	<4.3	4.078	3.788
91	<4.3	4.097	3.493
92	<4.3	4.408	4.392
93	<4.3	4.664	4.635
94	<4.3	4.906	4.739
95	<4.3	4.075	4.253
96	<4.3	4.686	4.631
97	<4.3	4.475	4.605
98	<4.3	6.015	6.692
99	<4.3	4.528	3.946
100	<4.3	4.487	3.981
101	<4.3	—	3.829
102	<4.3	—	3.372
103	<4.3	—	3.372
104	<4.3	—	3.947
105	<4.3	—	4.347
106	<4.3	—	4.393
107	<4.3	—	3.982
108	<4.3	—	3.789
109	<4.3	—	4.523
110	<4.3	—	3.789
111	<4.3	—	3.493
112	<4.3	—	4.393
113	<4.3	—	4.187
114	<4.3	—	4.212

(*continued on next page*)

Table 3 (*continued*)

No	Log(1/IC$_{50}$)		
	Exp.	Shen et al.[a]	This work
115	<4.3	—	4.111
116	<4.3	—	4.640
117	<4.3	—	5.373
118	<4.3	—	4.640
119	<4.3	—	4.961
120	<4.3	—	4.795
121	<4.3	—	4.640
122	<4.3	—	4.827
123	<4.3	—	4.879

[a] Taken from Ref. 12.

when a substituting group is present at 5-position, and a value of 0 is assigned otherwise. $B_{2,2',4,7}$ is another position index, which is assigned to 1 if a substituting group presents at the 2-, 2′-, 4-, or 7-position, and a value of 0 otherwise. The two position indices were introduced by referring to the experimental observations of Palmer and co-workers[3,4] on the effects of the position of substituents on inhibitory activity. The optimal weights x_i for the nonhydrogen atoms were also obtained by fitting the training set data when the QSAR model was developed, which are $x_C = 1.98$, $x_N = -2.97$, $x_O = -4.97$, $x_{Cl} = 0.89$, $x_{Br} = -0.25$, and $x_S = -0.55$ for the atoms of carbon, nitrogen, oxygen, chlorine, bromine, and sulfur, respectively. Positive values of x_i were obtained for carbon and chlorine atoms, indicating that their contribution is decreased when inhibitory activities are concerned. The other nonhydrogen atoms, on the other hand, showed negative x_i values, demonstrating that their contribution was increased when inhibitory activities are concerned. The above results illustrate clearly the flexibility of variable connectivity indices as structural descriptors.

The calculated results with the new model are shown in Table 2 and depicted in Figure 1. In order to evaluate the new model in more details, the statistical parameters of the new model were calculated, including the correlation coefficient (r^2), standard error (s), and Fisher's

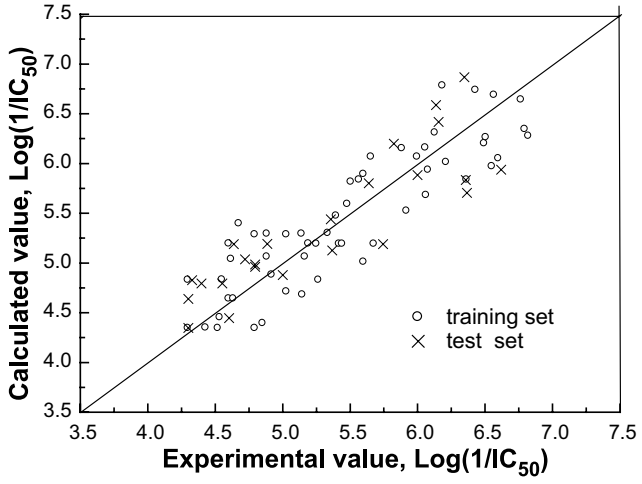


Figure 1. Calculated versus experimental Log(1/IC$_{50}$) for 79 active compounds using the new model.

F-value (F), as shown below the new model. From the statistical parameters given above and the results shown in Table 2 and Figure 1, it is evident that the new model works quite well for the estimation of inhibitory activities of 1-phenylbenzimidazoles as inhibitors of PDGFR. Shen et al.[12] proposed four correlations for the 1-phenylbenzimidazoles with substituents only on benzimidazole ring (75 compounds). To make a comparison with their work, the calculated results obtained by their best correlation are also listed in Table 2, in which the results for those 1-phenylbenzimidazoles with substituents only on phenyl ring and those on both phenyl and benzimidazole rings are missed because these two groups of 1-phenylbenzimidazoles were not included in the development of their models. Obviously, the new model works better than the best correlation obtained by Shen et al.,[12] and the new model can be applied to all the 1-phenylbenzimidazoles with wider application range. Furthermore, compared with the correlations of Shen et al., the new model requires variable connectivity indices as input parameters, which can be calculated easily as long as the molecular structure of the compound concerned is known after the weights have been determined.

The third-order variable cluster connectivity index, $^3\chi_c^f$, encodes the number of clusters and the nature of the nonhydrogen atoms constituting the clusters. In this work, since the parent compound is 1-phenylbenzimidazole for all the compounds concerned, $^3\chi_c^f$ denotes the number of substituents on phenyl and benzimidazole rings as well as the nature of the nonhydrogen atoms bound to them. The first-order variable connectivity index, $^1\chi^f$, can represent the size of the compound concerned, therefore, it can reflect the size of the substituents in this work. The positive coefficient of B_5 indicates that 5-position is an active position and substituent at 5-position can enhance inhibitory activity, while the negative coefficient of $B_{2,2',4,7}$ illustrates that 2-, 2′-, 4-, and 7-positions are inactive positions and thus substituents presented at them induce the inhibitory activity. Therefore, the new model, Eq. 4, demonstrates that the number of substituents, the nature of the atoms bound to the parent compound, the nature of the substituents, and the substituting position affect the inhibitory activity of a 1-phenylbenzimidazole derivative, which is useful for new drug design and development.

To validate the reliability and the predictive ability of the new model, it was used to predict inhibitory activities of another 24 active compounds. The predicted results are shown in Table 3 and depicted in Figure 1. The best results obtained by Shen et al.[12] are also given in Table 3 for comparison. The statistical parameters obtained by the new model for the test set compounds are $s = 0.427$, $F = 14.41$, and $r^2 = 0.75$. Evidently, the new model is reliable and gives good predictions for the 24 active compounds not included in the development of the new model. The results listed in Table 3 also show that the new model works better than the existing correlations.

In the work of Palmer and co-workers[3,4] more than forty 1-phenylbenzimidazoles were determined as inactive

compounds with $IC_{50} > 50\,\mu M$. It is a more serious test of the new model to use it to predict inhibitory activities for these inactive compounds since they were not included when the new model was developed. Since the definite values of activity for these compounds were not given by Palmer and co-workers,[3,4] a quantitative comparison is not possible. However, we can compare the calculated values of $Log(1/IC_{50})$ with 4.3 (corresponding to $IC_{50} = 50\,\mu M$) to determine if the compound is active or not. The predicted results, as shown in Table 3, show that most compounds are inactive with $Log(1/IC_{50})$ smaller than or around 4.3. Only a few compounds with $Log(1/IC_{50})$ much larger than 4.3, namely, compound 98 and several compounds with substituents at the 3′-position of the phenyl ring. Considering the experimental uncertainty (±15%), a predicted value of smaller than 5.0 is acceptable for inactive compounds. From the results shown in Table 3, only two compounds with $Log(1/IC_{50})$ larger than 5.0. The value for compound 117 is 5.371, which is also acceptable. However, compound 98 gives a value rather larger than 4.3, which seems to be an outlier. In the work of Shen et al.,[12] four correlations were proposed but compound 98 always performed as an outlier no matter it was placed in the training or test sets. In this work, we got the same conclusion. Therefore, compound 98 is either an outlier, which does not follow the general SAR for the 1-phenylbenzimidazoles or the experimental data for it is not so reliable.

The above calculation results, both correlations and predictions, show that the new model is reliable with good predictive accuracy. The effects of the structure on the inhibitory activity can be reflected and reproduced well with it. Compared with the existing models, it requires only variable connectivity indices and two position indices, which is easy to apply.

4. Conclusion

A QSAR study on inhibitory activities of 1-phenylbenzimidazoles against PDGFR was carried out in this work, and a QSAR model was developed using variable connectivity indices and two position indices as input parameters. The new model is simple, reliable and easy to apply, which is useful for new drug design and development. Compared with the existing models, the input parameters required by the new model are easier to obtain. Moreover, it is a general model applicable to all the 1-phenylbenzimidazole derivatives. The present work also shows the usefulness of the variable connectivity index in the fields of pharmaceutics and biochemistry, which can work quite well for those properties/activities where the traditional connectivity index fails. Compared with other structural descriptors, it is easy to calculate and flexible to adjust, resulting in simple correlations with good predictive accuracy. As a result, the present work demonstrates that the variable connectivity index should find more applications in the fields of pharmaceutics, biochemistry, etc.

Acknowledgements

The financial support of TRAPOYT and the Trans—Century Training Programme Foundation for the Talents by the MOE, P.R.C. is greatly appreciated.

References and notes

1. Claesson-Welsl, L. *Cytokines* **1993**, *5*, 31.
2. Meyer-Ingold, W.; Eichner, W. *Cell Biol. Int.* **1995**, *19*, 389.
3. Palmer, B. D.; Smaill, J. B.; Boyd, M.; Boschelli, D. H.; Doherty, A. M.; Hamby, J. M.; Khatana, S. S.; Kramer, J. B.; Kraker, A. J.; Panek, R. L.; Lu, G. H.; Dahring, T. K.; Winters, R. T.; Showalter, H. D. H.; Denny, W. A. *J. Med. Chem.* **1998**, *41*, 5457.
4. Palmer, B. D.; Kraker, A. J.; Hartl, B. G.; Panopoulos, A. D.; Panek, R. L.; Batley, B. L.; Lu, G. H.; Trumpp-Kallmeyer, S.; Showalter, H. D. H.; Denny, W. A. *J. Med. Chem.* **1999**, *42*, 2373.
5. Dolle, R. E.; Dunn, J. A.; Bobko, M.; Singh, B.; Kuster, J. E.; Baizman, E.; Harris, A. L.; Sawutz, S. G.; Miller, D.; Wang, S.; Faltynek, C. R.; Xie, W.; Sarup, J.; Bode, C. E.; Pagani, E. D.; Silver, P. J. *J. Med. Chem.* **1994**, *37*, 2627.
6. Maguire, M. P.; Sheets, K. R.; McVety, K.; Spada, A. P.; Ziberstein, A. A. *J. Med. Chem.* **1994**, *37*, 29.
7. Oblak, M.; Randić, M.; Solmajer, T. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 994.
8. Pierre, D.; Michel, L.; David, S. G. *J. Med. Chem.* **2000**, *43*, 4098.
9. Zhu, L. L.; Hou, T. J.; Chen, L. R.; Xu, X. J. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1032.
10. Kurup, A.; Garg, R.; Hansch, C. *Chem. Rev.* **2001**, *101*, 2573.
11. Naumann, T.; Matter, T. *J. Med. Chem.* **2002**, *45*, 2366.
12. Shen, Q.; Lu, Q.-Z.; Jiang, J.-H.; Shen, G.-L.; Yu, R.-Q. *Eur. J. Pharm. Sci.* **2003**, *20*, 63.
13. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic: New York, 1976.
14. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure–Activity Analysis*; Wiley: New York, 1986.
15. Zhong, C.; Hu, Q. *J. Pharm. Sci.* **2003**, *92*, 2284.
16. Yang, C.; Zhong, C. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1998.
17. He, J.; Zhong, C. *Eur. Polym. J.* **2003**, *39*, 1297.
18. Yang, C.; Zhao, L.; Zhong, C. *Int. J. Thermophys.* **2003**, *24*, 1651.
19. Zhong, C.; Yang, C. *J. Polym. Sci.: Polym. Phys. Ed.* **2002**, *40*, 401.
20. Zhong, C.; Yang, C.; Li, Q. *Ind. Eng. Chem. Res.* **2002**, *41*, 2826.
21. Liu, D.; Zhong, C. *Polym. J.* **2002**, *34*, 954.
22. Li, Y.; Hu, Q.; Zhong, C. *Ind. Eng. Chem. Res.*, in press.
23. Randić, M. *J. Comput. Chem.* **1991**, *12*, 970.
24. Randić, M. *J. Mol. Graph. Model* **2001**, *20*, 19.