A New Approach for the Tissue−Blood Partition Coefficients of Neutral and Ionized Compounds

Huabei Zhang*

Chemistry Department, Beijing Normal University, 100875 Beijing, P.R. China

Received September 20, 2004

A nonlinear model equation based on tissue composition (a content of lipids, proteins, and water) for the tissue−blood partition coefficients of compounds was further developed in this paper to account for the neutral and ionized forms of the compounds. The nonlinear model equation may be considered as the expressive form of the Hansch equation in nonlinear spaces (or multiphase system). Based on this model our nonlinear regression analysis for neutral and ionized compounds partitioning into kidney, brain, muscle, lung, liver, heart, and fat resulted in equations with high fitting power (training set: n=201, r=0.905, s=0.291, Q=0.890) and strong predictive power (test set: n=64, r=0.906, s=0.247). These results showed that the equilibrium distribution of a compound in a variety of tissues is essentially the equilibrium distribution of the compound in tissue (chemical) compositions. Neutral and ionic forms of a compound as well as in different tissue (chemical) compositions have a different mechanism of action in vivo.

INTRODUCTION

It was pointed out that in addition to unproved efficiency, toxicity, and adverse reactions, inadequate pharmacokinetic properties resulted in about 40% of withdrawals of drug leads from further development.[1,2] The ability of synthesizing a large number of diverse combinatorial chemistry libraries and the recent advances in automated high throughput screening techniques have reduced significantly the time required for identifying a lead compound with desired potency against a biological target. However, the bottleneck of faster drug discovery is shifting toward optimizing lead candidates so as to result in adequate ADME (absorption, distribution, metabolism, and excretion) properties.[3] Obviously evaluation of ADME properties in the early stages of drug discovery will save both time and expense.

Physiologically based pharmacokinetic (PBPK) models can be used as a tool in drug development. For PBPK modeling, the tissue−blood partition coefficients of the drug in various organs and tissues need to be known.[4] For example, to be effective as therapeutic agents, centrally acting drugs must cross the blood-brain (BBB). Conversely, to be devoid of unwanted central nervous system (CNS) effects, some drugs must show limited ability to cross the BBB. However, the experimental determination of the tissue−blood partition coefficients is difficult, expensive, and time-consuming, since it involves the direct measurement of the drug concentration in the tissue−blood of laboratory animals and obviously requires the synthesis of the compounds, often in a radiolabeled form. A reliable and accurate computational model for predicting tissue−blood partition coefficients will therefore have a significant impact on drug research and development.

Therefore it is important to predict the tissue−blood partition coefficient of compounds from physicochemical parameters directly. The use of the Hansch equation since its advent in 1962[5] has become increasingly helpful in understanding many aspects of chemical−biological interactions in drug and pesticide research as well as many areas of toxicology.[6,7] To obtain better results of QSAR the efforts of two aspects have been made: (1) Some new descriptors were developed and used,[8−13] such as Raevsky's[14] hydrogen bond descriptors, Abraham's,[15,16] solvatochromic descriptors, and polar surface area (PSA),[17,18] etc. It has been proven that PSA is a simple and significant descriptor for the prediction of logBB. (2) Some new models were developed, such as Balaz's nonlinear model[19] and Poulin's equation.[20,21] Poulin et al.[20,21] proposed the tissue−plasma partition coefficients of nonexcretory can be predicted from muscle−plasma partition coefficients and oil−water partition coefficients using a nonlinear equation. Obviously this method requires the synthesis of the compounds.

Recently we developed a nonlinear model equation based on tissue composition for the tissue−blood partition coefficients of neutral compounds. Based on this model our nonlinear regression analysis for neutral compounds partitioning into kidney, brain, muscle, lung, liver, heart, and fat resulted in equations with high fitting power (training set: n=166, r=0.922, s=0.260, Q=0.912) and strong predictive power (test set: n=49, r=0.922, s=0.246, Q=0.914).[22]

In the present report we extend the nonlinear model to the ionic form, and the new model bears a strong predictive power for tissue (kidney,brain, muscle, lung, liver, fat and heart)−blood partition coefficients of neutral and ionized compounds.

MATERIALS AND METHODS

Tissue−Blood Partition Coefficient. In our research the following two experimental data sets were used:

(1) The tissue−blood partition coefficients PC_t of 36 organic chemicals for human fat, liver, brain, kidney, muscle, lung, and heart were taken from ref 19 and are listed in Table 1(from no. 1 to no. 36 in Table 1). (2) The tissue−blood partition coefficients PC_t of 10 basic compounds for rabbit

* Corresponding author phone: +86-10-58805194; fax: +86-10-58800567; e-mail: hbzhang@bnu.edu.cn.

10.1021/ci049718e CCC: $30.25
Published on Web 12/24/2004

Table 1. Logarithm of Experimental and Calculated Tissue−Blood Partition Coefficients ($\log PC_t$) and pK_a Values

		kidney		brain		muscle		lung		liver		Heart		fat		
no.	compound	exp.	eq 5	exp.	eq 5	exp.	eq 5	exp.	eq 5	exp.	eq 5	exp.	eq 5	exp.	eq 5	pK_a
1	2,2-dimethylbutane	0.73	0.737	1.03	1.016	0.58	0.439	0.36	0.268	1.13	0.862	0.28	0.998	2.40	1.862	
2	pentane	0.20	0.566	0.76	0.826	0.27	0.298	0.12	0.156	0.74	0.682	−0.28	0.812	2.02	1.659	
3	2-Me-pentane	0.69	0.737	0.97	1.016	0.85	0.439	0.29	0.268	1.04	0.862	0.53	0.999	2.33	1.863	
4	3-Me-pentane	0.76	0.737	1.01	1.016	0.95	0.439	0.32	0.268	1.06	0.862	0.65	0.999	2.38	1.863	
5	hexane	0.57	0.737	0.80	1.016	0.80	0.439	0.10	0.268	0.81	0.862	0.54	0.999	2.11	1.863	
6	3-Me-hexane[a]	0.75	0.918	0.89	1.211	0.93	0.595		0.398	0.93	1.049	0.61	1.191	2.33	2.067	
7	heptane	0.67	0.918	0.80	1.211	0.80	0.595	0.12	0.397	0.76	1.049	0.51	1.191	2.31	2.067	
8	cyclopropane	0.15	0.202	0.08	0.392	0.12	0.028		−0.049	0.19	0.287		0.390	1.45	1.160	
9	Me-cy-pentane[a]	0.74	0.664	0.93	0.936	0.76	0.378	0.30	0.219	0.96	0.786	0.34	0.920	2.31	1.777	
10	cyclohexane	0.74	0.667	0.93	0.939	0.89	0.380	0.32	0.221	0.93	0.789	0.65	0.923	2.30	1.781	
11	benzene	0.27	0.328	0.45	0.524	0.40	0.139	0.27	0.056	0.56	0.420	0.34	0.527	1.82	1.311	
12	toluene[a]	0.24	0.449	0.55	0.669	0.53	0.231	0.31	0.128	0.67	0.552	0.46	0.668	1.97	1.479	
13	styrene		0.333		0.494	0.00	0.173		0.109	0.43	0.417		0.511	1.70	1.262	
14	CH_2Cl_2	−0.01	0.125	0.00	0.272	−0.10	−0.009	−0.01	−0.063	0.08	0.195	−0.14	0.282	1.15	0.996	
15	$CHCl_3$	0.14	0.242	0.40	0.425	0.18	0.072	0.06	−0.000	0.33	0.328	−0.11	0.428	1.54	1.193	
16	Me-CCl_3[a]	0.31	0.385	0.40	0.602	0.31	0.174	0.15	0.076	0.69	0.485	0.44	0.599	1.88	1.405	
17	CF_3-CH_2Cl	−0.16	0.253	0.08	0.458	0.17	0.063	0.17	−0.023	0.12	0.345		0.454	1.36	1.241	
18	teflurane		0.418	0.26	0.664	0.58	0.176		0.052	0.22	0.527		0.652	1.52	1.483	
19	halothane	0.41	0.480	0.39	0.729	0.54	0.230	0.24	0.101	0.57	0.592		0.718	1.97	1.554	
20	CF_2=CHCl[a]	0.06	0.243	0.10	0.445	0.06	0.056	0.06	−0.028	0.10	0.334		0.441	1.36	1.225	
21	CCl_2=CHCl	0.27	0.337	0.41	0.540	0.37	0.142	0.24	0.055	0.55	0.432	0.32	0.541	1.85	1.332	
22	MeOH	−0.08	−0.190	−0.17	−0.270	−0.09	−0.182	0.03	−0.180		−0.198		−0.206	−0.85	−0.765	
23	EtOH	−0.15	−0.161	−0.23	−0.252	−0.19	−0.155	−0.06	−0.152		−0.167		−0.176	−0.80	−0.683	
24	1-PrOH	−0.14	−0.128	−0.22	−0.227	−0.15	−0.124	−0.07	−0.120		−0.131		−0.140	−0.51	−0.532	
25	2-PrOH	−0.16	−0.126	−0.22	−0.221	−0.16	−0.123	−0.09	−0.120		−0.128		−0.135	−0.60	−0.444	
26	1-BuOH	−0.16	−0.090	−0.21	−0.193	−0.20	−0.089	−0.13	−0.085		−0.089		−0.096	−0.14	−0.333	
27	acetone[a]	−0.13	−0.083	−0.21	−0.102	−0.11	−0.110	−0.09	−0.118		−0.066		−0.045	−0.36	0.261	
28	butanone	−0.07	−0.021	−0.11	−0.014	−0.09	−0.065	−0.09	−0.078		0.007		0.041	0.11	0.472	
29	Et_2O	−0.04	0.057	0.03	0.128	−0.08	−0.024		−0.052	−0.04	0.106	0.0	0.164	0.77	0.757	
30	divinyl ether	−0.17	0.144	0.16	0.284	−0.04	0.014		−0.038	0.05	0.214		0.299	1.20	1.007	
31	flurane[a]	−0.03	0.296	0.16	0.502	0.16	0.102		0.013	0.16	0.390		0.499	1.39	1.290	
32	enflurane	0.28	0.314	0.26	0.523	0.43	0.116	0.30	0.026	0.45	0.410		0.520	1.87	1.314	
33	isoflurane	0.18	0.266	0.32	0.457	0.23	0.086	0.06	0.007	0.47	0.355		0.459	1.69	1.234	
34	methoxyflurane	0.18	0.251	0.29	0.415	0.29	0.095	0.06	0.031	0.40	0.332		0.427	1.80	1.173	
35	sevoflurane	0.52	0.617	0.34	0.891	0.40	0.329	0.34	0.167	0.70	0.739		0.874	1.86	1.732	
36	NO[a]	−0.07	−0.018	0.03	0.100	−0.07	−0.117	0.00	−0.157	−0.04	0.034		0.107	0.37	0.755	
37	pentazocin		1.324	0.708	1.055	0.806	1.317	1.506	1.331		1.352	0.806	1.331	0.398	1.220	8.5
38	nitrazepam		0.464	0.322	0.201	0.230	0.466	0.255	0.480		0.482	0.146	0.454	0.362	−0.088	3.4
39	haloperidol		1.287	0.914	1.118	0.857	1.256	1.728	1.262		1.326	1.155	1.328	1.441	1.609	7.8
40	biperiden[a]		1.365	1.410	1.200	0.929	1.333	2.117	1.338		1.405	1.536	1.408	2.079	1.701	8.8
41	diazepam		0.570	0.505	0.398	0.544	0.548	0.924	0.555		0.601	0.778	0.598	1.086	0.784	3.5
42	promethazine		1.327	1.301	1.094	1.188	1.311	2.180	1.323		1.358	1.544	1.345	2.122	1.416	9.1
43	trihexyphenidyl		1.361	1.326	1.196	1.120	1.330	1.871	1.335		1.401	1.358	1.404	1.883	1.698	8.7
44	chlorpromazine[a]		1.311	0.968	1.081	0.716	1.295	1.806	1.306		1.342	1.146	1.330	1.612	1.411	9.3
45	clotiazepam		0.712	0.505	0.608	0.204	0.668	1.041	0.667		0.756	0.415	0.772	0.771	1.172	3.6
46	clomipramine[a]		1.304	1.025	1.106	0.792	1.281	2.159	1.289		1.339	1.611	1.334	1.936	1.530	8.5

[a] Test set.

fat, brain, muscle, lung, and heart were taken from ref 23, and their pK_a values were taken from the same reference and are listed in Table 1 (from no. 37 to no. 46 in Table 1). Some research has shown that tissue−blood partition coefficients of human, rat, and rabbit are compatible and are often used in regression analysis together.[4,20,21]

It is difficult to obtain more data for further research because most of the data for drugs are proprietary and is not released by the drug manufacturers. However in the present data set there are 265 data points; we think that it is enough to establish and test our model.

Nonlinear Model. The tissue−blood(or plasma) partition coefficient PC_t is define as the ratio of the equilibrium concentrations C of the compound in the tissue and in blood or plasma.

The partition coefficient for a compounds partitioning between tissue (t) and blood or plasma (b) is

$$PC_t = C_t/C_b = (A_t/V_t)/C_b \quad (1)$$

with amount (A) in tissue $A_t = \Sigma C_{ij} V_i$ (i=l, p, w; j=ui, +, −)

where C_{ij} is the concentration of different ionic forms of the compounds in tissue composition, and V_i is the volume of the tissue composition (subscripts l, p, and w indicate lipid, protein, and water in tissue, respectively; subscripts ui, +, and − indicate neutral form, cationic form, and anionic form of a compound respectively).

We obtain

$$PC_t = (\Sigma C_{ij} V_i/V_t)/C_b$$

$$= \Sigma (V_i/V_t)(C_{ij}/C_b) \quad (i=l, p,w; j=ui, +, -) \quad (2)$$

Since

$$C_b = C_{bj}/f_j \quad (j=ui, +, -)$$

where f_j is the fraction of the compound in the neutral, cationic, and anionic forms at a given pH of the aqueous phase.

Therefore

$$PC_t = \Sigma f_j (V_i/V_t)(C_{ij}/C_{bj}) \quad (i=l, p, w; j=ui, +, -) \quad (3)$$

With

$$\text{volume fraction } \upsilon_i = V_i/V_t \text{ (i=l, p, w)}$$

$$\text{partition coefficient } P_{ij} = C_{ij}/C_{bj} \quad \text{(i=l, p, w; j=ui, +, −)}$$

Eq 3 is transformed to eq 4

$$\log PC_t = \log(\Sigma f_j\, 10^{\log P_{ij} + \log \upsilon_i}) \text{ (i=l, p, w; j=ui, +, −)} \tag{4}$$

By assuming that $\log P_{ij}$ can be linearly described by corresponding physicochemical descriptors X_{ij}, such as the partition coefficients compound of neutral and cationic forms in lipid, protein, and water, respectively:

$$\log P_{lui} = a_{1ui}X_{1ui} + a_{2ui}X_{2ui} + ... + a_{0ui}$$

$$\log P_{pui} = b_{1ui}X_{1ui} + b_{2ui}X_{2ui} + ... + b_{0ui}$$

$$\log P_{wui} = c_{1ui}X_{1ui} + c_{2ui}X_{2ui} + ... + c_{0ui}$$

$$\log P_{l+} = a_{1+}X_{1+} + a_{2+}X_{2+} + ... + a_{0+}$$

$$\log P_{p+} = b_{1+}X_{1+} + b_{2+}X_{2+} + ... + b_{0+}$$

$$\log P_{w+} = c_{1+}X_{1+} + c_{2+}X_{2ui} + ... + c_{0+}$$

In the present research only include neutral and cationic forms, therefore $f_- = 0$ in eq 4.

Physicochemical Descriptors and Methods. In the regression analysis the molecular polarizability (α) was used to characterize the overall contribution of the steric bulk effect or volume to the QSARs, and the maximum positive charge (q^{+max}) was used to characterize the electrostatic interaction. Other physicochemical descriptors employed were as follows: the sum of all positive partial atomic charges for all atoms in the molecule (ΣQ^+) as a descriptor of polarity, the sum of H-bond factor values for all acceptor substructures in the molecule (ΣCa), the sum of H-bond factor values for all donor atoms in a molecule (ΣCd), and the maximum H-bond acceptor descriptor in a molecule (Ca^{max}). All descriptors were calculated by the program package HYBOT/HYBOT-PLUS-98[14] which is based on methods described in several articles: H-bond descriptors,[24] partial atomic charges,[25] and polarizability.[26]

The fraction of un-ionized and ionized basic compounds were calculated at pH 7.4 using the following formula:[27]

$$f_{ui} = 1/(1 + 10^{pKa-7.4}), \quad f_+ = 1 - f_{ui}$$

All descriptor values are listed in Table 2.

Nonlinear regression analyses were performed using a standard regression program (GFA BASIC 4.38). In the regression equations *n* is the number of data points considered, *r* is the correlation coefficient, *s* is the standard error of the estimate, and *Q* is the cross-validated correlation coefficient derived from the predictive residual sum of squares (PRESS, leave-one-out method). Regression coefficients are given with their 95% confidence intervals.

RESULTS AND DISCUSSION

First of all, 46 compounds were randomly split into two data sets: 35 compounds in Table 1 as a training set (201 data points) and 11 compounds in Table 1 as a test set (64 data points).

For the analysis of the data of the training set in Table 1 we use eq 4 describing the distribution of a compound between seven tissues and blood. We tested the property descriptors provided by the HYBOT program (see Table 2), and the weight fractions were taken from ref 19 and listed in Table 3. The volume fractions (υ_l, υ_p, and υ_w) in eq 4 can be replaced by corresponding weight fractions (w_l, w_p, and w_w) approximately.[19,22] Some research has shown that the values of the weight fractions are very similar among rats, rabbits, and humans.[28] Therefore we use the weight fractions of human instead of those of rabbit in the equation.

We introduced HYBOT descriptors (α, q^{+max}, ΣQ^+, Ca^{max}, ΣCa, and ΣCd) into eq 4 in a stepwise manner until the statistical result cannot be further improved, and the following equations were obtained (for the method of a more detailed derivation see ref 22):

$$\log PC_t = \log(f_{ui}\,(10^{(\log P_{lui}+\log w_l)} + 10^{(\log P_{pui}+\log w_p)} + 10^{(\log P_{wui}+\log w_w)}) + f_+(10^{(\log P_{l+}+\log w_l)} + 10^{(\log P_{p+}+\log w_p)} + 10^{(\log P_{w+}+\log w_w)})) \tag{5}$$

$$\log P_{lui} = 0.083(\pm 0.021)\alpha_{ui} + 0.738(\pm 0.224)\Sigma Q^+_{ui} - 0.569(0.104)\Sigma Ca_{ui} + 1.133(\pm 0.559)\Sigma Cd_{ui} + 0.663(\pm 0.251) \tag{5a}$$

$$\log P_{pui} = 0.040(\pm 0.007)\alpha_{ui} \tag{5b}$$

$$\log P_{wui} = -0.267(\pm 0.171) \tag{5c}$$

$$\log P_{l+} = 1.157(\pm 0.251)Ca^{max}_{+} \tag{5a′}$$

$$\log P_{p+} = 7.086(\pm 0.622)q^{+max}_{+} \tag{5b′}$$

$$\log P_{w+} = 0 \tag{5c′}$$

$$n = 201\ r = 0.905\ s = 0.291\ Q = 0.890$$

$$\log PC_t = \log(f_{ui}\,(10^{(\log P_{lui}+\log w_l)} + 10^{(\log P_{pui}+\log w_p)} + 10^{(\log P_{wui}+\log w_w)}) + f_+(10^{(\log P_{l+}+\log w_l)} + 10^{(\log P_{p+}+\log w_p)} + 10^{(\log P_{w+}+\log w_w)})) \tag{6}$$

$$\log P_{lui} = 0.047(\pm 0.017)\alpha_{ui} + 0.800(\pm 0.238)\Sigma Q^+_{ui} - 0.893(\pm 0.179)\ Ca^{max}_{ui} + 0.725(\pm 0.534)\Sigma Cd_{ui} + 0.973(\pm 0.223) \tag{6a}$$

$$\log P_{pui} = 0.039(\pm 0.007)\alpha_{ui} \tag{6b}$$

$$\log P_{wui} = -0.257(\pm 0.169) \tag{6c}$$

$$\log P_{l+} = 1.156(\pm 0.259)Ca^{max}_{+} \tag{6a′}$$

$$\log P_{p+} = 7.090(\pm 0.630)q^{+max}_{+} \tag{6b′}$$

$$\log P_{w+} = 0 \tag{6c′}$$

$$n = 201\ r = 0.902\ s = 0.295\ Q = 0.887$$

$$\log PC_t = \log(f_{ui}\,(10^{(\log P_{lui}+\log w_l)} + 10^{(\log P_{pui}+\log w_p)} + 10^{(\log P_{wui}+\log w_w)}) + f_+(10^{(\log P_{l+}+\log w_l)} + 10^{(\log P_{p+}+\log w_p)} + 10^{(\log P_{w+}+\log w_w)})) \quad (7)$$

$$\log P_{lui} = 0.083(\pm 0.021)\alpha_{ui} + 0.739(\pm 0.227)\Sigma Q^+_{ui} - 0.569(\pm 0.105)\Sigma Ca_{ui} + 1.130(\pm 0.565)\Sigma Cd_{ui} + 0.662(\pm 0.255) \quad (7a)$$

$$\log P_{pui} = 0.040(\pm 0.007)\alpha_{ui} \quad (7b)$$

$$\log P_{wui} = -0.267(\pm 0.174) \quad (7c)$$

$$\log P_{l+} = 0.040(\pm 0.010)\,\alpha_+ \quad (7a')$$

$$\log P_{p+} = 7.174(\pm 0.616)q^{+max}_+ \quad (7b')$$

$$\log P_{w+} = 0 \quad (7c')$$

$$n = 201\ r = 0.902\ s = 0.295\ Q = 0.885$$

Here, the regression coefficients in parentheses are the regression coefficients of 95% confidence intervals.

Eq 5 was used to calculate log PC_t values of the training set and the test set in Table 1 for seven tissues. The results are presented in Table 1. The plots of calculated versus experimental logPC_t values of seven tissues for the training set (201 data points) and the test set (64 data points) are shown in Figures 1 and 2, respectively. All compounds in the training set and the test set can be usually described well.

The following descriptors were also used in regression analysis: the maximum negative charge (q^{-max}) and the maximum H-bond donor (Cd^{max}); however, they hardly improved these results further. Some relevant correlation matrices with acceptable descriptor intercorrelation coefficients are presented in Table 4. To ensure that the results of eqs 5−7 are not due to a chance correlation, some different training sets and test sets including different compounds in Table 1 were used to reproduce eqs 5−7 and Figure 2; all results are in high accordance with the present results. These show that eqs 5−7 are very robust.

Eq 4 means that the tissue−blood partition coefficient (PC_t) can be expressed as the weighed sum of three partition coefficients (lipid−blood, protein−blood, and water−blood) of neutral and ionic forms; therefore, eqs 5−7 can give excellent results. This shows that the equilibrium distribution of a compound in a variety of tissues is essentially the equilibrium distribution of the compound in several tissue (chemical) compositions.

Because the neutral form and the ionized form of a compound usually have different partition coefficients in different chemical compositions, therefore to divide a tissue based on their composition and consider ionized forms of compounds obviously are reasonable. The partition coefficient of a compound between tissue and blood obtained by the Hansch equation is an apparent partition coefficient of the compound including different ionic forms between several compositions and blood. Thus, the nonlinear equation can reflect a realistic situation better and usually obtain a better result than the linear equation for a multiphase system,[22] furthermore the most real systems are multiphase systems. The model equation can be considered as the expressive form of the Hansch equation in nonlinear spaces (multiphase system).

It is well-known that wet octanol as a compartment of pure water does not behave as it does in the pure component. However our present result and our recent research[22] seem to show that lipid and protein and water in tissues behave as they do in the pure components.

This can be explained that lipid and protein have fixed local distribution in vivo, and lipid and protein have a large difference in nature. Lipid mainly interacts with lipophilic molecules, and protein mainly interacts with hydrophilic and charged molecules; therefore, a compound either mainly interacts with lipid or with protein. It is seldom possible that a compound strongly interacts with lipid and protein simultaneously due to the difference of properties between lipid and protein as well as the space distribution of lipid and protein in vivo. Such as usually a receptor drug only strongly interacts with protein. Water and lipid have also a large difference, and a well absorbed drug is mainly affected by absorption of lipid and protein; therefore, some coupling interactions may be ignored.

In eq 5 five parameters were used to fit neutral compound partitioning into tissues, and only two parameters were used to fit ionic compound partitioning into tissues. As expected, for neutral compounds or the neutral form of compounds the tendency to partition into lipid (eq 5a) increases with an increasing polarizability α, which is mainly bulk descriptors and partly also a descriptor of van der Waals and dispersion interactions between unpolar substructures of compounds and lipids. Furthermore, both the ΣCa, term H-bond accepting substructures of compounds, and the ΣCd, term H-bond donor substructures of compounds, lead to reduced lipid absorption; obviously, hydrophilicity unfavors compounds into lipid. A positive effect on lipid absorption is also induced by polar groups ΣQ^+ in compounds. Polarizability α favors compounds to bind to protein because of protein with polarity, and the polarity−polarity interaction between a compound and protein favors their bind (see eq 5b). It is very similar with the relationship between the logarithm of tissue−blood partition of neutral compounds and HYBOT descriptors in our recent work.[22]

For ionic compounds or the ionic form of compounds the tendency to partition into lipid (eq 5a′) increases with increasing the maximum H-bond acceptor descriptor Ca^{max}; maybe this shows that the maximum H-bond acceptor of a group can slightly interact with the outer hydrated polar of lipid. The maximum positive charges (q^{+max}) play an important role in compounds partitioning into protein (eq 5b′). The maximum positive charges (q^{+max}) maybe interact with a negatively charged carboxyl of protein, which leads to a high absorption. The highly q^{+max}-coefficients in eq 5 are because electrostatic interaction is obviously a much stronger interaction than other interactions. Nowadays it is generally known that numerous well absorbed drugs appear as ions in vivo, which can be depicted by eqs 5−7.

Compared with eq 5, the H-bond accepting substructure ΣCa in eq 5a is replaced by the maximum H-bond acceptor descriptor Ca^{max} in eq 6a. The statistical significance of eq 6 is slightly worse than that of eq 5, and there is usually a high intercorrelation between ΣCa and Ca^{max}(r=0.897). This

Table 2. Physicochemical Property Data of Compounds

no.	compound		α_{Ui}/α_+	$q^{+max}{}_{ui}/q^{+max}{}_+$	$\Sigma Q^+{}_{ui}/\Sigma Q^+{}_+$	$Ca^{max}{}_{ui}/Ca^{max}{}_+$	$\Sigma Ca_{ui}/\Sigma Ca_+$	$\Sigma Cd_{ui}/\Sigma Cd_+$	f_{ui}/f_i
1	2,2-dimethylbutane	(ui)	11.784	0.035	0.435	0.000	0.000	0.000	1.000
2	pentane	(ui)	9.949	0.034	0.365	0.000	0.000	0.000	1.000
3	2-Me-pentane	(ui)	11.784	0.038	0.436	0.000	0.000	0.000	1.000
4	3-Me-pentane	(ui)	11.784	0.039	0.436	0.000	0.000	0.000	1.000
5	hexane	(ui)	11.784	0.034	0.436	0.000	0.000	0.000	1.000
6	3-Me-hexane	(ui)	13.619	0.039	0.508	0.000	0.000	0.000	1.000
7	heptane	(ui)	13.619	0.035	0.508	0.000	0.000	0.000	1.000
8	cyclopropane	(ui)	5.505	0.030	0.182	0.000	0.000	0.000	1.000
9	Me-cyclopentane	(ui)	11.010	0.039	0.407	0.000	0.000	0.000	1.000
10	cyclohexane	(ui)	11.010	0.034	0.411	0.000	0.000	0.000	1.000
11	benzene	(ui)	10.434	0.034	0.206	0.080	0.480	0.000	1.000
12	toluene	(ui)	12.269	0.036	0.267	0.088	0.528	0.000	1.000
13	styrene	(ui)	13.912	0.037	0.271	0.220	1.160	0.000	1.000
14	CH_2Cl_2	(ui)	6.465	0.040	0.080	0.150	0.300	0.000	1.000
15	$CHCl_3$	(ui)	8.393	0.050	0.088	0.080	0.240	0.000	1.000
16	Me-CCl_3	(ui)	10.228	0.043	0.149	0.070	0.210	0.000	1.000
17	CF_3-CH_2Cl	(ui)	6.099	0.358	0.569	0.386	0.446	0.000	1.000
18	teflurane	(ui)	6.706	0.389	0.737	0.174	0.323	0.000	1.000
19	halothane	(ui)	8.725	0.364	0.559	0.100	0.260	0.000	1.000
20	$CF_2=CHCl$	(ui)	5.998	0.384	0.465	0.167	0.324	0.000	1.000
21	$CCl_2=CHCl$	(ui)	10.036	0.042	0.079	0.100	0.220	0.000	1.000
22	MeOH	(ui)	3.246	0.102	0.248	1.630	1.630	−1.520	1.000
23	EtOH	(ui)	5.081	0.104	0.320	1.660	1.660	−1.420	1.000
24	1-PrOH	(ui)	6.916	0.105	0.391	1.610	1.610	−1.270	1.000
25	2-PrOH	(ui)	6.916	0.106	0.392	1.690	1.690	−1.060	1.000
26	1-BuOH	(ui)	8.751	0.105	0.461	1.570	1.570	−1.150	1.000
27	acetone	(ui)	6.433	0.102	0.312	1.950	1.950	0.000	1.000
28	butanone	(ui)	8.268	0.102	0.381	1.910	1.910	0.000	1.000
29	Et_2O	(ui)	8.751	0.052	0.446	1.540	1.540	0.000	1.000
30	divinyl ether	(ui)	8.004	0.099	0.438	0.750	0.970	0.000	1.000
31	flurane	(ui)	7.923	0.372	0.910	0.886	1.066	0.000	1.000
32	enflurane	(ui)	8.389	0.368	1.148	0.800	1.400	0.000	1.000
33	isoflurane	(ui)	8.389	0.390	1.125	0.952	1.512	0.000	1.000
34	methoxyflurane	(ui)	10.590	0.319	0.701	0.952	1.392	0.000	1.000
35	sevoflurane	(ui)	8.114	0.411	1.564	0.952	1.161	0.000	1.000
36	NO	(ui)	1.601	0.062	0.062	0.000	0.000	0.000	1.000
37	pentazocin	(ui)	34.174	0.112	1.284	2.000	3.696	−2.854	0.074
		(+)	34.561	0.295	2.593	1.000	1.662	−7.461	0.926
38	nitrazepam	(ui)	29.165	0.199	1.012	2.500	6.230	−2.529	1.000
39	haloperidol	(ui)	39.814	0.197	1.558	2.000	5.919	−1.867	0.258
		(+)	40.201	0.293	2.922	1.550	3.969	−6.177	0.715
40	biperiden	(ui)	37.433	0.113	1.374	1.998	4.041	−1.867	0.038
		(+)	37.820	0.293	2.693	1.521	2.043	−6.177	0.962
41	diazepam	(ui)	31.087	0.195	0.996	2.590	5.554	0.000	1.000
42	promethazine	(ui)	34.023	0.061	0.964	1.929	4.139	0.000	0.020
		(+)	34.410	0.291	2.314	1.200	2.210	−4.607	0.980
43	trihexyphenidyl	(ui)	36.564	0.113	1.421	1.998	3.849	−1.867	0.048
		(+)	36.951	0.293	2.740	1.521	1.851	−6.177	0.952
44	chlorpromazine	(ui)	35.951	0.066	0.974	1.933	4.258	0.000	0.012
		(+)	36.338	0.288	2.272	1.200	2.325	−4.690	0.988
45	clotiazepam	(ui)	33.979	0.197	1.030	2.250	5.271	0.000	1.000
46	clomipramine	(ui)	36.921	0.056	1.119	1.933	4.028	0.000	0.074
		(+)	37.308	0.288	2.416	1.200	2.095	−4.690	0.926

Table 3. Tissue Composition (Weight Fraction)[a]

tissue	w_l[b]	w_p[b]	w_w[b]
kidney	0.050	0.170	0.770
brain	0.107	0.079	0.790
muscle	0.020	0.170	0.790
lung	0.010	0.177	0.780
liver	0.070	0.180	0.720
fat	0.800	0.050	0.150
heart	0.100	0.167	0.727

[a] Taken from ref 19. [b] w_l, w_p, and w_w are weight fractions of lipid, protein, and water, respectively.

shows that the maximum H-bond acceptor descriptor Ca^{max} also leads to reduced lipid absorption. In eq 7 polarizability α in eq 7a′ is significant; this indicates that polarizability α favors ionic compounds into lipid as neutral compounds do.

The differences of these subequations indicate that neutral and ionic forms of a compound as well as they in different tissue (chemical) compositions have different mechanisms of action in vivo.

As pointed to by Liu et al.[1] due to the physical nature of the lipid bilayer, a successful model of drug transport must include descriptors accounting for both molecular hydrophilicity and hydrophobicity. However, Liu et al. failed to generate a statistically sound linear equation using molecular hydrophilicity and hydrophobicity to describe logBB values. Not only do eqs 5−7 include hydrophilic and hydrophobic descriptors but also they include electrostatic descriptors; therefore, it should bear a more universal suitability for the prediction tissue−blood partition coefficients of diverse compounds.

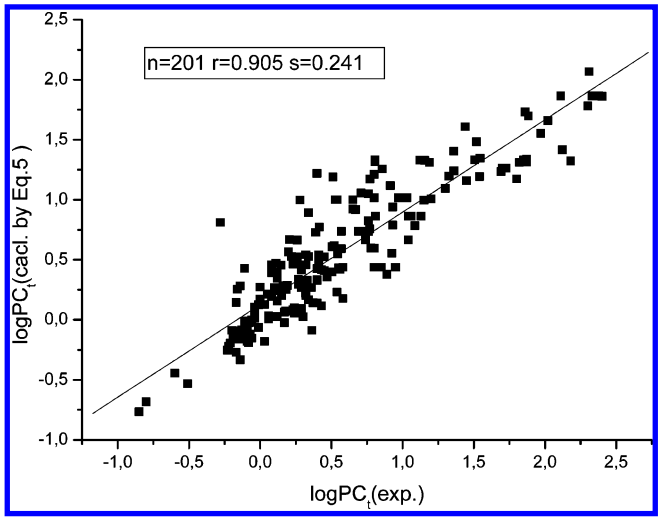

Figure 1. Calculated $\log PC_t$ values using eq 5 versus experimental $\log PC_t$ values for 201 data points in the training set.

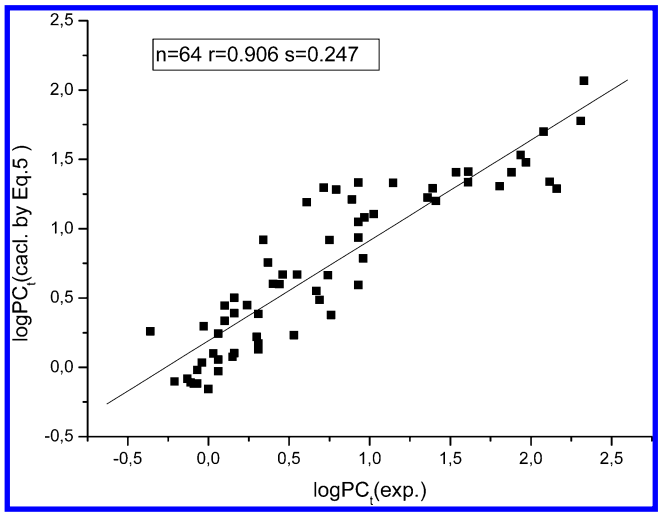

Figure 2. Calculated $\log PC_t$ values using eq 5 versus experimental $\log PC_t$ values for 64 data points in the test set.

Table 4. Some Correlation Matrices (*r*-Values) between Descriptors in Eqs 5−7

	α_{ui}	ΣQ^+_{ui}	Ca^{max}_{ui}	ΣCa_{ui}	ΣCd_{ui}
α_{ui}	1.000				
ΣQ^+_{ui}	0.653	1.000			
Ca^{max}_{ui}	0.590	0.556	1.000		
ΣCa_{ui}	0.830	0.664	0.897	1.000	
ΣCd_{ui}	−0.470	−0.374	−0.631	0.701	1.000

Poulin's model equation[4,20,21] can only predict the partition coefficients of nonexcretory tissues. However, not only can our model equation predict the partition coefficients of nonexcretory tissues but also our model equation can predict the partition coefficients of excretory tissues, such as partition coefficients of liver and kidney.

Obviously, eq 4 can be also used to calculate the partition coefficients of a single tissue, and this will result in a more accurate prediction for tissue−blood partition coefficients than results of a linear equation;[22] however, in the present research the work has not been performed due to including less ionic compounds in this data set.

In present work pK_a values were taken from experimental results; however, pK_a values can be calculated using some theoretic methods such as semiempirical and ab initio quantum mechanics method and linear free energy relationship (LFER).[28−31] LFER method implemented in ACD/pK_a DB software can give a reasonable pK_a value when the fragments of the molecule are in the training data and the calculation is very fast.[31] Therefore not only can the model perform an accurate calculation but also the model can be used in virtual high-throughput screening (VHTS).

Further work should be done to screen some more appropriate descriptors for the model equation so that it can perform a more accurate calculation and/or a faster calculation for a different purpose.

CONCLUSION

The model equation is the expressive form of the Hansch equation in nonlinear spaces (multiphase system). It can better reflect a real situation; therefore, it has a stronger predictive power than the Hansch equation for tissue−blood partition coefficients. Meanwhile this model will also help us to understand the absorption mechanism of compounds in vivo.

ACKNOWLEDGMENT

This research was supported in part by a Research Fund of the Ministry of Education of China for Homecoming Scholars.

REFERENCES AND NOTES

(1) Liu, R.; Sun, H.; Sung-Sau So, S.-S. Development of quantitative structure−property relationship models for early ADME evaluation in drug discovery. 2. blood-brain barrier penetration. *J. Chem. Inf. Comput. Sci*. **2001**, *41*, 1623−1632.
(2) Caldwell, G. W. Compound optimization in early- and late-phase drug discovery: acceptable pharmacokinetic properties utilizing combined physicochemical, in vitro and in vivo screens. *Curr. Opin. Drug Discuss.* **2000**, *3*, 30−41.
(3) Lin, J. H.; Lu, A. Y. Role of pharmcokinetics and metablism in drug discovery and development. *Pharmacol. Rev*. **1997**, *49*, 403−449.
(4) Bjorkman, S. Predivtion of the volume of distribution of a drug: which tissue-plasma partition coefficients are need? *J. Pharm. Pharmaco*. **2002**, *54*, 1237−1245.
(5) Hansch, C.; Maloney, P. P.; Fujita, T. Correlation of biological activity of phenoxyacetic acids with hammett substituent constants and partition coefficients. *Nature* **1962**, *194*, 178−180.
(6) Kurup, A.; Garg, R.; Hansch, C. Comparative QSAR Study of Tyrosine Kinase Inhibitors. *Chem. Rev.* **2001**, *101*, 2573−2600.
(7) Kurup, A.; Garg, R.; Hansch C. Comparative QSAR: Angiotensin II Antagonists. *Chem. Rev*. **2001**, *101*, 2727−2750.
(8) Rose, K.; Hall, L. H.; Kier, L. B. Modeling blood-brain barrier partitioning using the electrotopological state. *J. Chem. Inf. Comput. Sci*. **2002**, *42*, 651−666.
(9) Waterbeemd, H.; Testa, B. *Computer-assisted lead finding and optimization*; Folkers, G., Ed.; Verlag Helvetica Chimica Acta: Basel, 1997; pp 367−378.
(10) Iyer, M.; Mishru, R.; Han, Y.; Hopfinger, A. J. Predicting blood-brain barrier partitioning of organic molecules using membrane-interaction QSAR analysis. *Pharm. Res.* **2002**, *19*, 1611−1621.
(11) Bodor, N.; Huang, M. J. A new method for the estimation of the aqueous solubility of organic compounds. *J. Pharm. Sci*. **1992**, *81*, 954−960.
(12) Lobell, M.; Molnar L.; Keseru, G. Recent advances in the prediction of blood-brain partition from molecular structure. *J. Pharm. Sci.* **2003**, *92*, 360−370.
(13) Lombardo, F.; Blake, J. F. Curatolo, W. J. Computation of brain-blood partitioning of organic solutes via free energy calculations. *J. Med. Chem*. **1996**, *39*, 4750−4755.
(14) Raevsky, O. A. Hydrogen bond strength estimation by means of HYBOT. In *Computer-assisted lead find and optimization*; van de Waterbeemd, H., Testa, B., Folkers, G., Eds.; Verlag Helvetica Chimica Acta: Basel, 1997; pp 367−378.
(15) Abraham, M. H.; Taft, J. J.; Talor, P.; Laurence, C.; Berthelot, M.; Doherty, R. M.; Abboud, J. L. M.; Sraidi, K.; Guiheneuf, G. A general

treatment of hydrogen bond complexation constants in tetrachloromethane. *J. Am. Chem. Soc*. **1988**, *110*, 8534−8536.
(16) Abraham, M. H.; Chadha, H.; Michell, R. C. Hydrogen bonding.3 Factors that influence the distribution of solutes between blood and brain. *J. Pharm. Sci*. **1994**, *83*, 1257−1268.
(17) Kelder, J.; Grootenhuis, P. D. J.; Bayada, D. M.; Delbressine, L. P. C.; Ploemen, J.-P. Polar molecular surface as a dominating determinant for oral absorption and brain penetration of drugs. *Pharm. Res.* **1999**, *16*, 1514−1519.
(18) Clark, D. E. Rapid calculation of transport phenomena. 2. Prediction of blood-brain barrier penetration. *J. Pharm. Sci*. **1999**, *88*, 815−821.
(19) Balaz, S.; Lukacova, V. A model-based dependence of the human tissue/blood partition coefficient of chemicals on lipophicity and tissue composition. *Quant. Struct.-Act. Relat*. **1999**, *18*, 361−368.
(20) Poulin, P.; Schoenlein, K.; Theil, F. P. Prediction of adipose tissue: plasma partition coefficients for structurally unrelated drugs. *J. Pharm. Sci*. **2001**, *90*, 436−47.
(21) Poulin, P.; Theil, F. P. Prediction of pharmacokinetics prior to in vivo studies. II. Generic physiologically based pharmacokinetic models of drug disposition. *J. Pharm. Sci*. **2002**, *91*, 1358−1370.
(22) Zhang, H. A new nonlinear equation for the tissue/blood partition coefficients of neutral compounds. *J. Pharm. Sci*. **2004**, *93*, 1595−604.
(23) Yokogawa, K.; Nakashima, E.; Ishizaki, J. Relationships in the structure-tissue distribution of basic drugs in the rabbit. *Pharm. Res*. **1990**, *7*, 691−696
(24) Raevsky, O. A. Quantification of noncovalent interactions on the basis of the thermodynamic hydrogen bond parameters. *J. Phys. Org. Chem*. **1997**, *10*, 405−413.
(25) Kireev, D. B.; Fetisov, V. I.; Zefirov, N. S. Approximate molecular electrostatic potential computations: Application to quantitative structure−activity relationships. *J. Mol. Struct.* (THEOCHEM) **1994**, *304*, 143−150.
(26) Miller, K. J. Additivity methods in molecular polarizability. *J. Am. Chem. Soc*. **1990**, *112*, 8533−8538.
(27) Martin, A. N.; Swarbrick, J.; Cammarata, A. *physical pharmacy, physical chemical principles in the pharmaceutical sciences*; Lea & Febiger: Philadelphia, PA, 1969; ISBN 8121-01634.
(28) Dejongh, J.; Verhaar, H. J. D.; Hermens, J. L. M. A quantitive property-property relation [QPPR] approach of organic chemicals in rats and humans. *Arch Toxicol*. **1997**, *72*, 17−25.
(29) Tehan, B. G.; Lioyd, E. J.; Wong, M. G.; Pitt, W. R.; Montana, J. G.; Manallack D. T.; Gancia, E. Estimation of pK_a using semiempirical molecular orbital methods. Part 1: application to phenols and carboxylic acids. *Quant. Struct.-Act. Relat*. **2002**, *21*, 457−472.
(30) Liptak, M. D.; Gross, K.C.; Seybold P. G.; Feldgus, S.; Shields, G. C. Absolute pK_a determinations for substituted phenols. *J. Am. Chem. Soc*. **2002**, *124*, 6421−6427.
(31) ACD/Lab ACD Labs, version 5.0 Richmond St. West Suite 605, Toronto, ON M5H 2L3, Canada.

CI049718E