Prediction of Cytochrome P450 2B6-Substrate Interactions Using Pharmacophore Ensemble/Support Vector Machine (PhE/SVM) Approach

Max K. Leong* and Tzu-Hsien Chen

Department of Chemistry, National Dong Hwa University, Shoufeng, Hualien 97401, Taiwan

Abstract: An *in silico* model for predicting human cytochrome P450 2B6-substrate interactions was generated based on a novel scheme, which was initially devised to predict the hERG liability (reported in Leong, M. K., *Chem. Res. Toxicol.,* **2007**, *20*, 217.) using pharmacophore ensemble/support vector machine to take into account the protein conformational flexibility while interacting with structurally diverse substrates. This is of critical importance yet never being addressed by any analogue-based molecular modeling studies before. Thirty-seven molecules were chosen from the literature and scrutinized for structural integrity and data consistency, of which 26 were treated as the training set to generate models, which were subject to validation by the other 11 molecules as the test set. The predicted pK_m values by the final PhE/SVM model were in good agreement with observed values. In addition, this *in silico* model produced an r^2 of 0.84 and a 10-fold cross-validation q^2 of 0.66 for the training set and an r^2 of 0.87 for the test set, asserting the fact that this PhE/SVM model is an accurate model to predict the human P450 2B6-substrates interactions and can be used as a robust prediction tool to facilitate drug discovery.

Key Words: Pharmacophore ensemble, support vector machine, cytochrome 2B6, *in silico*, plasticity.

INTRODUCTION

The polymorphic cytochrome P450s (CYPs) constitute a superfamily of hemoproteins that are the principal enzymes involved in the biotransformation of many endogenous and xenobiotic compounds, including anticancer drugs and a variety of promutagens and procarcinogens [1-9]. These enzymes are expressed in many mammalian tissues [10] with the highest levels in liver [11, 12]. Individual P450s exhibit unique substrate specificity, region- and stereoselectivity profiles that reflect a different tertiary structure [13-16]. Of all CYP450 isozymes, CPY1A2, CYP1A6, CYP2B6, CYP2C9, CYP2C19, CYP2D6, CYP2E1 and CYP3A4 are involved in xenobiotic metabolism [17]. Additionally, the human CYP2B6 can metabolize about 3% of the clinical drugs [18] despite the fact that it only constitutes a small portion of total hepatic P450 complement (< 1%) [19]. Not until recently have its structure-function analyses begun as new clinically relevant substrates that have been identified [10, 20] and has heterologous expression been accomplished [21], indicating that CYP2B6 and CYP2B6-dependent substrates play increasingly important roles for drug metabolism [22-29]. Accordingly, it is of importance to develop an *in silico* model to predict the CYP2B6-substrate interaction in the process of drug discovery in the hope of reducing the attrition rates due to the adverse effects of the drug with cytochrome 2B6 since numerous studies have demonstrated that drug design can be substantially facilitated using computational approaches [30-33].

Since there is no CYP2B6 crystal structure available to date, it seems to be a plausible approach to build a CYP2B6 homology model based on recently published crystal structures of other mammalian CYP450s [34-41] similar to a recent 2C19 study [42]. In fact, Lewis *et al*. built a CYP2B6 homology model based on the CYP102 crystal structure, which, nevertheless, does not serve as a good homology modeling template since, most unreliably, it is a bacterial CYP, *viz*. a soluble cytosolic protein. Domanski *et al*. created a model derived from the CYP2B1 homology model [29]. Wang and Halpert took a more relatively realistic scheme to construct a CYP2B6 model by homology modeling based on the crystal structure of CYP2C5 [43]. However, all of CYP450 homology models published to date failed to take into account one critical fact that all CYP450 isozymes contain a metal iron, which are transitional-metal-containing systems and can only be modeled by quantum mechanics or QM/MM methods instead of traditional molecular mechanics [44] that has been further explained in detail elsewhere in the case of CYP450s [45].

In addition, it is believed that P450s have various degrees of flexibility [46], which can be manifested by the fact that CYP2C5, for example, can adapt different conformations in order to bind with various structurally diverse substrates [47, 48]. Furthermore, the helix I of CYP2B4 significantly moves about 15° to enlarge the binding pocket for larger substrates [49] and a recent study done by Ekroos and Sjögren also demonstrated that the binding pocket of CYP3A4 can substantially increases more than 80% by volume upon binding with small molecules [50]. In some cases, not only does the protein undergo conformation changes, but the substrate will take different orientations to bind to the same target protein [47], suggesting that the interaction between the substrate and protein is dynamic *per se* [51]. As a result, it will be an impractical approach, if not entirely flawless, to use a single protein conformation to model the protein-substrate interactions in the case of P450s [49] and the quality of homology

*Address correspondence to this author at the Department of Chemistry, National Dong Hwa University, Shoufeng, Hualien 97401, Taiwan. Tel.: +886 3 863 3609. Fax: +886 3 863 3570. E-mail: leong@mail.ndhu.edu.tw

1573-4064/08 $55.00+.00

models for membrane-bound proteins constructed from a single substrate will become relatively unreliable [52] and is far from perfect [53].

Alternatively, multiple protein conformations excerpted from multiple copies of the protein-ligand co-complex structures should be employed and lengthy molecular dynamics calculations should be carried out, if necessary, in order to cover broader "conformational space" or "conformational domain." In other words, the protein conformation ensemble, which is consisted of various protein conformations, should be employed so that more accurate protein-ligand interactions can be modeled that, nevertheless, will inevitably significantly increase amount of computing time and make the whole modeling calculation extremely difficult in some cases, if not absolutely infeasible.

In addition to structure-based calculations, numerous efforts have been devoted to ligand-based molecular modeling. Ekins *et al*., for example, developed a pharmacophore hypothesis and a QSAR model [54], Lewis *et al*. built a QSAR model [20], Wang and Halpert conducted a 3D-QSAR study [43], and a brief summary can be found elsewhere [55]. However, all of these proposed ligand-based models share one common characteristic, namely they all fail to address the issue of protein plasticity, which, as mentioned above, is pronounced in the case of CYP450s. As a result, these proposed models can accurately predict some compounds with a very specific chemotype, which corresponds to a specific protein conformation when interacting with those molecules within the model generation scope, and yet generates dramatic prediction deviations for any other compounds outside of the training scope, rendering a model's limited applicability.

On the other hand, every pharmacophore hypothesis, composed of a variety of chemical features, corresponds to a protein conformation or a set of protein conformations that are topologically close to one another. It can be expected that a group of pharmacophore hypotheses or pharmacophore ensemble should be constructed in case multiple protein conformations have to be taken into consideration and each pharmacophore represents at least one of the protein conformations under consideration. By doing so, the protein plasticity can be taken into account by using pharmacophore ensemble in the case of ligand-based modeling compared to protein ensemble in the case of structure-based modeling. In fact, the ensemble approach has been utilized for predicting protein subcellular localization [56-60], membrane protein types [61], enzyme functional classes [62] and signal peptides [63, 64]; and has been systematically introduced by two recent review articles [65, 66]. Recently, a novel scheme has be devised, in which a pharmacophore ensemble was constructed from a number of plausible pharmacophore candidates, corresponding to various protein conformations, and subject to regression by a support vector machine [67]. This novel PhE/SVM scheme performed well to predict the *hERG* liability for a variety of structurally highly diverse molecules. Consequently, it is a plausible approach to employ this novel PhE/SVM to build a predictive CYP2B6-substrate interaction model based on a variety of ligand structures because of the flexible nature of CYP450s.

THEORETICAL CALCULATIONS

The K_m values of cytochrome P450 2B6-substrate interaction were searched from the literature and gathered as many as possible to maximize the structural diversity. Experimental conditions were carefully inspected to assure assay consistency. The experimental variations between groups were minimized by trying to include biological data from the same research group. If there were two or more available biological data for a given compound and in a very close range, the average values were then taken in order to warrant better consistency. Furthermore, chemical structures were examined with caution and compounds without defined stereochemistry were excluded. Thirty-seven molecules were selected in this study, and their corresponding biological activities, and references to the literature are listed in Table **1**.

Table 1. Selected Compounds for This Study, Their K_m (nM) Values or Average Values if Applicable and References

Molecules	K_m(nM)	Refs.
Pentoxyresorufin	840	[5]
Benzyloxyresorufin	1,280	[54]
7-Ethoxy-4-trifluoromethylcoumarin	2,300	[25, 54]
Desmethylrabeprazole-thioether	3,800	[126]
7-Methoxy-4-trifluoromethylcoumarin	6,790	[127]
Efavirenz	12,400	[128]
Cinnarizine	17,200	[54]
8-Hydroxyefavirenz	20,200	[128]
RP73401	22,500	[129]
β-Arteether	28,000	[130]
S-Methadone	30,000	[15]
R-Deprenyl	33,000	[131]
Clotiazepam	33,100	[7]
4-Chloromethyl-7-ethoxycoumarin	33,700	[27]
N,N-diethyl-*m*-toluamide	40,200	[132]
S-Ketamine	44,400	[133]
Midazolam	46,100	[27]
Testosterone	50,500	[27]
Loperamide	65,600	[134]
3-Cyano-7-ethoxycoumarin	71,300	[27]
R-Methadone	73,000	[15]

(Table 1. Contd….)

Molecules	K_m(nM)	Refs.
R-Ketamine	74,400	[133]
Diazepam	113,000	[28]
7-Ethoxycoumarin	115,000	[54]
Amintriptyline	144,400	[54]
Styrene	180,000	[28]
S-Mephobarbital	264,000	[135]
Dextromethorphan	350,000	[54]
Meperidine	356,000	[136]
Imipramine	383,000	[54]
Lidocaine	537,600	[54]
S-Mephenytoin	564,000	[54]
Carbamazepine	700,000	[137]
Ifosfamide	1,730,000	[7]
SM-12502	1,767,000	[28]
Cyclophosphamide	1,890,000	[138]
1,2-Dibromoethane	9,700,000	[54]

Conformational analysis was done by *MacroModel* (Schrödinger, Portland, OR) using mixed Monte Carlo multiple minimum [68] (MCMM)/low mode [69] and the hydration effect was taken into account using the GB/SA algorithm [70]. Energy minimization was carried out by truncated-Newton conjugated gradient method (TNCG) with the MMFFs force field [71]. Water was chosen as solvent with a constant dielectric constant. The energy window was set to 83.7 KJ/mol (or 20 Kcal/mol) and the number of unique structures was limited to 255, which is the maximal conformation number that can be accommodated by *Catalyst*.

Twenty-six molecules were selected as samples in the training set to generate pharmacophore hypotheses based on compounds' activities and chemical structures in order to achieve statistic significance as suggested by Sprague [72] and Güner [73]; and the remaining eleven compounds were treated as the test set to validate those generated hypotheses from the training set. Tables **2** and **3** list compounds selected for the training set and the test set, respectively, and their corresponding negative logarithm K_m values, namely pK_m determined.

Automatic pharmacophore generation was carried out using the *Catalyst* package (Accelrys, San Diego, CA) and lipid hydrogen-bond acceptor, hydrophobe and aromatic hydrophobe chemical features were selected based on all of compounds in the training set. Assorted combinations of variable weight and variable tolerance options were employed to establish the hypothesis diversity. The generated pharmacophore hypotheses were, in turn, used to predict the pK_m values of those compounds in the test set.

The statistic parameters, namely root-mean-square deviation (rmsd), the correlation coefficient (r^2), maximum residual, average residual and standard deviation of residuals between the observed and predicted values, were calculated. In addition, the cost difference between the generated hypothesis and the null hypothesis, which gauges the statistic significance of a hypothesis, was also evaluated. Only those pharmacophore hypotheses with excellent statistic performance were pooled to construct the pharmacophore ensemble.

The predicted pK_m values of those compounds in the training set by those hypotheses in the pharmacophore ensemble were treated as input for the support vector machine calculations, which were carried out using *LIBSVM* [74], to generate regressed models. The resultant SVM models in turn were verified by those compounds in the test set. A Perl script was coded to systemically scan through those associated parameters, namely cost C, the width of the RBF kernel γ, ε in case of *ε-SVR* and ν in case of *ν-SVR*.

In addition, the generated prediction models are usually subject to independent dataset test in order to validate those models. Among the independent dataset test, sub-sampling (e.g., 5 or 10-fold cross-validation) test, and jackknife test, which are often used for examining the accuracy of a statistical prediction method, the jackknife test was deemed the most rigorous and objective [75] as demonstrated by an incisive analysis in a recent comprehensive review [75] and has been increasingly and widely adopted by investigators to test the power of various prediction methods [76-125]. Nevertheless, a 10-fold cross-validation scheme was employed in this study due to its relative speed and accuracy.

RESULTS AND DISCUSSION

Three pharmacophore hypotheses, denoted by Hypo A, Hypo B, and Hypo C, were selected from all generated pharmacophore hypotheses to constitute the PhE based on their prediction performances on every single compound and statistic evaluations in the training set (Table **2**) as well as the test set (Table **3**), and cost differences (Table **4**). All of these three candidate models in the ensemble possess the same chemical features, namely two lipid hydrogen-bond acceptors, one aromatic hydrophobe and one hydrophobe as shown in Table **5** along with their corresponding three-dimensional coordinates, interfeature distances, weights, and tolerances.

Fig. (**1**) illustrates the spatial arrangement of these three models in the PhE and it can be observed that there are some discrepancies among them. The distances between aromatic hydrophobe and one of lipid hydrogen-bond acceptor features, for instance, are 7.401 Å, 7.425 Å, and 7.517 Å in Hypo A, Hypo B, and Hypo C, respectively, showing only 0.1 Å variations among them. The other aromatic hydrophobe–lipid hydrogen-bond acceptor distances, nevertheless, vary from 4.142 Å (Hypo A) to 3.780 Å (Hypo B) and 3.849 Å (Hypo C), resulting in 0.4 Å change in length. Similar observation can also be found for the angles centered at aromatic hydrophobe and connecting to hydrophobe and lipid hydrogen-bond acceptor that undergo changes from 127.1° and 155.1° in Hypo A to 116.8° and 154.1° in Hypo B, and 116.3° and 150.5° in Hypo C. The observation of differences

Table 2. Experimentally Observed pK_m Values, Predicted pK_m Values by Hypo A, Hypo B and Hypo C and SVM Model for Those Compounds in the Training Set, the Corresponding Residual Values (Δ) and Their Associated Statistic Numbers (Correlation Coefficient) r^2, Root-Mean-Square Deviation (rmsd), Maximum Residual, Average Residual, Standard Deviation of Residual, and Cross-Validation Coefficient q^2

Molecules	Obs.	Hypo A		Hypo B		Hypo C		SVM Model	
	pK_m	pK_m	Δ	pK_m	Δ	pK_m	Δ	pK_m	Δ
Pentoxyresorufin	6.07	6.34	0.27	6.31	0.24	6.33	0.26	5.93	-0.14
Benzyloxyresorufin	5.89	5.50	-0.40	5.36	-0.54	5.46	-0.44	5.74	-0.16
7-Ethoxy-4-trifluoromethylcoumarin	5.64	4.92	-0.72	5.04	-0.60	4.96	-0.68	5.14	-0.49
7-Methoxy-4-trifluoromethylcoumarin	5.17	4.89	-0.28	5.01	-0.16	4.92	-0.25	5.12	-0.05
Efavirenz	4.91	4.64	-0.27	4.70	-0.21	4.60	-0.31	4.77	-0.14
Cinnarizine	4.76	4.27	-0.50	4.11	-0.66	4.24	-0.53	4.11	-0.65
RP73401	4.65	4.89	0.24	5.00	0.36	5.14	0.49	4.81	0.16
β-Arteether	4.55	4.09	-0.46	4.17	-0.39	4.19	-0.36	4.14	-0.41
Clotiazepam	4.48	4.47	-0.01	4.48	0.00	4.28	-0.20	4.57	0.09
4-Chloromethyl-7-ethoxycoumarin	4.47	4.89	0.41	4.96	0.49	4.89	0.41	5.09	0.62
N,N-diethyl-*m*-toluamide	4.40	3.64	-0.76	3.59	-0.81	3.64	-0.76	3.71	-0.69
Loperamide	4.18	4.34	0.15	4.25	0.07	4.34	0.15	4.23	0.05
R-Methadone	4.14	3.82	-0.31	3.92	-0.22	3.92	-0.22	3.99	-0.15
R-Ketamine	4.13	3.64	-0.49	3.59	-0.54	3.64	-0.49	3.71	-0.42
Diazepam	3.95	4.13	0.18	4.12	0.17	4.19	0.24	4.11	0.16
7-Ethoxycoumarin	3.94	4.27	0.33	4.04	0.10	4.10	0.16	4.07	0.13
Amintriptyline	3.84	4.21	0.37	4.24	0.40	4.29	0.45	4.19	0.35
Styrene	3.75	3.54	-0.21	3.46	-0.29	3.47	-0.28	3.55	-0.19
Dextromethorphan	3.46	4.02	0.56	4.07	0.62	4.08	0.63	4.08	0.62
Meperidine	3.45	3.62	0.17	3.59	0.14	3.62	0.17	3.70	0.25
Lidocaine	3.27	3.16	-0.11	3.25	-0.02	3.17	-0.10	3.12	-0.15
S-Mephenytoin	3.25	3.64	0.39	3.59	0.34	3.62	0.37	3.70	0.45
Carbamazepine	3.16	3.38	0.22	3.36	0.20	3.28	0.12	3.31	0.15
Ifosfamide	2.76	3.05	0.29	3.16	0.39	3.08	0.31	2.92	0.16
SM-12502	2.75	3.62	0.87	3.59	0.83	3.60	0.85	3.69	0.94
1,2-Dibromoethane	2.01	2.07	0.06	2.05	0.03	2.00	-0.01	2.13	0.11
r^2		0.82		0.82		0.82		0.84	
RMSD			0.40		0.41		0.41		0.38
Max			0.87		0.83		0.85		0.94
Average			0.35		0.34		0.35		0.30
Std. Dev.			0.21		0.24		0.21		0.24
q^2								0.66	

Table 3. Experimentally Observed pK_m Values, Predicted pK_m Values by Hypo A, Hypo B and Hypo C and SVM Model for Those Compounds in the Test Set, the Corresponding Residual Values (Δ) and Their Associated Statistic Numbers (Correlation Coefficient) r^2, Root-Mean-Square Deviation (rmsd), Maximum Residual, Average Residual, and Standard Deviation of Residual

Molecules	Obs.	Hypo A		Hypo B		Hypo C		SVM Model	
	pK_m	pK_m	Δ	pK_m	Δ	pK_m	Δ	pK_m	Δ
Desmethylrabeprazole-thioether	5.42	5.35	-0.07	4.85	-0.57	4.89	-0.53	5.56	0.14
8-Hydroxyefavirenz	4.70	4.29	-0.40	4.28	-0.41	4.29	-0.40	4.26	-0.43
S-Methadone	4.52	3.82	-0.70	3.89	-0.64	3.89	-0.64	3.96	-0.57
R-Deprenyl	4.48	3.64	-0.84	3.60	-0.88	3.64	-0.84	3.72	-0.76
S-Ketamine	4.35	3.64	-0.72	3.60	-0.75	3.64	-0.72	3.72	-0.63
Midazolam	4.34	3.72	-0.62	3.77	-0.57	3.77	-0.57	3.87	-0.47
Testosterone	4.30	4.14	-0.15	3.92	-0.38	4.06	-0.24	3.97	-0.33
3-Cyano-7-ethoxycoumarin	4.15	4.20	0.05	4.04	-0.11	4.04	-0.11	4.06	-0.09
S-Mephobarbital	3.58	3.14	-0.44	3.11	-0.46	3.48	-0.10	3.38	-0.20
Imipramine	3.42	3.13	-0.29	3.13	-0.29	3.13	-0.29	3.04	-0.38
Cyclophosphamide	2.72	2.92	0.20	2.82	0.10	3.03	0.31	2.61	-0.11
r^2		0.78		0.86		0.81		0.87	
RMSD			0.49		0.52		0.49		0.43
Max			0.84		0.88		0.84		0.76
Average			0.41		0.47		0.43		0.37
Std. Dev.			0.28		0.25		0.25		0.22

Table 4. Costs of Returned Hypotheses and Null Hypotheses and the Cost Differences (Δ) between the Former and the Later for the Pharmacophore Models Hypo A, Hypo B, and Hypo C

Cost	Hypo A	Hypo B	Hypo C
Null hypothesis	207.60	207.60	207.60
Returned hypothesis	137.26	136.07	136.10
Δ	70.34	71.53	71.50

among these three models can be facilitated by superimposing these three models as demonstrated in Fig. (**2**), in which Hypo A takes the seemingly opposite orientation as compared with Hypo B and Hypo C and the later two are topologically closer to each other. Such differences can be even more pronounced when fitting these three models into the most active compound pentoxyresorufin in the training set as shown in Fig. (**3**), in which pentoxyresorufin adopts various conformations to generate the best fit with these models, manifesting the fact that a PhE is needed to accommodate the conformational plasticity.

The maximal errors of Hypo A, Hypo B, and Hypo C in the training set, were unanimously resulted from the predictions of SM-12502 with residuals of 0.87, 0.83, and 0.85, respectively (Table **2**). Similarly, Hypo B perfectly predicted clotiazepam without any error, and Hypo A and Hypo C merely yielded small errors of 0.01 and 0.20, respectively. Generally, these three models gave rise to very close prediction trend as demonstrated in Fig. (**4**), resulting in the same correlation coefficient r^2 value of 0.82, and can be further exhibited by the scatter plot of observed *vs*. the predicted pK_m values as illustrated in Fig. (**4**).

Additionally, it can be observed from Table **2** that other statistical numbers, namely root-mean-square deviation, maximal residual, average residual, and standard deviation of residuals show little discrepancies among these three models,

Table 5. Chemical Features, Weights, Tolerances, Three-Dimensional Coordinates and Interfeature Distances of Pharmacophore Models Hypo A, Hypo B and Hypo C

		HBA Lipid		HBA Lipid		HYDROPHOB Aromatic	HYDROPHOBIC
Hypo A	Weights	2.24		2.24		1.57	3.58
	Tolerances	1.60	2.20	1.30	1.90	1.30	1.30
	Coords: X	-4.01	-4.27	-3.36	-1.31	0.02	4.00
	Coords: Y	1.36	3.28	5.40	4.95	0.64	-1.46
	Coords: Z	-0.64	1.67	-4.55	-6.69	0.00	-0.69
		O---------->		O---------->		O	O
	HBA lipid						
		3.0					
	HBA lipid	5.7	6.6				
		7.5	9.0	3.0			
	HYDROPHOB aromatic	4.1	5.3	7.4	8.1		
	HYDROPHOBIC	8.5	9.8	10.8	10.3	4.5	
Hypo B	Weights	1.55		2.21		1.55	3.54
	Tolerances	1.30	1.90	1.30	2.05	1.45	1.30
	Coords: X	-6.41	-6.81	-2.02	-1.35	0.97	3.43
	Coords: Y	1.25	0.38	0.90	3.03	1.26	0.31
	Coords: Z	0.91	3.75	-2.15	-4.16	0.13	3.86
		O---------->		O---------->		O	O
	HBA lipid						
		3.0					
	HBA lipid	5.4	7.6				
		7.4	10.0	3.0			
	HYDROPHOB aromatic	7.4	8.6	3.8	5.2		
	HYDROPHOBIC	10.3	10.2	8.1	9.7	4.6	
Hypo C	Weights	1.63		2.33		1.63	3.03
	Tolerances	1.45	1.90	1.30	2.05	1.30	1.30
	Coords: X	-6.47	-6.75	-2.03	-1.34	0.97	3.43
	Coords: Y	1.36	0.27	0.82	3.10	1.26	-0.09
	Coords: Z	0.94	3.72	-2.24	-4.07	0.13	3.86
		O---------->		O---------->		O	O
	HBA lipid						
		3.0					
	HBA lipid	5.5	7.6				
		7.4	9.9	3.0			
	HYDROPHOB aromatic	7.5	8.6	3.8	5.1		
	HYDROPHOBIC	10.4	10.2	8.2	9.8	4.7	

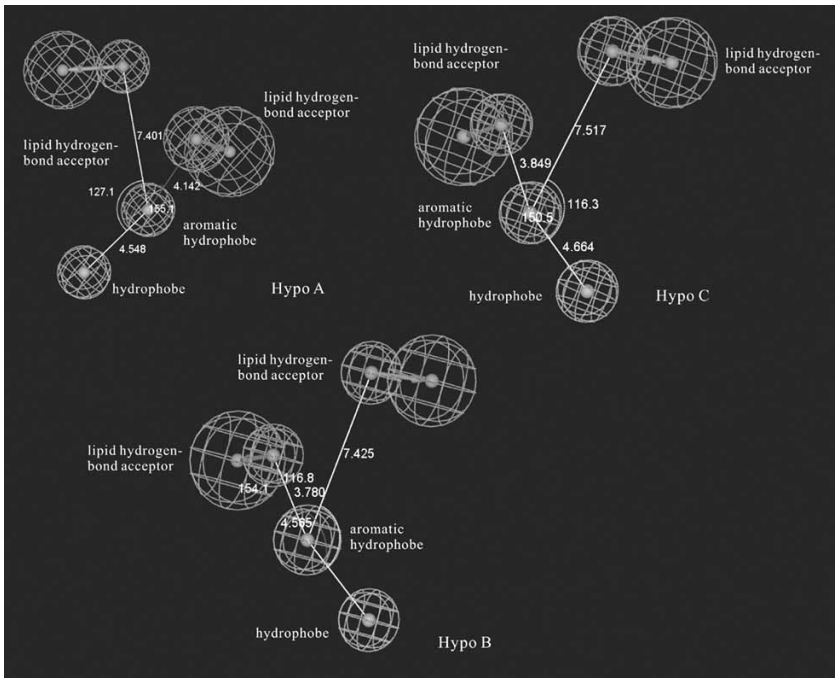

Fig. (1). Three pharmacophore candidate models Hypo A, Hypo B, and Hypo C in the PhE, possessing hydrogen-bond acceptor lipid, aromatic hydrophobe, and hydrophobe chemical features. The interfeature distances and angles among features are measured in Ångstroms and degrees, respectively.

suggesting that these three models functioned equally well in the training set. More importantly, the cost differences between the null hypothesis and returned hypotheses are 70.34, 71.53, and 71.50 for Hypo A, Hypo B, and Hypo C, respectively (Table **4**), all of which are larger than 60 that is the cost difference required to reach the level of a more than 90% chance to show the statistical correlation between the hypothesis and the input data as described in the *Catalyst*'s manual.

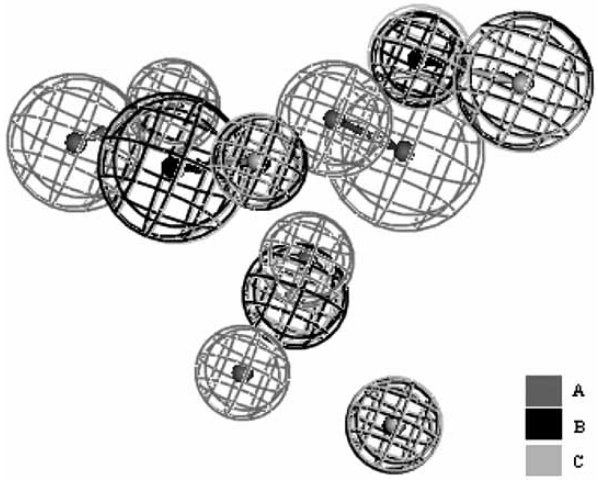

Fig. (2). Superimposition of three pharmacophore models Hypo A, Hypo B, and Hypo C.

When applied these three models in the PhE to predict those molecules in the test set, they showed slightly degraded performances in terms of statistical parameters, namely rmsd, average residual, and residual standard devia

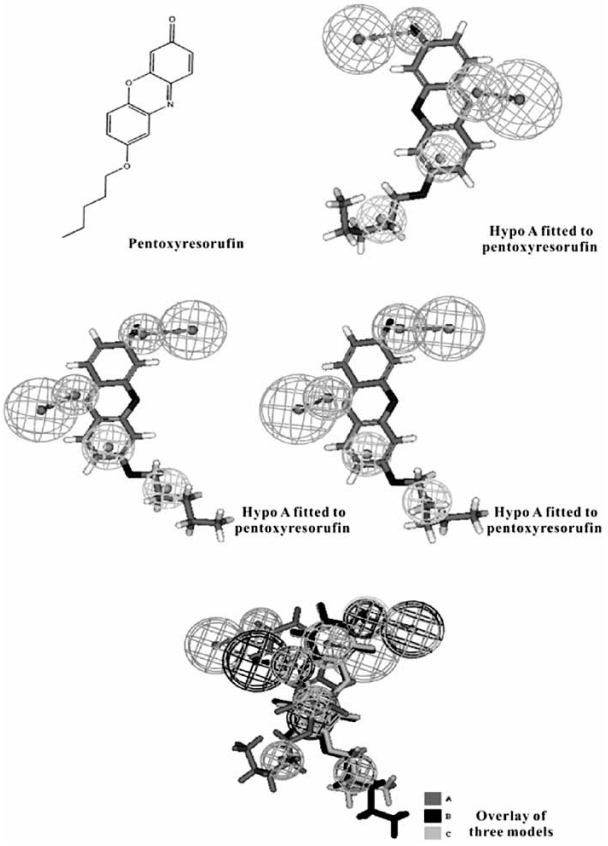

Fig. (3). Chemical structure of pentoxyresorufin, pharmacophore models Hypo A, Hypo B, and Hypo C fitted to pentoxyresorufin and overlay of these three models. The chemical features are described in Fig. (**1**).

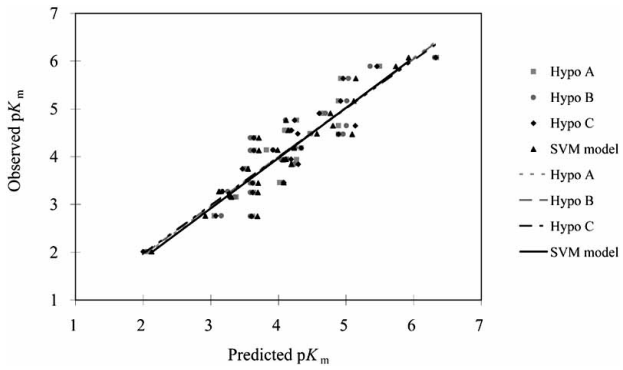

Fig. (4). The observed pK_m *vs*. the pK_m predicted by Hypo A, Hypo B, Hypo C, and SVM model for those molecules in the training set and their corresponding linear regression lines.

tion as shown in Table **3**, except Hypo B, whose r^2 values showed moderate improvement from 0.82 in the training set to 0.86 in the test set. Like maximal prediction errors in the training set, the predictions of *R*-deprenyl by these three models unanimously resulted in the maximal deviations from the observed value with residuals of 0.84, 0.88, and 0.84 by Hypo A, Hypo B, and Hypo C, respectively. Hypo A merely gave rise to a deviation error of 0.07 for the calculation of desmethylrabeprazole-thioether, whose residuals, nevertheless, were markedly 0.57 and 0.53 by Hypo B and Hypo C, respectively. Similar observation can be found for the prediction of *S*-mephobarbital, whose residuals were 0.44 and 0.46 by Hypo A and Hypo B, respectively, while substantially reduced to 0.10 by Hypo C. These variations in the prediction errors by these three models resulted in discrepancies in the prediction trends in the test set, which can be further demonstrated by Fig. (**5**) as compared with Fig. (**4**).

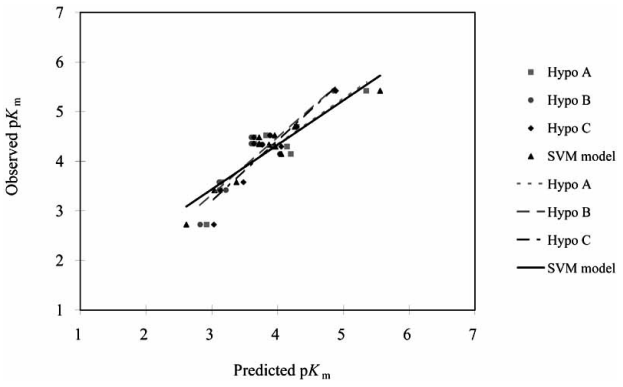

Fig. (5). The observed pK_m *vs*. the pK_m predicted by Hypo A, Hypo B, Hypo C, and SVM model for those molecules in the test set and their corresponding linear regression lines.

In general, the predictions of all molecules in both the training set and the test set are in good agreement with observed values as shown in Tables **2** and **3**. In addition, the insignificant r^2 differences between the training set and the test set by these three models suggest that they were statistically well-trained models in contrast to an over-trained model, which otherwise will result in a substantial r^2 difference between the training set and the test set. Based on the performances in the training set and the test set, it can be asserted that Hypo A, Hypo B, and Hypo C are eligible to be enlisted to compose the PhE.

The optimal SVM, generated from those hypotheses in the ensemble, was chosen from various runtime conditions based on the prediction calculations of those molecules in the training set, as shown in Table **2**, as well as its 10-fold cross-validation. Table **6** lists the input parameters for the final optimal SVM model. It can be found that the SVM model yielded smaller residuals than any members in the PhE for most of molecules in the training set. The prediction of ifosfamide by SVM, for example, showed an error of 0.16, while that generated relatively substantial residuals of 0.29, 0.39, and 0.31 by Hypo A, Hypo B, and Hypo C, respectively. The maximal prediction deviation in the training set by the SVM model was resulted from the estimation of SM-12502, which also gave rise to the maximal errors by Hypo A, Hypo B, and Hypo C except the fact that the former is slightly larger than the latter ones (0.94 *vs*. 0.87, 0.83, and 0.85, respectively). On the other hand, of 26 molecules in the training set, there were only 5 molecules in which the SVM model showed slightly worse performances than those three models in the PhE.

Table 6. The Optimal Runtime Parameters for the Final SVM Model

Parameter	Value
SVM type	ε-SVR
kernel type	radial basis function
γ	0.25
Cost	66
ε	0.15

In general, SVM marginally functioned better than those three models in PhE in the training set as indicated by the correlation coefficients r^2, root-mean-square deviations, and average residuals (Table **2**) that can be further confirmed by Fig. (**4**). Furthermore, the insignificant difference between the correlation coefficients q^2 of 0.66, calculated by the 10-fold cross-validation, and the correlation coefficient r^2 of 0.84 for SVM (Table **2**) signifies that SVM model shows highly statistical significance between the generated model and the input data and warrants its authenticity consequently.

The maximal prediction variation of SVM in the test set was generated from the estimation of *R*-deprenyl with a value of 0.76, which also gave rise to the largest calculation errors by Hypo A, Hypo B, and Hypo C. The similar observation can also be found for the training set except the fact that the SVM model has the lowest maximal residual in the test set yet the highest one in the training set and the SVM model, in general, gave rise to the lowest residuals for most of molecules in the test set compared with Hypo A, Hypo B, and Hypo C except that there are a few molecules in the test set, whose residuals by the SVM model are between the largest one and the smallest one. The prediction of desmeth-

ylrabeprazole-thioether by the SVM model, for example, deviates from the experimental value by 0.14, which is only slightly worse than Hypo A (0.07) but much better than both Hypo B (0.57) and Hypo C (0.53). Of these 4 models, it can be concluded that the SVM generated the lowest residuals or the second lowest residuals for all the molecules in the test set except imipramine, whose prediction residual is 0.29 by Hypo A, Hypo B, and Hypo C yet 0.38 by the SVM model.

It can be observed that Hypo A, Hypo B, Hypo C, and the SVM model showed little performance discrepancies between the training set and the test set as demonstrated by Tables **2** and **3**. The correlation coefficient r^2 values of Hypo B and SVM, for example, only marginally increase (0.04 and 0.03) from the training set to the test set, respectively, and those of Hypo A and Hypo C insignificantly decrease (0.04 and 0.01, respectively), manifesting the fact that these 4 models were well-trained or no over-training was observed as mentioned above. Nevertheless, it can be clearly seen that the maximum residual of SVM was substantially reduced from the training set (0.94) to the test set (0.76), which was approximately maintained the same level in both the training set and the test set by Hypo A, Hypo B, and Hypo C. Similarly, of these 4 models, the SVM model also has the lowest rmsd, average residual, and residual standard deviation in the test set, suggesting that the SVM model is the best prediction model compared with the other models in the test set based on those parameters mentioned above.

It is very important and practical to evaluate the performance of an *in silico* model not only in the training set but in the test set. As a result, it can be concluded that the SVM model is superior to Hypo A, Hypo B, and Hypo C based on the prediction and statistical performances in both sets mentioned above presumably because of the fact that the PhE/SVM approach cannot only take into account the protein conformational flexibility and but also gives rise to the more realistic final model and that usually can not be achieved by traditional ligand-based modeling schemes. Furthermore, in case some structurally distinct substrates which interact with different protein conformations not taken into account in the model generation, i.e. different pharmacophore models, are to be studied, the PhE/SVM scheme can provide additional capacity for robust and fast adaptation to accommodate such extension of structure domain by simply implementing the corresponding pharmacophore models generated from those novel molecules into the ensemble and generating a new SVM model. As a result, the PhE/SVM scheme only requires small fraction of computational time compared with the initial model construction. In contrast, traditional ligand-based modeling methods will demand new model generations and require at least the same amount of computational time compared with the initial model building to accommodate such adaptation. More importantly, it can be expected that the predictability of a new model consequently will be deteriorated since they have to compromise accuracy in favor of extension of the conformational domain resulted from augmentation of structurally distinct molecules.

In addition to accurate prediction capacity, the generated PhE/SVM model can be utilized as an effective filter to remove any compounds, which may potentially interact with CYP2B6 and cause adverse interactions, from a large compound library since the PhE/SVM scheme is primarily based on PhE and the pharmacophore is believed to be an efficient and fast screening tool [73].

In addition to understanding the binding affinity variations among those substrates using this novel analogue-based scheme, the enzyme-substrate interaction details are expected. Consequently, this investigation will be further extended in the future using the *Cell-PLoc* package [65], for example, to predict its subcellular location based on its amino acid sequence in the hope that the interactions between the substrates and 2B6 enzyme can be elucidated.

CONCLUSIONS

This PhE/SVM scheme, which was generated using the support vector machine based on pharmacophore ensemble to take into consideration protein plasticity while interacting with structurally diverse molecules, was successfully applied to predict CYP2B6-substrate interactions for those 26 and 11 molecules in the training set and the test set, respectively, and showed excellent statistical significance. Other than prediction, the PhE/SVM also can be employed as a primary filter to conduct virtual screening for compound libraries in order to eliminate those compounds with a potential to cause adverse effects with the cytochrome 2B6 enzyme because of the speed nature of this PhE/SVM scheme. In contrast to the traditional ligand-based modeling approaches, this PhE/SVM scheme also provides excellent adaptability without significantly losing accuracy to accommodate more diverse protein conformations once more structurally novel molecules are to be taken into account. It can be further expected that this PhE/SVM scheme can work well to model protein-substrate interactions of other CPY450s, which are conformationally highly flexible *per se* while interacting with structurally diverse substrates.

ACKNOWLEDGEMENT

This work was supported by the National Science Council, Taiwan. Parts of calculations were performed at the National Center for High-Performance Computing, Taiwan. The authors wish to thank all the anonymous reviewers for their helpful comments.

REFERENCES

[1] Lewis, D.F.; Tamburini, P.P.; Gibson, G.G. *Chem. Biol. Interact.,* **1986**, *58*, 289.
[2] Roberts, G.F.; Mehta, I.; Murray, M. *Mol. Pharmacol.,* **1989**, *35*, 736.
[3] Forrester, L.M.; Henderson, C.J.; Glance, M.J.; Back, D.J.; Park, B.K.; Ball, S.E.; Kitteringham, N.R.; McLaren, A.W.; Miles, J.S.; Skett, P. *Biochem. J.,* **1992**, *281*, 359.
[4] Lewis, D.F. *Curr. Med. Chem.,* **2003**, *10*, 1955.
[5] Micuda, S.; Mundlova, L.; Anzenbacherova, E.; Anzenbacher, P.; Chladek, J.; Fuksa, L.; Martinkova, J. *Eur. J. Clin. Pharmacol.,* **2004**, *60*, 583.
[6] Villeneuve, J.P.; Pichette, V. *Curr. Drug Metab.,* **2004**, *5*, 273.
[7] Niwa, T.; Shiraga, T.; Ishii, I.; Kagayama, A.; Takagi, A. *Biol. Pharm. Bull.,* **2005**, *28*, 1711.
[8] VandenBranden, M.; Wrighton, S.A.; Ekins, S.; Gillespie, J.S.; Binkley, S.N.; Ring, B.J.; Gadberry, M.G.; Mullins, D.C.; Strom, S.C.; Jensen, C.B. *Drug Metab. Dispos.,* **1998**, *26*, 1063.
[9] Walsky, R.L.; Obach, R.S. *Drug Metab. Dispos.,* **2004**, *32*, 647.
[10] Domanski, T.L.; Halpert, J.R. *Curr. Drug Metab.,* **2001**, *2*, 117.
[11] Lewis, D.F.V.; Sheridan, G. *Sci. World J.,* **2001**, *1*, 151.

[12] Roymans, D.; Annaert, P.; Van Houdt, J.; Weygers, A.; Noukens, J.; Sensenhauser, C.; Silva, J.; Van Looveren, C.; Hendrickx, J.; Mannens, G.; Meuldermans, W. *Drug Metab. Dispos.,* **2005**, *33*, 1004.

[13] Lewis, D.F.V.; Modi, S.; Dickins, M. *Drug Metab. Rev.,* **2002**, *34*, 69.

[14] Lewis, D.F.V.; Dickins, M. *Drug Discov. Today,* **2002**, *7*, 918.

[15] Gerber, J.G.; Rhodes, R.J.; Gal, J. *Chirality,* **2004**, *16*, 36.

[16] Lewis, D.F.; Ito, Y.; Goldfarb, P.S. *Curr. Med. Chem.,* **2006**, *13*, 2645.

[17] Lewis, D.F. *Exp. Toxicol. Pathol.,* **1999**, *51*, 369.

[18] Rendic, S.; Di Carlo, F.J. *Drug Metab. Rev.,* **1997**, *29*, 413.

[19] Shimada, T.; Yamazaki, H.; Mimura, M.; Inui, Y.; Guengerich, F.P. *J. Pharmacol. Exp. Ther.,* **1994**, *270*, 414.

[20] Lewis, D.F.V.; Lake, B.G.; Dickins, P.M.; Eddershaw, J.; Tabbit, M.H.; Goldfarb, P.S. *Xenobiotica,* **1999**, *29*, 361.

[21] Hanna, I.H.; Reed, J.R.; Guengerich, F.P.; Hollenberg, P.F. *Arch. Biochem. Biophys.,* **2000**, *376*, 206.

[22] Mimura, M.; Baba, T.; Yamazaki, H.; Ohmori, S.; Inui, Y.; Gonzalez, F.J.; Guengerich, F.P.; Shimada, T. *Drug Metab. Dispos.,* **1993**, *21*, 1048.

[23] Cho, J.-Y.; Lim, H.-S.; Chung, J.-Y.; Yu, K.-S.; Kim, J.-R.; Shin, S.-G.; Jang, I.-J. *Drug Metab. Dispos.,* **2004**, *32*, 1341.

[24] Sridar, C.; Kent, U.M.; Notley, L.M.; Gillam, E.M.J.; Hollenberg, P.F. *J. Pharmacol. Exp. Ther.,* **2002**, *301*, 945.

[25] Code, E.L.; Crespi, C.L.; Penman, B.W.; Gonzalez, F.J.; Chang, T.K.; Waxman, D.J. *Drug Metab. Dispos.,* **1997**, *25*, 985.

[26] Ekins, S.; VandenBranden, M.; Ring, B.J.; Wrighton, S.A. *Pharmacogenetics,* **1997**, *7*, 165.

[27] Ekins, S.; Vandenbranden, M.; Ring, B.J.; Gillespie, J.S.; Yang, T.J.; Gelboin, H.V.; Wrighton, S.A. *J. Pharmacol. Exp. Ther.,* **1998**, *286*, 1253.

[28] Ekins, S.; Wrighton, S.A. *Drug Metab. Rev.,* **1999**, *31*, 719.

[29] Domanski, T.L.; Schultz, K.M.; Roussel, F.; Stevens, J.C.; Halpert, J.R. *J. Pharmacol. Exp. Ther.,* **1999**, *290*, 1141.

[30] Chou, K.-C. *Anal. Biochem.,* **1996**, *233*, 1.

[31] Chou, K.-C. *Curr. Med. Chem.,* **2004**, *11*, 2105.

[32] Chou, K.-C.; Wei, D.Q.; Du, Q.S.; Sirois, S.; Zhong, W.Z. *Curr. Med. Chem.,* **2006**, *13*, 3263.

[33] Chou, K.-C. In *Frontiers in Medicinal Chemistry*. Atta-ur-Rahman; Reitz, A.B. Eds.; Bentham Science Publishers: The Netherlands, **2006**; Vol. *3*, p. 455.

[34] Williams, P.A.; Cosme, J.; Sridhar, V.; Johnson, E.F.; McRee, D.E. *Mol. Cell.,* **2000**, *5*, 121.

[35] Scott, E.E.; He, Y.A.; Wester, M.R.; White, M.A.; Chin, C.C.; Halpert, J.R.; Johnson, E.F.; Stout, C.D. *Proc. Natl. Acad. Sci. USA,* **2003**, *100*, 13196.

[36] Williams, P.A.; Cosme, J.; Ward, A.; Angove, H.C.; Vinkovi, D.M.; Jhoti, H. *Nature,* **2003**, *424*, 464.

[37] Schoch, G.A.; Yano, J.K.; Wester, M.R.; Griffin, K.J.; Stout, C.D.; Johnson, E.F. *J. Biol. Chem.,* **2004**, *279*, 9497.

[38] Wester, M.R.; Yano, J.K.; Schoch, G.A.; Yang, C.; Griffin, K.J.; Stout, C.D.; Johnson, E.F. *J. Biol. Chem.,* **2004**, *279*, 35630.

[39] Williams, P.A.; Cosme, J.; Vinkovi, D.M.; Ward, A.; Angove, H.C.; Day, P.J.; Vonrhein, C.; Tickle, I.J.; Jhoti, H. *Science,* **2004**, *305*, 683.

[40] Yano, J.K.; Wester, M.R.; Schoch, G.A.; Griffin, K.J.; Stout, C.D.; Johnson, E.F. *J. Biol. Chem.,* **2004**, *279*, 38091.

[41] Yano, J.K.; Hsu, M.-H.; Griffin, K.J.; Stout, C.D.; Johnson, E.F. *Nat. Struct. Mol. Biol.,* **2005**, *12*, 822.

[42] Wang, J.-F.; Dong-Qing, W.; Li, L.; Zheng, S.-Y.; Li, Y.-X.; Chou, K.-C. *Biochem. Biophys. Res. Commun.,* **2007**, *355*, 513.

[43] Wang, Q.; Halpert, J.R. *Drug Metab. Dispos.,* **2002**, *30*, 86.

[44] Bersuker, I.B.; Leong, M.K.; Boggs, J.E.; Pearlman, R.S. *Int. J. Quantum Chem.,* **1997**, *63*, 1051.

[45] Friesner, R.A.; Guallar, V. *Annu. Rev. Phys. Chem.,* **2005**, *56*, 389.

[46] Poulos, T.L. *Proc. Natl. Acad. Sci. USA,* **2003**, *100*, 13121.

[47] Wester, M.R.; Johnson, E.F.; Marques-Soares, C.; Dansette, P.M.; Mansuy, D.; Stout, C.D. *Biochemistry,* **2003**, *42*, 6370.

[48] Wester, M.R.; Johnson, E.F.; Marques-Soares, C.; Dijols, S.; Dansette, P.M.; Mansuy, D.; Stout, C.D. *Biochemistry,* **2003**, *42*, 9335.

[49] Zhao, Y.; White, M.A.; Muralidhara, B.K.; Sun, L.; Halpert, J.R.; Stout, C.D. *J. Biol. Chem.,* **2006**, *281*, 5973.

[50] Ekroos, M.; Sjögren, T. *Proc. Natl. Acad. Sci. USA,* **2006**, *103*, 13682.

[51] Hunter, P. *EMBO Rep.,* **2006**, *7*, 249.

[52] Lightfoot, T.; Ellis, S.W.; Mahling, J.; Ackland, M.J.; Blaney, F.E.; Bijloo, G.J.; de Groot, M.J.; Vermeulen, N.P.E.; Blackburn, G.M.; Lennard, M.S.; Tucker, G.T. *Xenobiotica,* **2000**, *30*, 219.

[53] Lehnert, U.; Xia, Y.; Royce, T.E.; Goh, C.-S.; Liu, Y.; Senes, A.; Yu, H.; Zhang, Z.L.; Engelman, D.M.; Gerstein, M. *Q. Rev. Biophys.,* **2004**, *37*, 121.

[54] Ekins, S.; Bravi, G.; Ring, B.J.; Gillespie, T.A.; Gillespie, J.S.; Vandenbranden, M.; Wrighton, S.A.; Wikel, J.H. *J. Pharmacol. Exp. Ther.,* **1999**, *288*, 21.

[55] de Groot, M.J.; Kirton, S.B.; Sutcliffe, M.J. *Curr. Top. Med. Chem.,* **2004**, *4*, 1803.

[56] Chou, K.-C.; Shen, H.-B. *Biochem. Biophys. Res. Commun.,* **2006**, *347*, 150.

[57] Shen, H.-B.; Chou, K.-C. *Biochem. Biophys. Res. Commun.,* **2007**, *355*, 1006.

[58] Chou, K.-C.; Shen, H.-B. *J. Proteome Res.,* **2006**, *5*, 1888

[59] Chou, K.-C.; Shen, H.-B. *J. Proteome Res.,* **2007**, *6*, 1728

[60] Shen, H.-B.; Chou, K.-C. *Biopolymers,* **2006**, *85*, 233.

[61] Chou, K.-C.; Shen, H.-B. *Biochem. Biophys. Res. Commun.,* **2007**, *360*, 339.

[62] Shen, H.-B.; Chou, K.-C. *Biochem. Biophys. Res. Commun.,* **2007**, *364*, 53.

[63] Shen, H.-B.; Chou, K.-C. *Biochem. Biophys. Res. Commun.,* **2007**, *363*, 297.

[64] Chou, K.-C.; Shen, H.-B. *Biochem. Biophys. Res. Commun.,* **2007**, *357*, 633.

[65] Chou, K.-C.; Shen, H.-B. *Anal. Biochem.,* **2007**, *370*, 1.

[66] Shen, H.-B.; Yang, J.; Chou, K.-C. *Expert. Rev. Proteomics.,* **2007**, *4*, 453.

[67] Leong, M.K. *Chem. Res. Toxicol.,* **2007**, *20*, 217.

[68] Chang, G.; Guida, W.C.; Still, W.C. *J. Am. Chem. Soc.,* **1989**, *111*, 4379.

[69] Kolossvary, I.; Guida, W.C. *J. Am. Chem. Soc.,* **1996**, *118*, 5011.

[70] Still, W.C.; Tempczyk, A.; Hawley, R.C.; Hendrickson, T. *J. Am. Chem. Soc.,* **1990**, *112*, 6127.

[71] Halgren, T.A. *J. Comput. Chem.,* **1996**, *17*, 490.

[72] Sprague, P.W. *Perspect. Drug Discov. Des.,* **1995**, *3*, 1.

[73] Güner, O.F. *Pharmacophore Perception, Development, and Use in Drug Design*. International University Line: La Jolla, California, **2000**.

[74] Chang, C.-C.; Lin, C.-J. *LIBSVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm*, 2001.

[75] Chou, K.-C.; Zhang, C.T. *Crit. Rev. Biochem. Mol. Biol.,* **1995**, *30*, 275.

[76] Zhou, G.P. *J. Protein Chem.,* **1998**, *17*, 729.

[77] Zhou, G.-P.; Doctor, K. *Proteins,* **2003**, *50*, 44.

[78] Pan, Y.-X.; Zhang, Z.-Z.; Guo, Z.-M.; Feng, G.-Y.; Huang, Z.-D.; He, L. *J. Protein Chem.,* **2003**, *22*, 395.

[79] Huang, Y.; Li, Y. *Bioinformatics,* **2004**, *20*, 21.

[80] Chen, C.; Zhou, X.; Tian, Y.; Zou, X.; Cai, P. *Anal. Biochem.,* **2006**, *357*, 116.

[81] Chen, C.; Tian, Y.-X.; Zou, X.-Y.; Cai, P.-X.; Mo, J.-Y. *J. Theor. Biol.,* **2006**, *243*, 444.

[82] Zhang, S.W.; Pan, Q.; Zhang, H.C.; Shao, Z.C.; Shi, J.Y. *Amino Acids,* **2006**, *30*, 461.

[83] Du, P.; Li, Y. *BMC Bioinformatics,* **2006**, *7*, 518.

[84] Mondal, S.; Bhavna, R.; Mohan Babu, R.; Ramakumar, S. *J. Theor. Biol.,* **2006**, *243*, 252.

[85] Lin, H.; Li, Q.-Z. *Biochem. Biophys. Res. Commun.,* **2007**, *354*, 548.

[86] Lin, H.; Li, Q.-Z. *J. Comput. Chem.,* **2007**, *28*, 1463.

[87] Shi, J.Y.; Zhang, S.W.; Pan, Q.; Cheng, Y.M.; Xie, J. *Amino Acids,* **2007**, *33*, 69.

[88] Cao, Y.; Liu, S.; Zhang, L.; Qin, J.; Wang, J.; Tang, K. *BMC Bioinformatics,* **2006**, *7*, 20.

[89] Zhang, Z.-H.; Wang, Z.-H.; Zhang, Z.-R.; Wang, Y.-X. *FEBS Lett.,* **2006**, *580*, 6169.

[90] Gao, Q.-B.; Wang, Z.-Z.; Yan, C.; Du, Y.-H. *FEBS Lett.,* **2005**, *579*, 3444.

[91] Gao, Q.-B.; Wang, Z.-Z. *Protein Eng. Des. Sel.,* **2006**, *19*, 511.

[92] Guo, J.; Lin, Y.; Liu, X. *Proteomics,* **2006**, *6*, 5099.

[93] Kedarisetti, K.D.; Kurgan, L.; Dick, S. *Biochem. Biophys. Res. Commun.,* **2006**, *348*, 981.

[94] Jahandideh, S.; Abdolmaleki, P.; Jahandideh, M.; Asadabadi, E.B. *Biophys. Chem.,* **2007**, *128*, 87.
[95] Chen, Y.-L.; Li, Q.-Z. *J. Theor. Biol.,* **2007**, *248*, 377.
[96] Zhou, X.-B.; Chen, C.; Li, Z.-C.; Zou, X.-Y. *J. Theor. Biol.,* **2007**, *248*, 546.
[97] Diao, Y.; Li, M.; Feng, Z.; Yin, J.; Pan, Y. *J. Theor. Biol.,* **2007**, *247*, 608.
[98] Diao, Y.; Ma, D.; Wen, Z.; Yin, J.; Xiang, J.; Li, M. *Amino Acids,* **2007**, *34*, 111.
[99] Li, F.M.; Li, Q.Z. *Amino Acids,* **2007**, *34*, 119.
[100] Fang, Y.; Guo, Y.; Feng, Y.; Li, M. *Amino Acids,* **2007**, *34*, 103.
[101] Du, Q.; Wei, D.; Chou, K.-C. *Peptides,* **2003**, *24*, 1863.
[102] Du, Q.-S.; Jiang, Z.-Q.; He, W.-Z.; Li, D.-P.; Chou, K.-C. *J. Biomol. Struct. Dyn.,* **2006**, *23*, 635.
[103] Cai, Y.-D.; Chou, K.-C. *Biochem. Biophys. Res. Commun.,* **2003**, *305*, 407.
[104] Wang, M.; Yang, J.; Liu, G.-P.; Xu, Z.-J.; Chou, K.-C. *Protein Eng. Des. Sel.,* **2004**, *17*, 509.
[105] Cai, Y.-D.; Zhou, G.-P.; Chou, K.-C. *J. Theor. Biol.,* **2005**, *234*, 145.
[106] Gao, Y.; Shao, S.; Xiao, X.; Ding, Y.; Huang, Y.; Huang, Z.; Chou, K.-C. *Amino Acids,* **2005**, *28*, 373.
[107] Xiao, X.; Shao, S.; Ding, Y.; Huang, Z.; Chou, K.-C. *Amino Acids,* **2006**, *30*, 49.
[108] Cai, Y.D.; Chou, K.C. *J. Proteome Res.,* **2005**, *4*, 967.
[109] Shen, H.-B.; Chou, K.-C. *Biochem. Biophys. Res. Commun.,* **2005**, *334*, 288.
[110] Liu, H.; Yang, J.; Wang, M.; Xue, L.; Chou, K.-C. *Prot. J.,* **2005**, *24*, 385.
[111] Chou, K.-C.; Shen, H.-B. *J. Cell. Biochem.,* **2006**, *99*, 517.
[112] Xiao, X.; Shao, S.; Ding, Y.; Huang, Z.; Huang, Y.; Chou, K.-C. *Amino Acids,* **2005**, *28*, 57.
[113] Sun, X.D.; Huang, R.B. *Amino Acids,* **2006**, *30*, 469.
[114] Guo, Y.Z.; Li, M.; Lu, M.; Wen, Z.; Wang, K.; Li, G.; Wu, J. *Amino Acids,* **2006**, *30*, 397.
[115] Wen, Z.; Li, M.; Li, Y.; Guo, Y.; Wang, K. *Amino Acids,* **2007**, *32*, 277.
[116] Shen, H.-B.; Chou, K.-C. *Amino Acids,* **2007**, *32*, 483.
[117] Liu, D.Q.; Liu, H.; Shen, H.-B.; Yang, J.; Chou, K.-C. *Amino Acids,* **2007**, *32*, 493.
[118] Chen, J.; Liu, H.; Yang, J.; Chou, K.-C. *Amino Acids,* **2007**, *33*, 423.
[119] Shen, H.-B.; Yang, J.; Chou, K.-C. *Amino Acids,* **2007**, *33*, 57.
[120] Zhang, T.L.; Ding, Y.S. *Amino Acids,* **2007**, *33*, 623.
[121] Ding, Y.-S.; Zhang, T.-L.; Chou, K.-C. *Protein Peptide Lett.,* **2007**, *14*, 811.
[122] Niu, B.; Cai, Y.-D.; Lu, W.-C.; Li, G.-Z.; Chou, K.-C. *Protein Peptide Lett.,* **2006**, *13*, 489.
[123] Wang, M.; Yang, J.; Chou, K.-C. *Amino Acids,* **2005**, *28*, 395.
[124] Xiao, X.; Chou, K.-C. *Protein Peptide Lett.,* **2007**, *14*, 871.
[125] Tan, F.; Feng, X.; Fang, Z.; Li, M.; Guo, Y.; Jiang, L. *Amino Acids,* **2007**, *33*, 669.
[126] Miura, M.; Satoh, S.; Tada, H.; Habuchi, T.; Suzuki, T. *Eur. J. Clin. Pharmacol.,* **2005**, *62*, 821.
[127] Ghosal, A.; Hapangama, N.; Yuan, Y.; Lu, X.; Horne, D.; Patrick, J.E.; Zbaida, S. *Biopharm. Drug Dispos.,* **2003**, *24*, 375.
[128] Ward, B.A.; Gorski, J.C.; Jones, D.R.; Hall, S.D.; Flockhart, D.A.; Desta, Z. *J. Pharmacol. Exp. Ther.,* **2003**, *306*, 287.
[129] Stevens, J.C.; White, R.B.; Hsu, S.H.; Martinet, M. *J. Pharmacol. Exp. Ther.,* **1997**, *282*, 1389.
[130] Grace, J.M.; Aguilar, A.J.; Trotman, K.M.; Brewer, T.G. *Drug Metab. Dispos.,* **1998**, *26*, 313.
[131] Hidestrand, M.; Oscarson, M.; Salonen, J.S.; Nyman, L.; Pelkonen, O.; Turpeinen, M.; Ingelman-Sundberg, M. *Drug Metab. Dispos.,* **2001**, *29*, 1480.
[132] Usmani, K.; Rose, R.L.; Goldstein, J.A.; Taylor, W.G.; Brimfleld, A.A. *Drug Metab. Dispos.,* **2002**, *30*, 289.
[133] Yanagihara, Y.; Kariya, S.; Ohtani, M.; Uchino, K.; Aoyama, T.; Yamamura, Y.; Iga, T. *Drug Metab. Dispos.,* **2001**, *29*, 887.
[134] Kim, K.-A.; Chung, J.; Jung, D.-H.; Park, J.-Y. *Eur. J. Clin. Pharmacol.,* **2004**, *60*, 575.
[135] Kobayashi, K.; Abe, S.; Nakajima, M.; Shimada, N.; Tani, M.; Chiba, K.; Yamamoto, T. *Drug Metab. Dispos.,* **1999**, *27*, 1429.
[136] Ramirez, J.; Innocenti, F.; Schuetz, E.G.; Flockhart, D.A.; Relling, M.V.; Santucci, R.; Ratain, M.J. *Drug Metab. Dispos.,* **2004**, *32*, 930.
[137] Pearce, R.E.; Vakkalagadda, G.R.; Leeder, J.S. *Drug Metab. Dispos.,* **2002**, *30*, 1170.
[138] Huanga, Z.; Roya, P.; Waxman, D.J. *Biochem. Pharmacol.,* **2000**, *59*, 961.

Received: 18 December, 2007 **Revised: 17 April, 2008** **Accepted: 25 April, 2008**