Quantitative structure–activity relationship studies of mushroom tyrosinase inhibitors

Chao-Bin Xue · Wan-Chun Luo · Qi Ding ·
Shou-Zhu Liu · Xing-Xiang Gao

Abstract Here, we report our results from quantitative structure–activity relationship studies on tyrosinase inhibitors. Interactions between benzoic acid derivatives and tyrosinase active sites were also studied using a molecular docking method. These studies indicated that one possible mechanism for the interaction between benzoic acid derivatives and the tyrosinase active site is the formation of a hydrogen-bond between the hydroxyl (^{a}OH) and carbonyl oxygen atoms of Tyr^{98}, which stabilized the position of Tyr^{98} and prevented Tyr^{98} from participating in the interaction between tyrosinase and ORF378. Tyrosinase, also known as phenoloxidase, is a key enzyme in animals, plants and insects that is responsible for catalyzing the hydroxylation of tyrosine into o-diphenols and the oxidation of o-diphenols into o-quinones. In the present study, the bioactivities of 48 derivatives of benzaldehyde, benzoic acid, and cinnamic acid compounds were used to construct three-dimensional quantitative structure–activity relationship (3D-QSAR) models using comparative molecular field (CoMFA) and comparative molecular similarity indices (CoMSIA) analyses. After superimposition using common substructure-based alignments, robust and predictive 3D-QSAR models were obtained from CoMFA ($q^2 = 0.855$, $r^2 = 0.978$) and CoMSIA ($q^2 = 0.841$, $r^2 = 0.946$), with 6 optimum components. Chemical descriptors, including electronic (Hammett σ), hydrophobic (π), and steric (MR) parameters, hydrogen bond acceptor (H-acc), and indicator variable (I), were used to construct a 2D-QSAR model. The results of this QSAR indicated that π, MR, and H-acc account for 34.9, 31.6, and 26.7% of the calculated biological variance, respectively. The molecular interactions between ligand and target were studied using a flexible docking method (FlexX). The best scored candidates were docked flexibly, and the interaction between the benzoic acid derivatives and the tyrosinase active site was elucidated in detail. We believe that the QSAR models built here provide important information necessary for the design of novel tyrosinase inhibitors.

Keywords Benzoic acid derivatives · CoMFA · CoMSIA · Molecular docking · QSAR · Tyrosinase inhibitors

C.-B. Xue · W.-C. Luo (✉) · S.-Z. Liu
College of Plant Protection, Key Laboratory of Pesticide Toxicology and Application Technique, Shandong Agricultural University, Tai′an 271018, P.R. China
e-mail: wcluo@sdau.edu.cn

Q. Ding
Safety Evaluation Center, Shenyang Research Institute of Chemical Industry, Shenyang 110021, P.R. China

X.-X. Gao
Institute of Plant Protection, Shandong Academy of Agricultural Sciences, Jinan 250100, P.R. China

Introduction

Tyrosinase (EC 1.14.18.1), also known as phenoloxidase (PO), is a copper-containing, mixed-function oxidase widely distributed in microorganisms, animals, and plants [1, 2] that is a key enzyme involved in the browning that occurs upon bruising or long-term storage of plants. Tyrosinase catalyzes both the hydroxylation of monophenols and the oxidation of o-diphenols into o-quinones, and is involved in the formation of pigments such as melanins [3]. Tyrosinase may play a role in human neuromelanin formation in the substantia nigra portion of the brain, it may also be central to dopamine neurotoxicity and may contribute to Parkinson

disease-related neurodegeneration [4]. In mammals, tyrosinase is responsible for skin pigmentation defects and abnormalities such as flecks [5]. In insects, tyrosinase is uniquely associated with three different physiologically important biochemical processes, including sclerotization of the insect cuticle, defensive encapsulation and melanization of foreign organisms, and wound healing [6]. Tyrosinase inhibitors have become increasingly important in medicinal [7], agricultural, and cosmetic [8] products, primarily in relation to its roles in hyperpigmentation. Accordingly, there is an urgent need to design and screen tyrosinase inhibitors with higher bioactivities, and to elucidate the mechanisms of inhibition.

We previously sought to determine the mechanisms underlying PO inhibition and kinetics in specific compounds, including kojic acid, apigenin, quercetin, 4-hexylresorcinol, and 4-dodecylresorcinol [9–13]. Three-dimensional quantitative structure–activity relationships (3D-QSAR) and the molecular docking of insect tyrosinase inhibitors with the active site of *Streptomyces* tyrosinase were also investigated [14], both of which are crucial to the design of effective inhibitory molecules. In order to provide further insight into potential tyrosinase inhibitors, we performed a 3D-QSAR analysis for mushroom tyrosinase inhibitors using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) [15, 16]. A 2D-QSAR analysis was also carried out by the Hansch–Fujita approach. Our results from these studies represent continued and extensive efforts to describe the QSAR of tyrosinase inhibitors.

Materials and methods

Data set and structures

Previously reported in vitro data from one series of benzaldehydes (CQ1-CQ19) [17, 18], one series of benzoic acids (CS1-CS26) [19–21], and one series of cinnamic acids (CS27-CS29) [22] were used in the study (see the structures in Chart shown in Table 1). The IC$_{50}$ values of 48 molecules were segregated into groups of 38 and 10 as a training and test set, respectively. The test compounds were selected manually such that diverse structures and a wide range of activity were included. IC$_{50}$ values were converted into pIC$_{50}$ (pIC$_{50}$ = −lgIC$_{50}$), and the pIC$_{50}$ value of each of 48 molecules were used for the construction of 3D-QSAR models. In addition, pIC$_{50}$ values of 37 compounds (CQ1-15, CQ19, CS1-6, CS9-17, and CS21-26) were used for the construction of 2D-QSAR models to determine physicochemical parameters in the absence of others compounds (Table 2). The structures and biological activities of the compounds are reported in Table 1.

Molecular modeling and alignment rule

Molecular modeling studies were performed using SYBYL 6.9 running on a Silicon Graphics Origin 300 server. The compounds CQ1-19 and CS1-29 were sketched, and each structure was energy-minimized using a conjugate gradient minimization algorithm with the Tripos force field until a gradient convergence of 0.001 kcal/(mol Å) was achieved. Gasteiger–Hückel charges were calculated for all compounds, and were used to construct CoMFA and CoMSIA models. A common substructure-based alignment was adopted that attempted to align molecules to the template molecule on a common backbone, as illustrated in Fig. 1. For manual alignment of the inhibitors, the structure of compound CQ16 (4-hexylbenzaldehyde) was used as a template.

CoMFA descriptors

CoMFA steric and electrostatic interaction fields were calculated at each lattice intersection on a regularly-spaced grid of 2.0 Å. The grid pattern was generated automatically by the SYBYL/CoMFA routine and extended 4.0 Å units beyond the dimensions of each molecule in the X, Y, and Z directions. An sp^3 carbon atom with a van der Waals radius of 1.52 Å and a + 1.0 charge was used as a probe to calculate the steric (Lennard–Jones 6–12 potential) and electrostatic (Coulombic potential) fields with a distance-dependent dielectric at each lattice point. Values of the steric and electrostatic fields were truncated to 30.0 kcal/mol. The CoMFA steric and electrostatic fields generated were scaled by the CoMFA standard option available in SYBYL. The electrostatic fields were ignored at lattice points with maximal steric interactions.

CoMSIA descriptors

CoMSIA calculates similarity indices at the intersections of a surrounding lattice. The similarity index $A_{F,k}$ for a molecule j with atoms at the grid point q is determined by the equation

$$A_{F,k}^q(j) = \sum_i \omega_{\text{probe},k} \omega_{ik} e^{-\alpha r_{iq}^2} \tag{1}$$

where ω_{ik} is the actual value of the physicochemical property k of atom i; $\omega_{\text{probe},k}$ is the probe atom with a charge of +1, radius of 1 Å, hydrophobicity of +1, hydrogen bond donating capacity of +1, and hydrogen bond accepting capacity of +1; and r_{iq} is the mutual distance between the probe atom at grid point q and the atom i of the molecule. Five physicochemical properties k (steric, electrostatic, hydrophobic, hydrogen bond acceptor and donor) were evaluated, using a common charged,

Table 1 Molecular structures, experimental activities (Obs) and predicted activities (Calcd) with residues (Res) by CoMFA, CoMSIA and Hansch–Fujita of tyrosinase inhibitors (CQ1-CQ19 (I), CS1-CS26 (II) and CS27-CS29 (III))
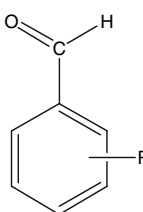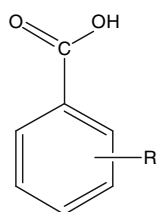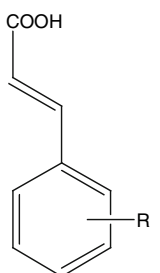

CQ1 to **CQ19**   **CS1** to **CS26**   **CS27** to **CS29**

Compd	Group	R	IC$_{50}$ (mmol/L)	pIC$_{50}$						
				Obs	Calcd[a]	Res[b]	Calcd[c]	Res[d]	Calcd[e]	Res[f]
Training set										
CQ1	I	–	0.80	3.10	3.28	−0.18	3.37	−0.27	3.26	−0.16
CQ2		2-OH	2.85	2.55	2.68	−0.13	2.63	−0.08	2.87	−0.32
CQ4		4-OH	1.15	2.94	3.00	−0.06	2.97	−0.03	2.87	0.07
CQ5		3-OMe	1.80	2.74	2.70	0.04	2.60	0.14	2.91	−0.17
CQ6		4-OMe	0.35	3.46	3.44	0.02	3.44	0.02	3.21	0.25
CQ8		3-Me	0.45	3.35	3.49	−0.14	3.45	−0.10	3.47	−0.12
CQ9		4-Me	0.12	3.92	3.78	0.14	3.89	0.03	3.55	0.37
CQ10		4-C$_2$H$_5$	0.095	4.02	4.07	−0.05	4.10	−0.08	3.87	0.15
CQ11		4-C$_3$H$_7$	0.075	4.12	3.89	0.23	4.23	−0.11	4.07	0.05
CQ13		4-C(CH$_3$)$_3$	0.038	4.42	4.41	0.01	4.17	0.25	4.52	−0.10
CQ14		4-C$_4$H$_9$	0.038	4.42	4.58	−0.16	4.59	−0.17	4.22	0.20
CQ15		4-C$_5$H$_{11}$	0.0135	4.87	4.87	0.00	4.72	0.15	4.42	0.45
CQ16		4-C$_6$H$_{13}$	0.008	5.10	5.14	−0.04	4.95	0.16	–	–
CQ18		4-C$_8$H$_{17}$	0.01	5.00	5.02	−0.02	5.17	−0.17	–	–
CQ19		4-CH = CH$_2$	0.023	4.64	4.30	0.34	4.09	0.55	4.30	0.34
CS1	II	–	1.00	3.00	3.00	0.00	3.01	−0.01	2.80	0.20
CS2		4-OH	1.30	2.89	2.73	0.16	2.62	0.27	2.41	0.48
CS4		4-OC$_2$H$_5$	1.10	2.96	2.83	0.13	2.63	0.33	2.84	0.12
CS5		4-OC$_3$H$_7$	1.85	2.73	2.78	−0.05	2.79	−0.05	2.80	−0.07
CS6		4-OC$_4$H$_9$	1.65	2.78	2.66	0.12	2.81	−0.03	3.11	−0.32
CS7		4-OC$_5$H$_{11}$	1.40	2.85	2.86	−0.01	3.02	−0.17	–	–
CS9		2-Me	2.56	2.59	2.58	0.01	2.87	−0.28	3.09	−0.49
CS10		3-Me	0.58	3.24	3.33	−0.09	3.10	0.14	3.01	0.23
CS11		4-Me	0.35	3.46	3.51	−0.05	3.54	−0.08	3.09	0.38
CS13		4-C$_3$H$_7$	0.235	3.63	3.63	0.00	3.88	−0.25	3.61	0.02
CS14		4-CH(CH$_3$)$_2$	0.225	3.65	3.55	0.10	3.78	−0.13	3.66	−0.01
CS15		4-C(CH$_3$)$_3$	0.215	3.67	3.78	−0.11	3.83	−0.16	4.06	−0.39
CS16		4-C$_4$H$_9$	0.165	3.78	3.85	−0.07	3.77	0.01	3.76	0.02
CS17		4-C$_5$H$_{11}$	0.12	3.92	3.90	0.02	3.84	0.08	3.96	−0.04
CS19		4-C$_7$H$_{15}$	0.095	4.02	4.06	−0.04	3.97	0.05	–	–
CS20		4-C$_8$H$_{17}$	0.082	4.09	4.11	−0.02	4.01	0.08	–	–
CS22		2-OH; 4-Ome	2.28	2.64	2.63	0.01	2.73	−0.09	2.61	0.03
CS23		2-OH; 5-Ome	7.90	2.10	2.11	−0.01	2.03	0.07	2.22	−0.12
CS24		2-OH; 4-Me	1.65	2.78	2.93	−0.15	2.82	−0.04	2.97	−0.18
CS26		4-CH = CH$_2$	0.33	3.48	3.50	−0.02	3.45	0.03	3.84	−0.36

Table 1 continued

Compd	Group	R	IC$_{50}$ (mmol/L)	pIC$_{50}$						
				Obs	Calcd[a]	Res[b]	Calcd[c]	Res[d]	Calcd[e]	Res[f]
CS27	III	–	2.10	2.68	2.77	−0.09	2.85	−0.17	–	–
CS28		4-OH	0.50	3.30	3.12	0.18	3.12	0.18	–	–
CS29		4-OMe	0.42	3.38	3.42	−0.04	3.43	−0.05	–	–
Test set										
CQ3	I	3-OH	3.70	2.43	2.89	−0.46	2.55	−0.12	2.50	−0.07
CQ7		2-Me	2.60	2.59	2.87	−0.28	3.21	−0.62	3.55	−0.96
CQ12		4-CH(CH$_3$)$_2$	0.075	4.12	3.82	0.30	4.13	−0.01	4.12	0.00
CQ17		4-C$_7$H$_{15}$	0.0088	5.06	5.10	−0.04	5.09	−0.03	–	–
CS3	II	4-Ome	0.42	3.38	2.71	0.67	2.54	0.84	2.75	0.63
CS8		4-OC$_6$H$_{13}$	1.15	2.94	2.96	−0.02	3.19	−0.25	–	–
CS12		4-C$_2$H$_5$	0.29	3.54	3.81	−0.26	3.74	−0.20	3.41	0.13
CS18		4-C$_6$H$_{13}$	0.11	3.96	3.97	−0.01	3.92	0.04	–	–
CS21		2-OH	4.30	2.37	2.93	−0.56	3.22	−0.85	2.41	−0.04
CS25		2-OH; 5-Me	2.15	2.67	2.38	0.29	2.51	0.16	2.89	−0.22

[a] Calculated using the CoMFA model based on the common substructure-based alignment

[b] Calculated residues using CoMFA model

[c] Calculated using the CoMSIA model with the steric and electrostatic fields

[d] Calculated residues using CoMSIA model

[e] Calculated using the Hansch–Fujita model

[f] Calculated residues using Hansch–Fujita model

hydrophobic, hydrogen-bond-accepting probe atom with a radius of 1 Å. A Gaussian-type distance dependence was considered to exist between the grid point q and each atom i of the molecule. The value of the so-called attenuation factor α was set to 0.3. A lattice with 2 Å grid spacing was generated automatically.

A partial least-squares (PLS) approach [23–25], an extension of multiple linear regression analysis, was used to derive the 3D-QSAR in which the CoMFA and CoMSIA descriptors were used as independent variables and pIC$_{50}$ values were dependent variables. Cross-validation with the Leave-One-Out (LOO) option and a column filtering value of 2.0 kcal/mol was carried out to obtain the optimal number of components to be used in the final analysis. After the optimal number of components was determined, a non-cross-validated analysis was performed without column filtering. The q^2 (cross-validated r^2), S_{press} (the root mean predictive error sum of squares), r^2 (non-cross-validated r^2), F- (F-test), and the standard error of estimate (SEE) values were computed according to the definitions in SYBYL. These values are listed in Table 3, where $Pr^2 = 0$ indicates the probability of obtaining the observed F ratio value by chance alone, if the target and the explanatory variables themselves are truly uncorrelated. If $Pr^2 = 0$ is zero, then the results are significant and did not occurred by chance.

In SYBYL/QSAR, the intensity of the cross-validation process is controlled by selecting the numbers of groups or times the cross-validation step is to be carried out while predicting all rows (at each stage of model development). For a more rigorous statistical test, several runs of cross-validation using five groups were performed in which each target property value was predicted by a model based on $\sim 80\%$ of the available data. To further obtain statistical confidence limits for the analysis, a bootstrapping analysis (100 runs) was performed.

A common test to check the consistency of the models is to randomize the biological data and repeat the model derivation process, thereby allowing detection of possible chance correlations. Following randomization, very low or negative q^2 value were observed in all of the PLS analyses.

Physicochemical parameters for 2D-QSAR

The physicochemical parameters used in the present study were obtained from literature [26]. The hydrophobic parameter (π) describes the contribution of a substituent to the lipophilicity of a compound. The molecular refractivity (MR) reflects the molar volume corrected by the refractive index and measures the size and polarizability of the substituent. H-acc and H-don represent the hydrogen bond acceptor and donor parameters of molecules, respectively. The Hammett electronic parameter (σ) was used to represent the electron donating

Table 2 Physicochemical parameters of tyrosinase inhibitors

NO	Substitutes	Hydrophobic parameters π	H-bond parameters		Steric parameters MR	Electronic parameters σ
			H-Acceptor	H-Donor		
CQ1	H	0.00	0	0	1.03	0
CQ2	2-OH	−0.67	1	1	2.85	−0.37
CQ3	3-OH	−0.67	1	1	2.85	0.12
CQ4	4-OH	−0.67	1	1	2.85	−0.37
CQ5	3-OMe	−0.20	1	0	7.87	0.12
CQ6	4-OMe	−0.20	1	0	7.87	−0.27
CQ7	2-Me	0.56	0	0	5.65	−0.17
CQ8	3-Me	0.56	0	0	5.65	−0.07
CQ9	4-Me	0.56	0	0	5.65	−0.17
CQ10	4-C_2H_5	1.02	0	0	10.30	−0.15
CQ11	4-C_3H_7	1.55	0	0	14.96	−0.13
CQ12	4-$CH(CH_3)_2$	1.53	0	0	14.96	−0.15
CQ13	4-$C(CH_3)_3$	1.98	0	0	19.62	−0.20
CQ14	4-C_4H_9	2.13	0	0	19.61	−0.16
CQ15	4-C_5H_{11}	2.67	0	0	24.26	−0.16
CQ19	4-$CH = CH_2$	0.82	0	0	10.99	−0.02
CS1	H	0.00	0	0	1.03	0
CS2	4-OH	−0.67	1	1	2.85	−0.37
CS3	4-OMe	−0.20	1	0	7.87	−0.27
CS4	4-OC_2H_5	0.38	1	0	12.47	−0.24
CS5	4-OC_3H_7	1.05	1	0	17.06	−0.25
CS6	4-OC_4H_9	1.55	1	0	21.66	−0.32
CS9	2-Me	0.56	0	0	5.65	−0.17
CS10	3-Me	0.56	0	0	5.65	−0.07
CS11	4-Me	0.56	0	0	5.65	−0.17
CS12	4-C_2H_5	1.02	0	0	10.30	−0.15
CS13	4-C_3H_7	1.55	0	0	14.96	−0.13
CS14	4-$CH(CH_3)_2$	1.53	0	0	14.96	−0.15
CS15	4-$C(CH_3)_3$	1.98	0	0	19.62	−0.20
CS16	4-C_4H_9	2.13	0	0	19.61	−0.16
CS17	4-C_5H_{11}	2.67	0	0	24.26	−0.16
CS21	2-OH	−0.67	1	1	2.85	−0.37
CS22	2-OH; 4-OMe	−0.87	2	1	10.72	−0.64
CS23	2-OH; 5-Ome	−0.87	2	1	10.72	−0.13
CS24	2-OH; 4-Me	−0.12	1	1	8.50	−0.54
CS25	2-OH; 5-Me	−0.12	1	1	8.50	−0.44
CS26	4-$CH = CH_2$	0.82	0	0	10.99	−0.02

or accepting properties of a substituent; the conjugation effect of all the substituents (Fig. 2) was considered in the formula Eq. 2.

$$\sigma = \sigma_p R_1 + \sigma_m R_2 + \sigma_p R_3 \qquad (2)$$

where R_1, R_2 and R_3 are the substituents of *ortho*, *meta*, and *para* positions, respectively; and σ_m and σ_p are the electronic parameters of the substituents at the *meta* and *para* positions, respectively. All substituent parameters used in the QSAR analyses are listed in Table 2.

Molecular docking using FlexX

FlexX is a fast and flexible docking method that uses an incremental construction algorithm to place ligands into an active site [27, 28]. The coordinates of the tyrosinase crystal structure used in this research were derived from PDB entry 1WX2, as retrieved from the Protein Data Bank (PDB http://www.rcsb.org/pdb), which is the suitable tyrosinase crystal structure currently available. The structure consists of the enzyme complexed with the caddie protein ORF378.
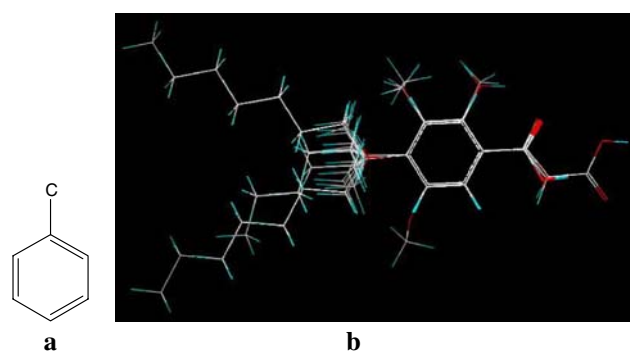
Fig. 1 Superposition of compounds in the training and test sets using the common substructure-based alignment rules. (**a**) Common substructure-based alignment rules; (**b**) Superposition of compounds in the training and test sets

Table 3 Summary of results from the CoMFA and CoMSIA

	CoMFA	CoMSIA[a]	CoMSIA[b]
Component	6	6	6
q^2	0.855	0.841	0.832
S_{press}	0.144	0.140	0.150
r^2	0.978	0.946	0.950
SEE	0.121	0.190	0.182
F	230.735	90.896	98.906
$Pr^2 = 0$	0.000	0.000	0.000
Fraction			
Steric	0.587	0.364	0.342
Electrostatic	0.413	0.636	0.500
H-Acceptor			0.159
r^2(bs)	0.985	0.968	0.967
SD[c]	0.063	0.082	0.084

[a] CoMSIA using the steric and electrostatic fields

[b] CoMSIA using the steric, electrostatic and hydrogen bond acceptor fields

[c] Results from 100 runs of bootstrapped analyses
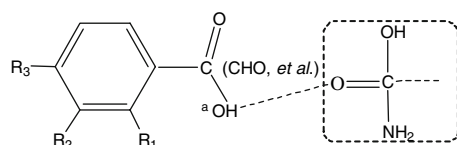
Fig. 2 Interaction model of benzoic acid derivatives with Tyr[98] of the tyrosinase active site

Any amino acid residue within 6.5 Å of the Cu^A and Cu^B was included in the substrate-binding pocket. The crystal structure was stripped of water molecules, hydrogen atoms were generated in standard geometry, and metal atoms were retained in the protein description. Further target-specific details are given below. Twelve compounds of benzoic acid derivatives were selected and docked into the tyrosinase active site using the FlexX approach.

Results

3D-QSAR models

CoMFA and CoMSIA are widely used for studying quantitative structure–activity relationships at the 3D level [29–33]. CoMFA and CoMSIA models were derived from a set of 38 structurally similar compounds, and 10 compounds were selected as a test set for model validation. The PLS analysis results are summarized in Table 3, and indicate that all of the statistical indices are reasonably high and that the constructed CoMFA and CoMSIA models have good predictability.

In the CoMFA model, PLS analysis yielded a q^2 value of 0.855 with an SEE of 0.121; the noncross-validated PLS analysis gave a conventional r^2 of 0.978 with an S_{press} of 0.144; and the model yielded 6 optimum components based on the CoMFA standard options.

CoMSIA analyses were performed according to the same procedure as CoMFA, and used the following descriptor fields in different combinations: steric, electrostatic, and hydrophobic parameters, and hydrogen bond acceptor and donor. CoMSIA results are summarized in Table 3. A q^2 of 0.841 and a conventional r^2 of 0.946 were obtained when the steric and electrostatic fields were considered. The F and SEE values were 90.896 and 0.190, respectively. These data also indicate that a reliable CoMSIA model was successfully constructed.

Because the LOO cross-validation method can lead to high q^2 values that do not necessarily reflect the general predictability of the model, we performed cross-validation using five groups. As randomness of the cross-validation group formation may significantly impact the results, our cross-validation was performed 30 times for the CoMFA and CoMSIA analyses, based on the alignment. The results of cross-validation using five groups are reported in Table 4, and indicate that although the mean q^2 values were slightly lower compared to the values obtained by the LOO method, both of the q^2 values were above 0.809. The results suggested that there was a good internal consistency in the underlying data set.

The predicted activities of the training (38 compounds) and test (10 compounds) sets derived from the constructed 3D-QSAR model and actual activities are listed in Table 1. Correlation of the predicted and actual activities given by the CoMFA and CoMSIA models, considering the steric and electrostatic fields, are presented in Figs. 3 and 4, respectively. The greatest advantage of CoMFA and CoMSIA is that the effect of a parameter on a target property can be viewed as a 3D coefficient contour plot. These coefficient contour plots were helpful in identifying important regions where changes in the steric or electrostatic fields could affect biological activity, and could also help identify possible interaction sites. Figures 5–8 give

Table 4 Results of analysis with Cross-validation using five groups

	q^{2a}		
	CoMFA	CoMSIA[b]	CoMSIA[c]
Mean	0.836	0.819	0.809
High	0.876	0.878	0.861
Low	0.783	0.696	0.684

[a] Cross-validated q^2 using five groups with optimum number of components average of 30 runs

[b] CoMSIA using the steric and electrostatic fields

[c] CoMSIA using the steric, electrostatic and hydrogen bond acceptor fields

isocontour diagrams obtained from the CoMFA and CoM-SIA analyses using the steric and electrostatic fields. In this figures, the field contributions ("stdev*coeff") of different properties are illustrated together with exemplary ligands.

CoMFA and CoMSIA contour maps

The CoMFA contour map of steric contribution is depicted in Fig. 5; the green and yellow polyhedra describe regions of space around the molecules where an increase in steric bulk results in enhanced or diminished activity, respectively. The electrostatic contribution of CoMFA is illustrated in Fig. 6; the blue contour defines a region where increasing positive charge will result in increased activity, whereas the red contour defines a region of space where increasing electron density is favorable. Green contours embedded in yellow contours indicate that there

Fig. 4 CoMSIA predicted as experimental pIC_{50} values. Open circles represent predictions for the training set; solid circles represent predictions for the test set
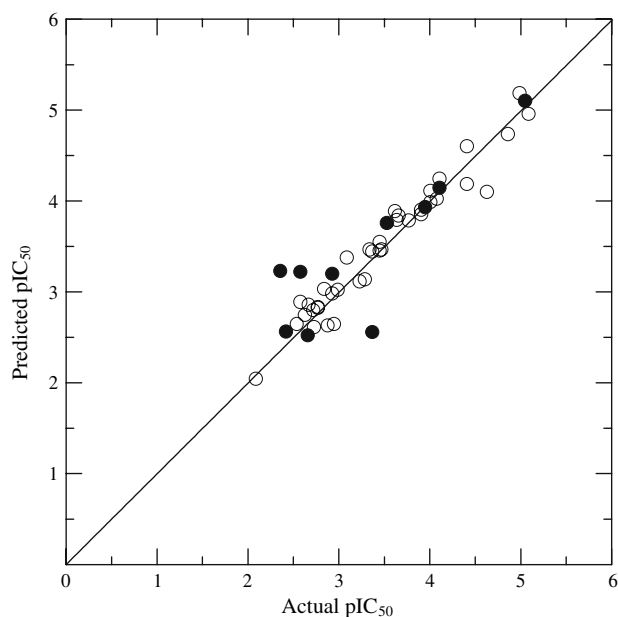

Fig. 3 CoMFA predicted as experimental pIC_{50} values. Open circles represent predictions for the training set; solid circles represent predictions for the test set
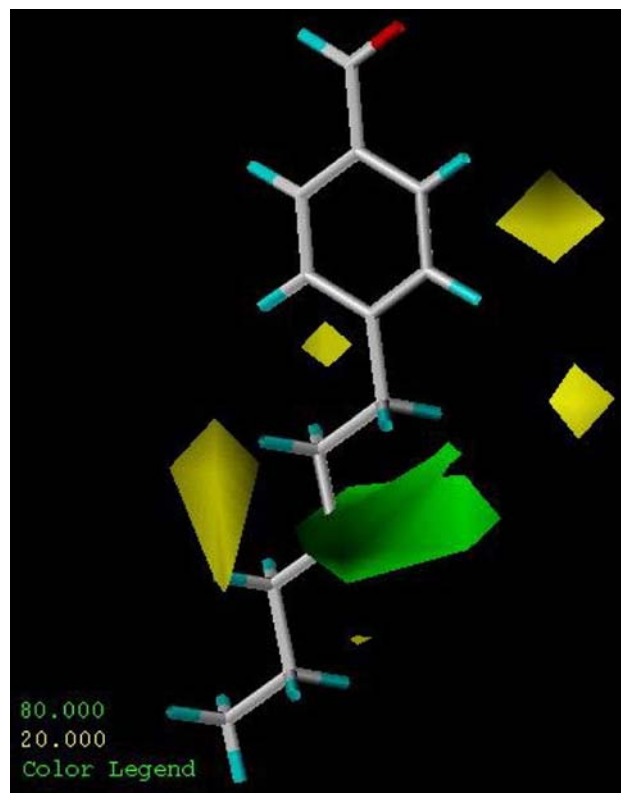

Fig. 5 Steric maps from the CoMFA model. Compound **CQ16** (4-hexylbenzaldehyde) is shown inside the field. Sterically favored areas (contribution level of 80%) are represented by green polyhedra. Sterically disfavored areas (contribution level of 20%) are represented by yellow polyhedra
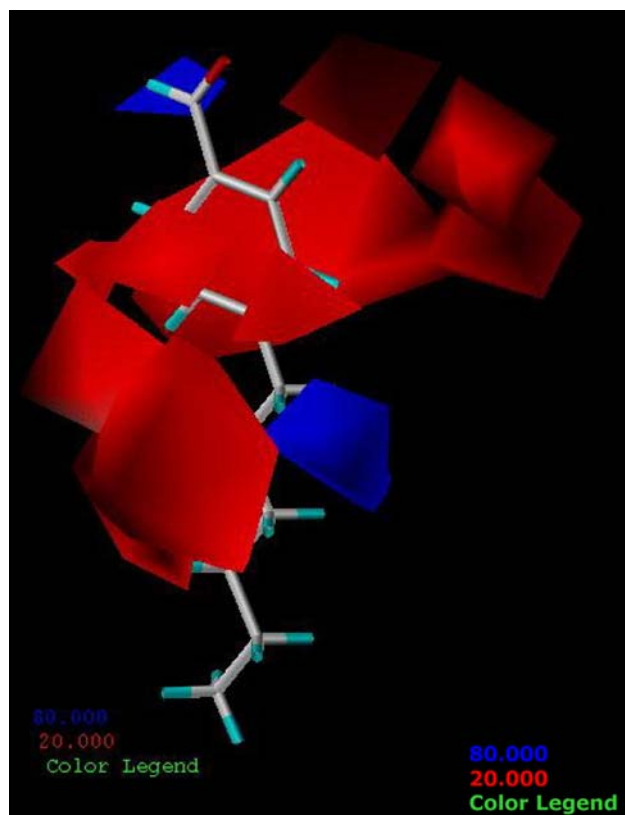
Fig. 6 Electrostatic maps from the CoMFA model. Compound **CQ16** is shown inside the field. Blue contours (80% contribution) encompass regions where an increase of positive charge will enhance affinity, whereas in red contoured areas (20% contribution) more negative charges are favorable for binding properties

Fig. 7 Steric maps from the CoMSIA model using the steric and electrostatic fields. Compound **CQ16** is shown inside the field. Green contours (actual values 0.017) enclose areas where steric bulk will enhance affinity, and yellow contours (actual values −0.010) highlight areas that should be kept unoccupied

exists an optimal value for the steric effect, and blue contours embedded in red contours indicate that there exists an optimal electrostatic effect.

In Fig. 5 and 7, the compound CQ16 (4-hexylbenzaldehyde) is displayed in the CoMFA and CoMSIA map to aid visualization. The green contour around the *para* position of the aldehyde group indicates that steric bulk is favored, especially in CoMSIA map. We can infer that the presence of bulky substitutes in the *para* position is very important for inhibitory activity; for example, presence of CQ10-18 and CS16-20 in this position resulted in a relatively large activity increase compared to other compounds. Conversely, when the *ortho* position contained
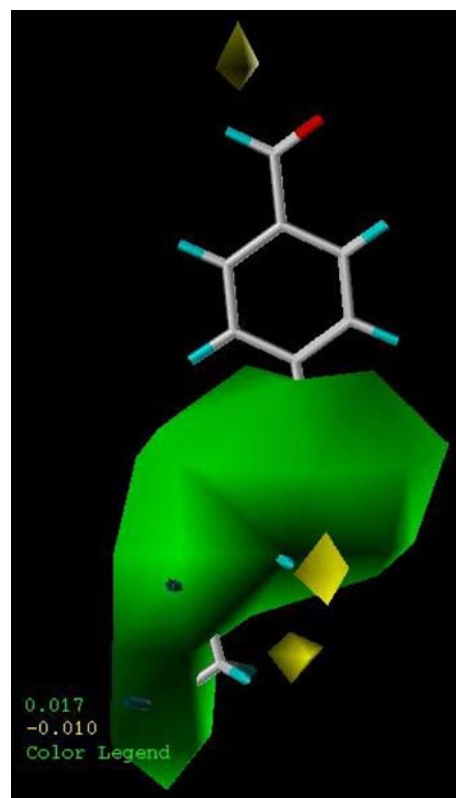
CQ7, CS9, and CS21). In Fig. 6, the red contours around the aromatic ring indicate that negative charges are favored in those regions. In Fig. 8, the red contours embedded in blue contours around the aromatic ring indicate that an optimal electrostatic effect is favored.

2D-QSAR analysis and model validation

Variations in the inhibitory activity of 37 compounds against mushroom tyrosinase were analyzed using the physicochemical parameters listed in Table 2. The pIC_{50} equation with best correlation coefficients obtained by a Hansch–Fujita approach is as follows:

$$pIC_{50} = 2.546 - 1.776\,\pi - 2.317\,\text{H-acc} + 0.249\,\text{MR} - 0.785\,\sigma + 0.460\,\text{I} \tag{3}$$
$$\phantom{pIC_{50} = 2.546}(0.056)\quad(0.211)\qquad(0.707)\qquad(0.120)\quad(0.104)$$

a bulky substitutent, such as a hydroxyl or methyl group, the bioactivities decreased consequently (compounds CQ2,

where n (number of compounds) = 37, $r = 0.900$, $q^2 = 0.733$, s (standard error) = 0.328, and $F = 26.047$.
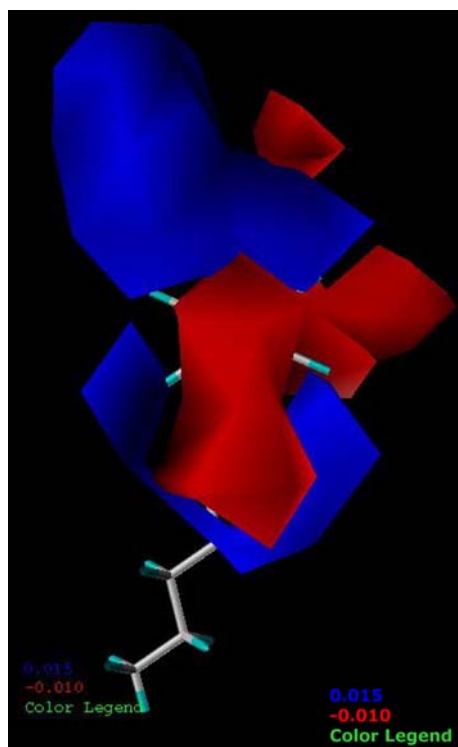
Fig. 8 Electrostatic field map of CoMSIA using the steric and electrostatic fields. Compound **CQ16** is shown inside the field. Blue contours (actual values 0.015) encompass regions where an increase of positive charge will enhance affinity, whereas in red contoured areas (actual values −0.010) more negative charges are favorable for binding properties

Table 6 Pearson correlation matrix for variables used to drive Eq. 1

	π	H-acceptor	MR	I	σ
π	1.000				
H-acceptor	0.284	1.000			
MR	0.115	0.047	1.000		
I	0.001	0.046	0.114	1.000	
σ	0.065	0.279	0.0001	0.100	1.000
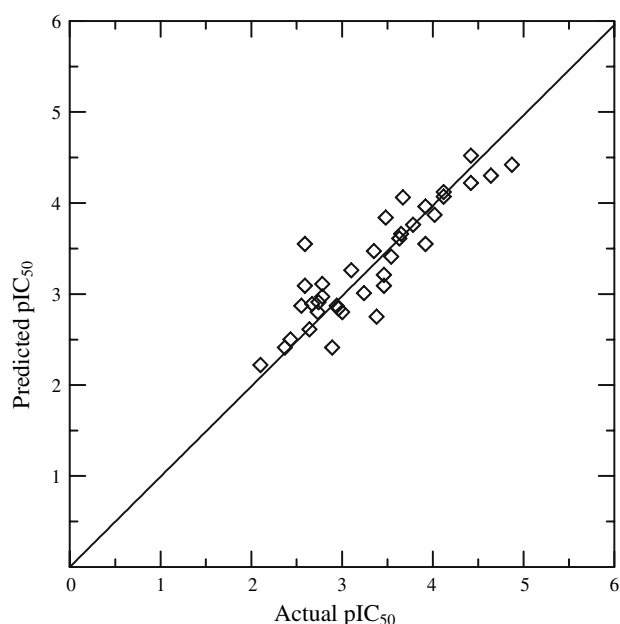
Fig. 9 Hansch–Fujita model prediction pIC_{50} values of the tyrosinase inhibitors

The numbers in parentheses under each coefficient are the 95% confidence intervals of the regression coefficient.

The stepwise development of this equation and the Pearson correlation matrix of variables are shown in Tables 5 and 6, respectively. The in vitro biological activities calculated by Eq. 3 are listed in Table 1, and the correlation of the Hansch–Fujita model between the predicted and actual activities is presented in Fig. 9. I is an indicator variable assigned to be 1 for benzaldehyde derivatives and 0 for benzoic acid derivatives. The coefficient of I term indicates that benzaldehyde derivatives have a much stronger activity than benzoic acid derivatives, given the same electronic, steric, and hydrogen bond accepter substituents.

MR had a positive coefficient in Eq. 3, significantly contributing to activity; but the negative coefficients of π, H-acc, and σ indicated that low values of any of these variables would increase activity. The parameters of H-acc and π had relatively high negative coefficients, which clearly indicated that the activity was sensitive to the hydrogen acceptor and hydrophobic properties, and that slight changes in these values could have a significant effect on activity. Equation 3 also indicated that π, MR, and H-acc accounts for 34.9, 31.6, and 26.7% of the calculated biological variance, respectively. In addition, from

Table 5 Development of QSAR of Eq.1

Intercept	π	H-acceptor	MR	I	σ	r	q^2	s	F
2.993	0.513					0.762	0.537	0.455	48.487
3.238	0.374	−0.340				0.791	0.574	0.436	28.467
2.928	−1.659	−2.192	0.233			0.839	0.600	0.394	26.220
2.661	−1.625	−2.076	0.234	0.414		0.888	0.707	0.339	29.688
2.546	−1.776	−2.317	0.249	0.460	−0.785	0.900	0.733	0.328	26.047

Table 5 we find that the introduction of the I and σ terms led to a significant improvement in the correlation, indicating that the indicator variable and electronic parameters were also important.

The interaction between benzoic acid derivatives and the 1WX2 active site

By default, FlexX produced 30 docked structures for each benzoic acid derivative. The conformation with the lowest docking energy in the most populated cluster was selected as the most likely "active" conformation against the 1WX2 active site. In the present study, 12 compounds of benzoic acid derivatives were successfully docked into the 1WX2 site, the active conformation of which is presented in Fig. 10.

The distance between hydrogen atom of the hydroxyl (ᵃOH) and the carbonyl oxygen atom of Tyr[98] was 1.827 Å (Fig. 2), indicating the likely presence of a hydrogen bond. Formation of this hydrogen bond stabilized the position of Tyr[98], preventing Tyr[98] from participating in the interaction between tyrosinase and ORF378. ORF378 is consequently unable to serve as a transporter of Cu(II) ions to the catalytic site of tyrosinase, a necessary step in tyrosinase catalytic activity. Figure 10 also illustrates an additional condition in which a substituent that was *para* to the aldehyde group and interacted with surrounding amino

acid residues (including Trp[184], Arg[185], and Pro[102]) indirectly led to a loss of function in the active center of tyrosinase.

Discussion and conclusions

As tyrosinase inhibitors become more important in medicinal, agricultural and cosmetic products, the need to design and screen tyrosinase inhibitors with higher bioactivities becomes pressing. Our previous study [14] indicated the existence of a large green polyhedron occupying the *para* position of the aldehyde (carboxyl, thiosemicarbazone) group, suggesting a favorable location for bulky substitutes; similar results were obtained in this paper. Furthermore, studies with *Sacrophaga neobelliaria* phenoloxidase inhibitors [34] suggest that substituent hydrophobicity at this position also plays a key role in inhibition activity. Likewise, very similar results were obtained from the 2D-QSAR study in the present paper, which determined that the hydrophobic parameter π accounts for 34.9% of the calculated biological variance. A hydroxyl group at the *ortho* position of the aldehyde group was also predicted to contribute to higher inhibition activity [34]; however, our results indicate that a hydroxyl group at this position decreases inhibitory acitivity.

p-butylbenzaldehyde thiosemicarbazone was previously reported to inhibit tyrosinase through interaction of the

Fig. 10 Molecular docking of benzoic acid derivatives with the active center of tyrosinase
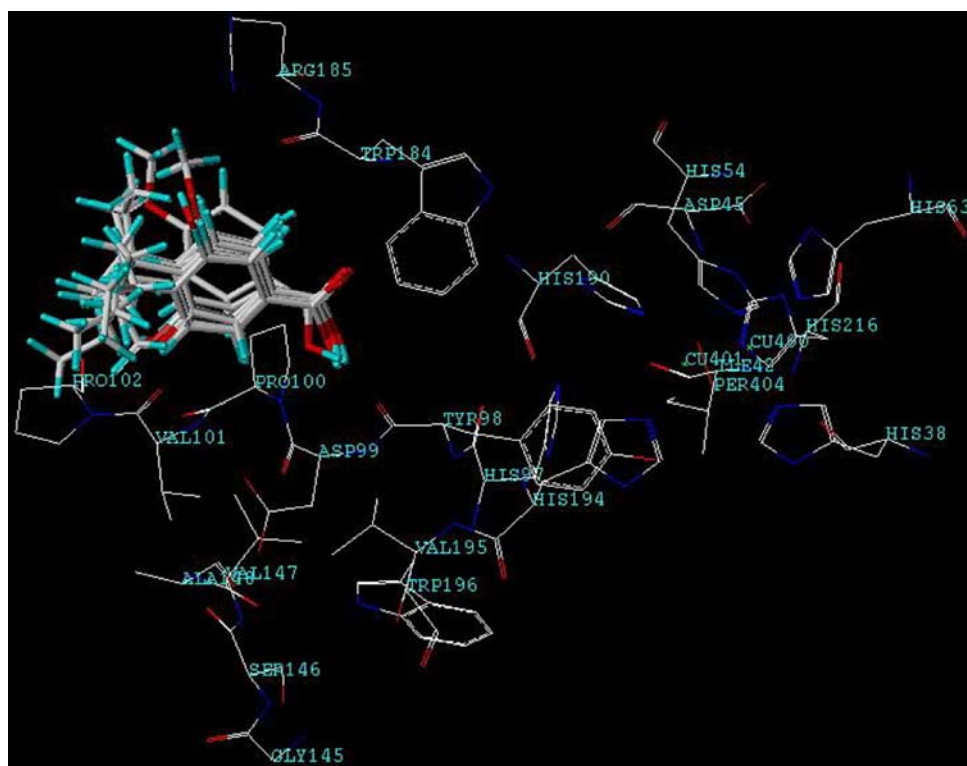
"clamp" structure (H^{21}–N^9–C^{10}–N^{12}–H^{22}) with Tyr^{98} of the enzyme active site [14]. Here, we report that one possible principal mechanism of interaction of benzoic acid derivatives with the tyrosinase active site is the formation of a hydrogen-bond between the hydroxyl (^{a}OH) and the carbonyl oxygen atoms of Tyr^{98}, which stabilize the position of Tyr^{98} and prevent Tyr^{98} from participating in the interaction between tyrosinase and ORF378.

The QSAR models reported here had good statistical results in terms of their q^2 and r^2 values. The information derived from this study provides necessary information for predicting the affinity of related compounds and for guiding further structural modification and synthesis of potent new tyrosinase inhibitors. These results may provide a basis for the use of tyrosinase inhibitors as novel, environmentally-friendly agricultural chemicals or food fresh-keeping agents.

Literature Cited

1. Sánchez-Ferrer A, Rodríguez-López JN, García-Cánovas F, García-Carmona F (1995) Biochim Biophys Acta 1247:1
2. Chase MR, Raina K, Bruno J, Sugumaran M (2000) Insect Biochem Molec 30:953
3. Chen QX, Liu XD, Huang H (2003) Biochemistry-Moscow 68:644
4. Xu Y, Stokes AH, Freeman WM, Kumer SC, Vogt BA, Vrana KET (1997) Mol Brain Res 45:159
5. Oetting WS (2000) Pigm Cell Res 13:320
6. Ashida M, Brey P (1995) Proc Natl Acad Sci USA 92:10698
7. Fitzpatrick TB, Eisen AZ, Wolff K, Freedberg IM, Austen KF (1983) Dermatology in general medicine. New York, NY
8. Maeda K, Fukuda M (1991) J Soc Cosmet Chem 42:361
9. Wang SD, Luo WC, Gao XX (2004) Agr Sci China 3:923
10. Wang SD, Luo WC, Xu SJ, Ding Q (2005) Pestic Biochem Phys 82:52
11. Wang SD, Luo WC, Xiao T, Xue CB (2005) Biopestic Int 1:82
12. Wang XY, Liu CY, Zhang JD, Luo WC (2005) Insect Sci 12:231
13. Xue CB, Luo WC, Chen QX, Wang Q, Ke LN (2006) Insect Sci 13:251
14. Xue CB, Zhang L, Luo WC, Xie XY, Jiang L, Xiao T (2007) Bioorg Med Chem 15:2006
15. Cramer RD III, Patterson DE, Bunce JD (1988) J Amer Chem Soc 110:5959
16. Klebe G, Abraham U, Mietzner T (1994) J Med Chem 37:4130
17. Chen QX, Song KK, Wang Q, Huang H (2003) J Enzym Inhib Med Ch 18:491
18. Huang H, Liu XD, Chen QX (2003) J Xiamen Univer 42:97
19. Huang XH, Chen QX, Wang Q, Song KK, Wang J, Li S, Guan X (2006) Food Chem 94:1
20. Zhang JP, Chen QX, Song KK, Xie JJ (2006) Food Chem 95:579
21. Chen QX, Song KK, Qiu L, Liu XD, Huang H, Guo HY (2005) Food Chem 91:269
22. Shi Y, Chen QX, Wang Q, Song KK, Qiu L (2005) Food Chem 92:707
23. Wold S, Albano C, Dunn WJ III, Edlund U, Esbensen K, Geladi P, Hellberg S, Johanson E, Lindberg W, Sjostrom M (1984) NATO ASI Ser Ser C 138:17
24. Wold S, Rhue A, Wold H, Dunn WJI (1984) SIAM J Sci Stat Comput 5:735
25. Clark M, Cramer RD III (1993) Quant Struct-Act Relat 12:137
26. Li RL (2004) In structure activity relationships of drug. The Medicine Science and Technology Press, Beijing, BJ
27. Rarey M, Kramer B, Lengauer T, Kleb G (1996) J Mol Biol 261:470
28. Rarey M, Kramer B, Lengauer T (1997) J Comput Aid Mol Des 11:369
29. Yang GF, Jiang XH, Yang HZ (2002) Pest Manag Sci 58:1063
30. Yang GF, Liu HY, Yang HZ (1999) Pestic Sci 55:1143
31. Yang GF, Jiang XH, Ding Y, Yang HZ (2002) Acta Chim Sinica 60:134
32. Yang GF, Huang XQ (2006) Curr Pharm Design 12:4601
33. Wei DG, Yang GF, Wan J, Zhan CG (2005) J Agric Food Chem 53:1604
34. Li W, Kubo I (2004) Bioorg Med Chem 12:701