Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry 15 (2007) 2006–2015
Bioorganic & Medicinal Chemistry

3D-QSAR and molecular docking studies of benzaldehyde thiosemicarbazone, benzaldehyde, benzoic acid, and their derivatives as phenoloxidase inhibitors

Chao-Bin Xue,[a,b] Li Zhang,[b] Wan-Chun Luo,[a,*] Xian-Ye Xie,[c] Lin Jiang[c] and Ting Xiao[a]

[a]*College of Plant Protection, Key Laboratory of Pesticide Toxicology and Application Technique, Shandong Agricultural University, Shandong Tai'an 271018, PR China*
[b]*Key Laboratory of Pesticide and Chemical Biology, Ministry of Education of China and College of Chemistry, Central China Normal University, Hubei Wuhan 430079, PR China*
[c]*College of Chemistry and Material Science, Shandong Agricultural University, Shandong Tai'an 271018, PR China*

Received 15 November 2006; revised 22 December 2006; accepted 23 December 2006
Available online 30 December 2006

Abstract—Phenoloxidase (PO), also known as tyrosinase, is a key enzyme in insect development, responsible for catalyzing the hydroxylation of tyrosine into *o*-diphenols and the oxidation of *o*-diphenols into *o*-quinones. Inhibition of PO may provide a basis for novel environmentally friendly insecticides. In the present study, we determined the inhibitory activities and IC_{50} values of 57 compounds belonging to the benzaldehyde thiosemicarbazone, benzaldehyde, and benzoic acid families against phenoloxidase from *Pieris rapae* (Lepidoptera) larvae. In addition, the inhibitory kinetics of 4-butylbenzaldehyde thiosemicarbazone against PO was measured in air-saturated solutions for the oxidation of L-3,4-dihydroxyphenylalanine (L-DOPA). The results indicated that the compound is a reversible noncompetitive inhibitor. The bioactivity results were used to construct three-dimensional quantitative structure–activity relationship (3D-QSAR) models using two molecular field analysis techniques: comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA). After carrying out superimposition using common substructure-based alignment, robust and predictive 3D-QSAR models were obtained from CoMFA ($q^2 = 0.926$, $r^2 = 0.986$) and CoMSIA ($q^2 = 0.933$, $r^2 = 0.984$) with six optimum components. The 3D-QSAR model built here will provide hints for the design of novel PO inhibitors. The molecular interactions between the ligands and the target were studied using a flexible docking method (FlexX). The best scored candidates were docked flexibly, and the interaction between the representative compound 4-butylbenzaldehyde thiosemicarbazone and the active site was elucidated in detail.

1. Introduction

Phenoloxidase (PO) (EC 1.14.18.1), also known as tyrosinase, a copper containing bifunctional enzyme, is highly conserved, present in microorganisms, animals, and plants.[1,2] It is the key enzyme involved in the browning that occurs upon bruising or long-term storage. Phenoloxidase catalyzes both hydroxylation of monophenols and oxidation of *o*-diphenols into *o*-quinones, and is involved in the formation of pigments such as melanins.[3] In mammals, the enzyme is responsible for skin pigmentation abnormalities, such as flecks, and defects.[4] Recently, the enzyme was reported to be linked to Parkinson's and other neurodegenerative diseases.[5,6] In insects, PO 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.[7] It is possible that inhibition of PO could lead to abrogation of insect defense mechanisms or abnormal body softening, both of which could be used in pest control. Therefore, PO is quite significant in the fields of industry, medicine, and agriculture.

Numerous reports on the purification, characterization, cloning, and kinetics of phenoloxidase have appeared

Keywords: 3D-QSAR; Molecular docking; Phenoloxidase; Benzaldehyde thiosemicarbazone derivatives; CoMFA; CoMSIA.
* Corresponding author. Tel.: +86 538 8242983; fax: +86 538 8242983; e-mail: wcluo@sdau.edu.cn

0968-0896/$ - see front matter © 2007 Published by Elsevier Ltd.
doi:10.1016/j.bmc.2006.12.038

over the past three decades. However, only recently was the first crystal structure published of a PO from any source, prokaryotic or eukaryotic. The enzyme isolated from *Streptomyces castaneoglobisporus* is a copper-bound enzyme complexed with ORF378, a 'caddie' protein assisting in the transport of two Cu(II) ions into the PO catalytic center. Two Cu(II) ions (Cu^A and Cu^B) are clearly defined in the active site of the PO crystal structure: Cu^A binds to the ε-nitrogen atoms of His^{38}, His^{54}, and His^{63}, and Cu^B binds to those of His^{190}, His^{194}, and His^{216}.[8]

An attractive theory states that interaction between PO and ORF378 is mediated by the accommodation of Tyr^{98} in the substrate-binding pocket of PO. The phenol ring of Tyr^{98} undergoes π-stacking with the imidazole ring of His^{194} of PO, one of the Cu^B ligands. In addition, its phenolic hydroxyl group forms a hydrogen bond with solvent species (water molecules or peroxide ions), forming a bridge structure with Cu^A and Cu^B in the active center. In a manner similar to the complexation of catechol oxidase with the potent inhibitor phenylthiourea (PTU), the Tyr^{98} ring is aligned perfectly with the aromatic ring of the inhibitor. These results suggest that Tyr^{98} of ORF378 functions as a competitive inhibitor to the substrate tyrosine.[8]

There is an urgent need to design and screen PO inhibitors with higher bioactivities and to elucidate the mechanisms of inhibition. In previous studies, we investigated the inhibition and inhibitory kinetics of compounds such as kojic acid, apigenin, quercetin, 4-hexylresorcinol, and 4-dodecylresorcinol against PO activity, with the intent of determining the inhibitory mechanism.[9–13] To our knowledge, there have been no reports concerning the three-dimensional quantitative structure–activity relationship (3D-QSAR) or the molecular docking of PO inhibitors with the active site, both of which are crucial to the design of effective inhibitory molecules. In the present study, the inhibitory activities of 57 compounds against phenoloxidase were investigated, and 3D-QSAR analysis was performed using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA).[14,15] In addition, the interactions between the ligands and the PO active site were also determined using the FlexX program.

2. Results and discussion

2.1. Effects of the 57 compounds on PO activity

The effects of the 57 compounds on the PO activity were examined by measuring the oxidation of L-DOPA. The inhibitory course of enzyme activity by inhibitor was found to be concentration-dependent as shown in Figure 1 (compound **N18**, 4-butylbenzaldehyde thiosemicarbazone). With increasing concentrations of the inhibitor, the residual enzyme activity decreased rapidly, although it was not completely suppressed at any concentration. The chemical concentration to inhibit 50% of the enzyme activity (IC_{50}) were estimated for all the compounds and were summarized in Table 1.

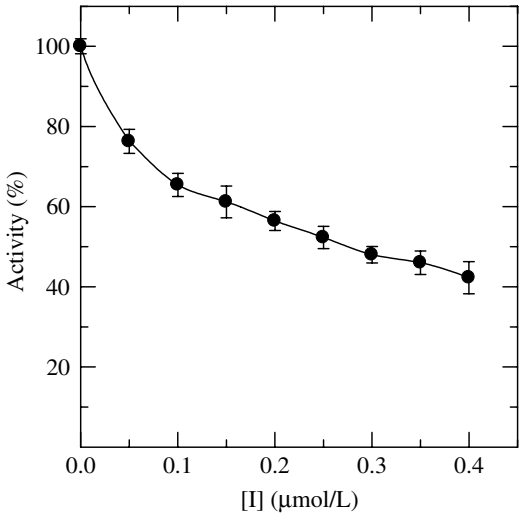


Figure 1. Inhibition of compound **N18** (4-butylbenzaldehyde thiosemicarbazone) on the PO from the 5th larvae of *P. rapae* for the catalysis of L-DOPA.

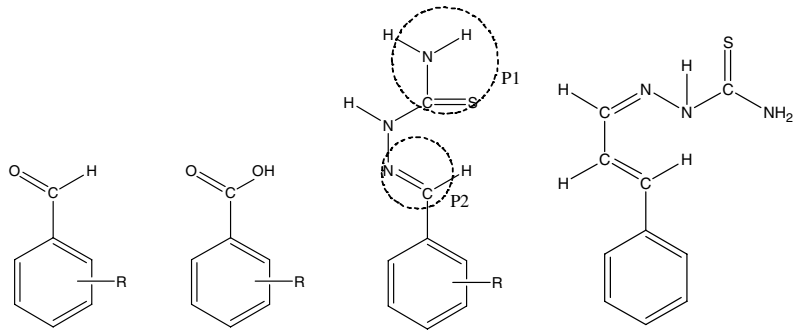
2.2. Reversible inhibition kinetics of compound N18 on PO activity

We investigated the inhibition mechanisms of **N18** on the enzyme for the L-DOPA oxidation. The plots of the remaining enzyme activity combined with different inhibitor concentrations revealed a series of straight lines, which all passed through the origin (Fig. 2). Increasing the inhibitor concentration resulted in the descending of the slope of the line, indicating that the inhibition of the enzyme was reversible. The inhibitor did not reduce the amount of the enzyme, but rather, resulted in the inhibition of activity of the enzyme for involved in oxidation of L-DOPA, which indicated that the inhibitory reaction was reversible.

2.3. Inhibitory type and inhibition constant of compound N18

Under the conditions employed in the present study, L-DOPA oxidation by the enzyme followed Michaelis–Menten kinetics. Compound **N18** behaved as a noncompetitive inhibitor (Fig. 3), which means that the inhibitor could combine with both the free enzyme (E) and the enzyme–substrate (ES) complex. Thus the interaction between **N18** and the enzyme is independent from the interaction between the substrate and the enzyme, and the inhibitor does not change substrate–enzyme affinity. As with increasing concentration, a series of lines was revealed with a common intercept on the 1/[S] axis but with different slopes. The apparent value of V_m decreased with no effect on K_m. The equilibrium constant for inhibitor binding, K_I, and the enzyme–substrate complex, K_{IS}, were the same, which were obtained from a plot of the apparent Michaelis–Menten constant ($K_{m\ app}$) versus the concentration of **N18**, which was linear, as shown in the inset of Figure 3. The obtained constant ($K_I = K_{IS}$) was 0.40 μmol/L.

Table 1. Molecular structures, experimental activities (Obs), and predicted activities (Calcd) with residues (Res) by CoMFA and CoMSIA of phenoloxidase inhibitors (**Q1**–**Q17**(I), **S1**–**S18** (II), **N1**–**N21** (III), and **N22** (IV))



Q1 to **Q17** **S1** to **S18** **N1** to **N21** **N22**

Compound	Group	R	IC_{50}(mmol/L)	pIC_{50}				
				Obs	Calcd[a]	Res[b]	Calcd[c]	Res[d]
Training set								
Q1	I	—	5.90	2.23	2.29	−0.06	2.08	0.15
Q2		2-OH	4.14	2.38	1.98	0.40	2.07	0.31
Q3		3-OH	9.62	2.02	2.36	−0.35	2.40	−0.38
Q5		2,4-OH	4.88	2.31	2.10	0.21	2.20	0.12
Q6		3,4-OH	9.00	2.05	2.04	0.01	2.25	−0.21
Q8		2-OMe	107.20	0.97	1.16	−0.19	1.00	−0.03
Q9		3-OMe	42.88	1.37	1.41	−0.04	1.44	−0.08
Q11		2-OH; 3-OMe	5.75	2.24	1.89	0.35	1.73	0.51
Q13		3-OH; 4-OMe	24.90	1.60	1.82	−0.22	2.01	−0.40
Q14		4-OH; 3-OMe	50.50	1.30	1.54	−0.24	1.70	−0.40
Q15		4-Propyl-*i*	0.86	3.07	3.07	−0.01	3.02	0.05
Q16		4-Butyl	0.74	3.13	3.00	0.13	3.06	0.07
Q17		4-Butyl-*t*	0.72	3.14	2.72	0.42	2.75	0.39
S2	II	2-OH	33.80	1.47	1.55	−0.08	1.71	−0.24
S4		4-OH	16.25	1.79	2.09	−0.30	1.95	−0.16
S5		2,4-OH	30.50	1.52	1.66	−0.15	1.83	−0.31
S6		2,5-OH	32.50	1.49	1.41	0.08	1.54	−0.05
S8		3-OH; 4-NH_2	25.50	1.59	1.97	−0.38	1.57	0.03
S9		4-OH; 3-NH_2	3.80	2.42	2.07	0.35	1.71	0.71
S10		3-OMe	13.30	1.88	1.88	−0.01	1.71	0.16
S11		4-OMe	5.55	2.26	2.33	−0.08	2.16	0.09
S12		3-OH; 4-OMe	47.50	1.32	1.47	−0.14	1.61	−0.29
S13		4-OH; 3-OMe	69.00	1.16	1.33	−0.17	1.48	−0.31
S14		4-Propyl	1.75	2.76	2.46	0.30	2.40	0.36
S16		4-Butyl	1.05	2.98	2.82	0.16	3.05	−0.07
S17		4-Butyl-*t*	2.10	2.68	2.66	0.02	2.81	−0.13
S18		4-Hexyl	1.18	2.93	2.78	0.15	2.79	0.14
N1	III	—	0.84×10^{-3}	6.08	6.17	−0.09	6.00	0.07
N3		3-OH	0.53×10^{-3}	6.28	6.09	0.19	6.05	0.22
N5		2,4-OH	1.05×10^{-3}	5.98	5.91	0.07	5.78	0.20
N6		3,4-OH	0.29×10^{-3}	6.54	6.35	0.19	6.44	0.10
N7		2,5-OH	2.25×10^{-3}	5.65	5.43	0.22	5.52	0.13
N8		2-OMe	21.60×10^{-3}	4.67	4.51	0.16	4.34	0.33
N9		3-OMe	1.92×10^{-3}	5.72	6.29	−0.58	6.09	−0.37
N10		4-OMe	0.60×10^{-3}	6.22	6.26	−0.04	6.12	0.10
N12		3-OH; 4-OMe	0.26×10^{-3}	6.59	6.22	0.37	6.38	0.20
N13		4-OH; 3-OMe	3.30×10^{-3}	5.48	5.10	0.38	5.48	0.00
N14		2,4-OMe	15.60×10^{-3}	4.81	4.94	−0.13	4.84	−0.03
N16		2,4,5-OMe	100.00×10^{-3}	4.00	3.97	0.03	4.00	0.00
N17		4-Propyl-*i*	0.35×10^{-3}	6.46	6.48	−0.03	6.66	−0.20
N18		4-Butyl	0.28×10^{-3}	6.55	6.73	−0.18	6.74	−0.18
N19		4-Butyl-*t*	0.56×10^{-3}	6.25	6.48	−0.23	6.51	−0.26
N20		3,5-Butyl-*i*; 2-OH	72.00×10^{-3}	4.14	4.46	−0.31	4.44	−0.30
N21		4-$N(CH_3)_2$	0.36×10^{-3}	6.44	6.46	−0.01	6.43	0.01
N22	IV	—	0.65×10^{-3}	6.19	6.37	−0.19	6.25	−0.07

Table 1 (*continued*)

Compound	Group	R	IC_{50}(mmol/L)	pIC_{50}				
				Obs	Calcd[a]	Res[b]	Calcd[c]	Res[d]
Test set								
Q4	I	4-OH	3.12	2.51	2.40	0.11	2.21	0.30
Q7		2,5-OH	1.05	2.98	2.27	0.71	1.99	0.99
Q10		4-OMe	1.90	2.72	2.64	0.08	2.19	0.53
Q12		2-OH; 4-OMe	2.00	2.70	2.62	0.08	2.19	0.51
S1	II	—	14.20	1.85	1.98	−0.13	1.82	0.03
S3		3-OH	16.25	1.79	2.05	−0.26	2.13	−0.34
S7		3,4-OH	29.50	1.53	2.16	−0.63	2.13	−0.60
S15		4-Propyl-*i*	1.92	2.72	2.27	0.45	2.34	0.37
N2	III	2-OH	2.20×10^{-3}	5.66	5.58	0.08	5.69	−0.03
N4		4-OH	0.44×10^{-3}	6.36	6.27	0.09	6.12	0.23
N11		2-OH; 3-OMe	2.08×10^{-3}	5.68	5.93	−0.24	5.72	−0.03
N15		3,4-OMe	2.70×10^{-3}	5.57	5.79	−0.22	5.79	−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, hydrophobic, and hydrogen bond acceptor fields.
[d] Calculated residues using the CoMSIA model.

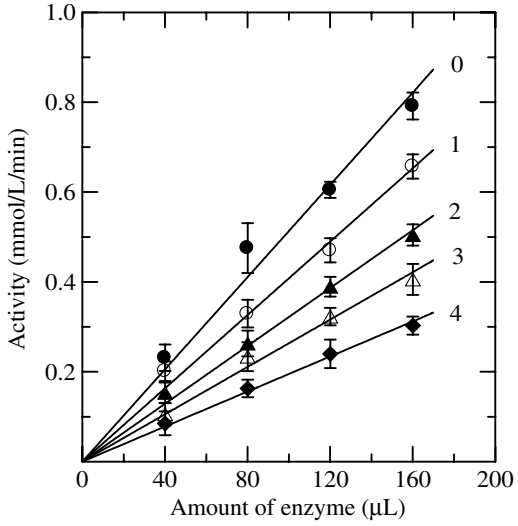


Figure 2. Relationship of the catalytic activity of the PO from the 5th larvae of *P. Rapae* with the enzyme concentrations at different concentrations of compound **N18** (4-butylbenzaldehyde thiosemicarbazone). Concentrations of **N18** for curves 0–4 were 0, 0.05, 0.15, 0.25, and 0.35 μmol/L, respectively.

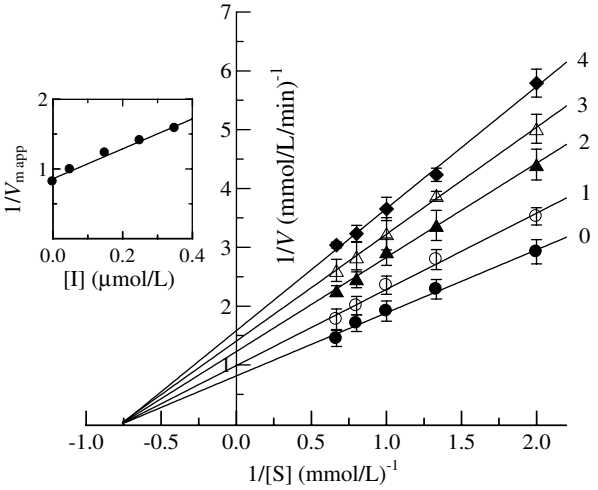


Figure 3. Lineweaver–Burk plot for inhibition of compound **N18** (4-butylbenzaldehyde thiosemicarbazone) on the PO from the 5th larvae of *P. rapae* for the catalysis of L-DOPA at 37 °C, pH 6.8. Concentrations of **N18** for curves 0–4 were 0, 0.05, 0.15, 0.25, and 0.35 μmol/L, respectively. The inset represents the plot of $K_{m\ app}$ versus the **N18** concentration for determining the inhibition constants (K_I).

2.4. 3D-QSAR models

CoMFA and CoMSIA are widely used for studying quantitative structure–activity relationships at the three-dimensional level.[21–25] In this study, CoMFA and CoMSIA models were derived from a set of 45 structurally similar compounds, and 12 compounds were selected as a test set for model validation. PLS analysis results are summarized in Table 2, and indicate that all of the statistical indices are reasonably high and the constructed CoMFA and CoMSIA models have good predictability. In the CoMFA model, PLS analysis yielded a cross-validated correlation coefficient q^2 of 0.926 with a standard error of estimate (SEE) of 0.250, the noncross-validated PLS analysis gave a conventional r^2 of 0.986 with the S_{press} of 0.257, and the model yielded

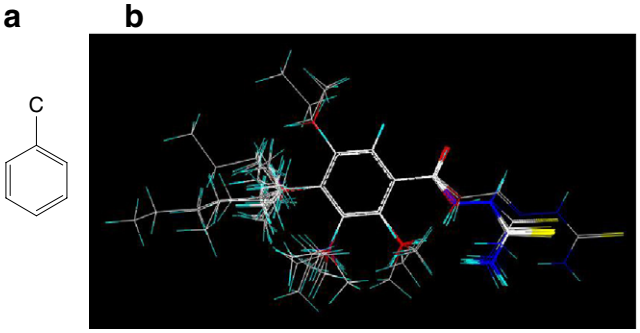


Figure 4. Superposition of compounds in the training and test sets using the common substructure-based alignment rules. (a) Common substructure-based alignment rules; (b) superposition of compounds in the training and test sets.

Table 2. Summary of results from the CoMFA and CoMSIA

	CoMFA	CoMSIA[a]	CoMSIA[b]	CoMSIA[c]	CoMSIA[d]	CoMSIA[e]
Compound	6	4	6	4	6	4
q^2	0.926	0.919	0.933	0.928	0.924	0.917
S_{press}	0.257	0.216	0.257	0.228	0.264	0.205
r^2	0.986	0.968	0.984	0.972	0.982	0.977
SEE	0.250	0.370	0.271	0.347	0.284	0.313
F	448.557	302.119	381.764	346.278	346.726	426.252
$Pr^2 = 0$	0.000	0.000	0.000	0.000	0.000	0.000
Fraction						
Steric	0.614	0.284	0.256	0.213	0.146	0.104
Electrostatic	0.386	0.716		0.462	0.364	0.241
Hydrophobic			0.354	0.324	0.203	0.160
Acceptor			0.390		0.287	0.175
Donor						0.320
r^2(bs)	0.993	0.977	0.989	0.981	0.990	0.983
SD[f]	0.176	0.312	0.213	0.280	0.204	0.262

[a] CoMSIA using the steric and electrostatic fields.
[b] CoMSIA using the steric, hydrophobic, and hydrogen bond acceptor fields.
[c] CoMSIA using the steric, electrostatic, and hydrophobic fields.
[d] CoMSIA using the steric, electrostatic, hydrophobic, and hydrogen bond acceptor fields.
[e] CoMSIA using the steric, electrostatic, hydrophobic, hydrogen bond acceptor, and hydrogen bond donor fields.
[f] Results from 100 runs of bootstrapped analyses.

Table 3. Results of analyses with cross-validation using five groups

	q^{2a}	
	CoMFA	CoMSIA[b]
Mean	0.917	0.927
High	0.944	0.938
Low	0.895	0.909

[a] Cross-validated q^2 using five groups with optimum number of components average of 30 runs.
[b] CoMSIA using the steric, hydrophobic and hydrogen bond acceptor fields.

six optimum components based on the CoMFA standard options. CoMSIA were performed using the following descriptor fields: steric, electrostatic, hydrophobic, and hydrogen bond acceptor and donor in different combinations. The same procedure used for the CoMFA study was also used for CoMSIA. The CoMSIA results are summarized in Table 2. A cross-validated value, q^2 of 0.933 and a conventional r^2 of 0.984 were obtained when the three descriptors steric, hydrophobic, and hydrogen bond acceptor were considered. The *F* value and SEE value were 381.764 and 0.271, respectively. These data also indicate that a reliable CoMSIA model was successfully constructed.

The LOO (Leave-One-Out) cross-validation method might lead to high q^2 values, which do not necessarily reflect a general predictability of the models. Therefore, cross-validation using five groups was performed. Because the randomness of the formation of the cross-validation groups may have a significant effect on the results, our cross-validation was performed 30 times for the analyses of CoMFA and CoMSIA using the steric, hydrophobic, and hydrogen bond acceptor fields based on the alignment. The results of cross-validation using five groups are reported in Table 3. Although the mean q^2 values were slightly lower as compared to the values obtained in the LOO method, both of the q^2 values were above 0.910. The results suggested that there was a good internal consistency in the underlying data set.

The predicted activities of the training set (45 compounds) and the test set (12 compounds) derived from the constructed 3D-QSAR model and the actual activities are listed in Table 1, and the correlation of CoMFA and CoMSIA models considering the three descriptors steric, hydrophobic, and hydrogen bond acceptor between the predicted activities and the actual activities is presented in Figures 5 and 6, respectively. The greatest advantage of CoMFA and CoMSIA is that the field effect on the target property can be viewed as 3D coefficient contour plots. The coefficient contour plots are helpful to identify important regions where any

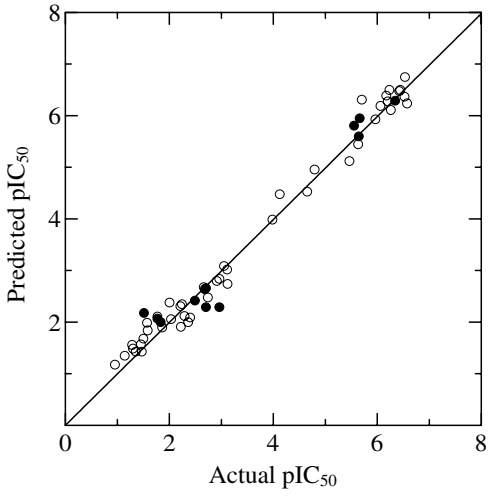


Figure 5. CoMFA predicted as experimental pIC_{50} values. Open circles represent predictions for the training set; solid circles represent predictions for the test set.

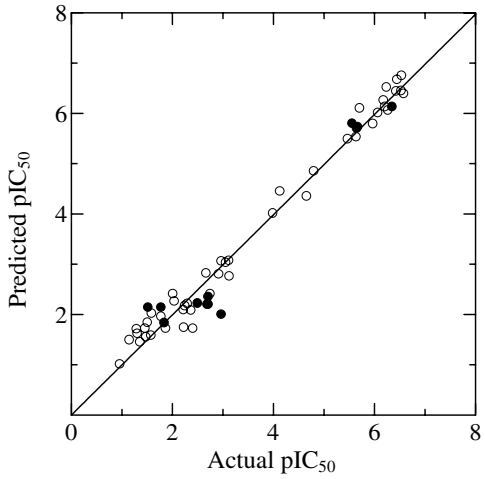


Figure 6. CoMSIA predicted as experimental pIC_{50} values. Open circles represent predictions for the training set; solid circles represent predictions for the test set.

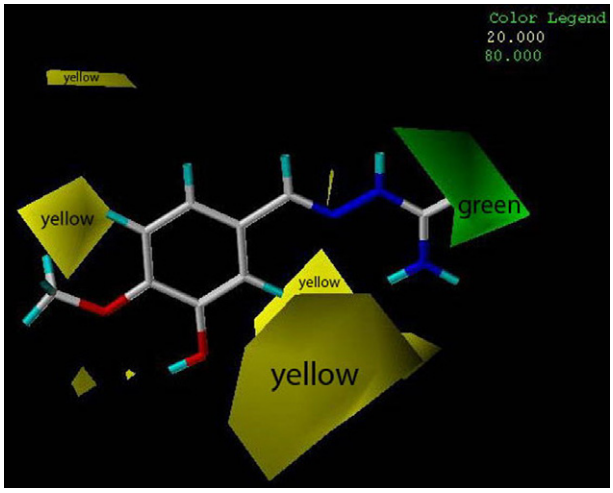


Figure 7. Steric maps from the CoMFA model. Compound **N12** (3-hydroxy-4-methoxy benzaldehyde thiosemicarbazone) is shown inside the field. Sterically favored areas (contribution level of 80%) are represented by green polyhedra. Sterically disfavored areas (contribution level of 20%) are represented by yellow polyhedra.

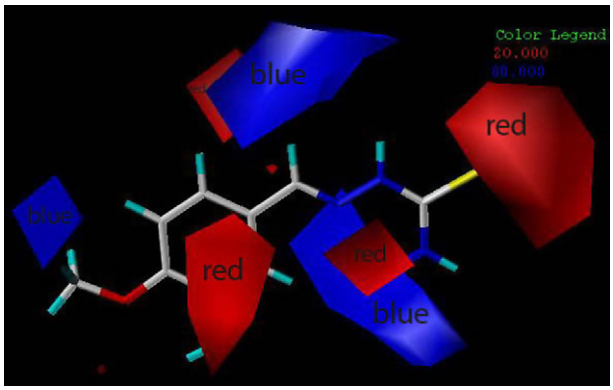


Figure 8. Electrostatic maps from the CoMFA model. Compound **N12** is shown inside the field. Blue contours (80% contribution) encompass regions where an increase of positive charge will enhance affinity, whereas in red contoured areas (20% contribution) more negative charges are favorable for binding properties.

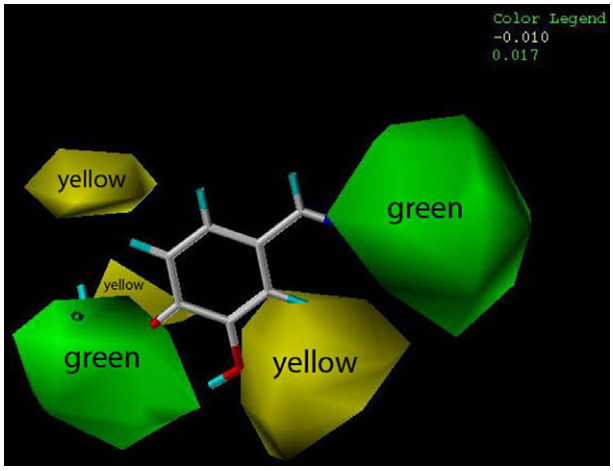


Figure 9. Steric maps from the CoMSIA model using the steric, hydrophobic, and hydrogen bond acceptor fields. Compound **N12** is shown inside the field. Green contours (actual values 0.017) enclose areas where steric bulk will enhance affinity, and yellow contours (actual values −0.010) highlight areas that should be kept unoccupied.

change in the steric, electrostatic, and hydrophobic fields may affect the biological activity, and they may also help to identify the possible interaction sites. In Figures 7–11, the isocontour diagrams of the field contributions ('stdev*coeff') of different properties obtained from the CoMFA and CoMSIA (using the steric, hydrophobic, and hydrogen bond acceptor fields) are illustrated together with exemplary ligands.

2.4.1. CoMFA contour map. The CoMFA contour map of steric contribution is depicted in Figure 7; the green and yellow polyhedra describe regions of space around the molecules where an increase in steric bulk enhances or diminishes the activity, respectively. The electrostatic contribution of CoMFA is illustrated in Figure 8; 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. Obviously, green contours embedded in yellow contours indicate that there exists an optimal value for the steric effect, and blue contours embedded in red contours indicate that there exists an optimal electrostatic effect.

In Figure 7, compound **N12** (3-hydroxy-4-methoxy benzaldehyde thiosemicarbazone) is displayed in the map to aid visualization. The green contour around the sulfur atom region of thiosemicarbazone indicates that steric bulk is favored there. In the present investigation, the three analogous classes of compound were studied: derivatives of benzaldehyde thiosemicarbazone, benzaldehyde, and benzoic acid. It is obvious that the presence of the outstretched thiosemicarbazone structure in benzaldehyde thiosemicarbazone derivatives (compounds **N1**–**N22**) resulted in a great increase of activity as compared to the two latter classes (com-

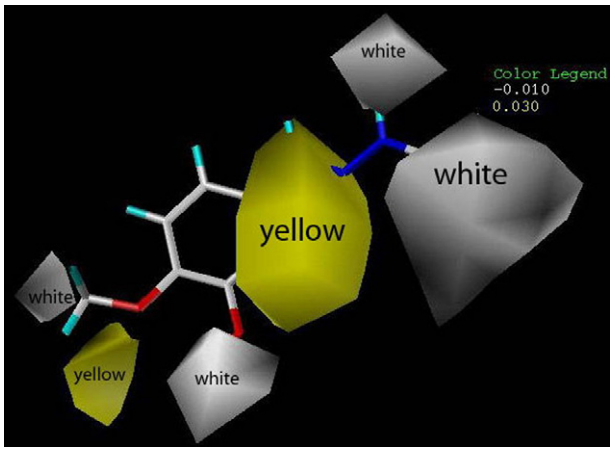


Figure 10. Hydrophobic field map of CoMSIA using the steric, hydrophobic, and hydrogen bond acceptor fields. Compound **N12** is shown inside the field. Yellow regions (actual values 0.030) indicate areas where hydrophobic groups increase activity, and white regions (actual values −0.010) indicate areas where hydrophobic groups decrease activity.

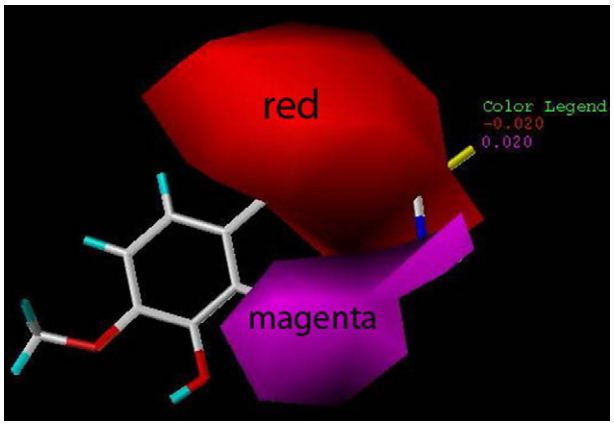


Figure 11. Hydrogen bond acceptor field map of CoMSIA using the steric, hydrophobic, and hydrogen bond acceptor fields. Compound **N12** is shown inside the field. Magenta regions (actual values 0.020) indicate areas where hydrogen bond acceptor groups increase activity, and red regions (actual values −0.020) indicate areas where hydrogen bond acceptor groups decrease activity.

pounds **Q1**–**Q17** and **S1**–**S18**). Additionally, when the *ortho*-position contains bulky substitutents such as the methoxy group, the bioactivities decreased consequently (compounds **Q8** and **N8**). In Figure 8, the red contours around the sulfur atom region of thiosemicarbazone and the benzene ring indicate that negative charges are favored in those regions. When positively charged groups are present at the *para*-position of the benzene ring, a significant increase in activity is observed. This is observed with compounds **Q15**–**Q17**, **S14**–**S18**, and **N17**–**N19**.

2.4.2. CoMSIA contour map. The steric contour map of the CoMSIA model is depicted in Figure 9. To aid in visualization, the most potent inhibitor **N12** has once again been overlaid on the map. The results of CoMSIA modeling are similar to those obtained with CoMFA, and the analysis is as described for the CoMFA maps. One feature of note is the large green polyhedron occupying the *para*-position of the benzene ring, suggesting that this is a favorable location for bulky substitutes. This is observed for compounds **Q15**–**Q17**, **S14**–**S18**, **N17**–**N19**, and **N21**. From both models, we can infer that the presence of bulky substitutes in the *para*-position is very important for inhibitory activity.

The hydrophobic field is presented in Figure 10. White and yellow contours highlight areas where hydrophilic and hydrophobic properties are preferred, respectively. A large yellow polyhedron surrounding the P2 position (see the region in Chart 1, III) indicates that hydrophobic groups in this area promote enhanced activity. Two white isopleths around the P1 position and the *meta*-position of the benzene ring reveal that hydrophilic groups at these areas also produce increased activity. This is borne out by the fact that the inhibitory activities of compounds **N1**–**N22** are 1000-fold higher than those of **Q1**–**Q17** and **S1**–**S18**.

Figure 11 highlights areas where hydrogen bond acceptors on the ligands promote or decrease binding affinities. Magenta isopleths encompass regions where hydrogen bond donors on the receptor site are expected, while proton acceptors in the ligands directing to red regions decrease the binding affinities. A red polyhedron around the P1 position, but not including the sulfur region, suggests that the subunit bearing hydrogen donors will increase the activity. A large magenta isopleth beside the *meta*-position indicates that hydrogen bond donors on the receptor promote activity.

The combination of CoMFA and CoMSIA methods leads to a better interpretation for QSAR at the 3D

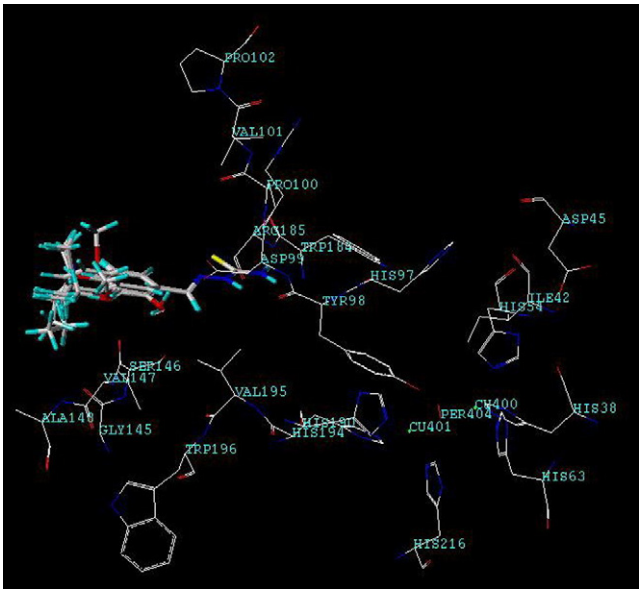


Figure 12. Position of docked 14 compounds in the active pocket of 1WX2.

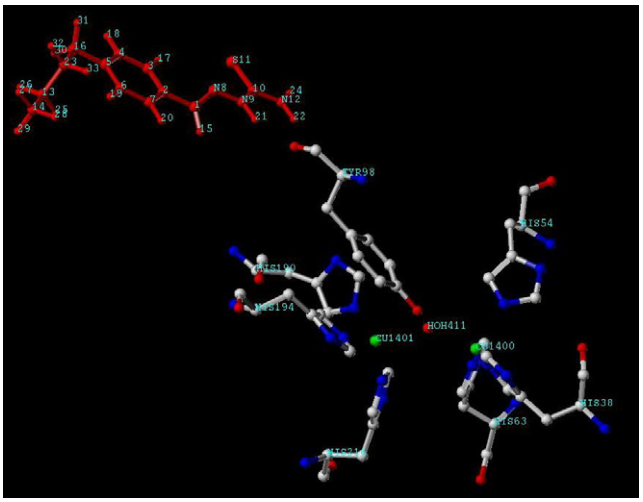


Figure 13. Interaction of compound **N18** with the active site of 1WX2.

level. However, it must be pointed out that the features derived from a comparative molecular field analysis depend on the structural variations inherently present in the selected data set. Selecting another structurally deviating data set might result in different results leading to alternative conclusions.

2.5. The interaction between benzaldehyde thiosemicarbazone derivatives and the 1WX2 active site

FlexX is a fast, flexible docking method that uses an incremental construction algorithm to place ligands into an active site. By default, FlexX produces 30 docked structures for each benzaldehyde thiosemicarbazone derivative. The conformation with the lowest docking energy in the most populated cluster is selected as the possible 'active' conformation against the 1WX2 active site. In the present study, 14 compounds were successfully docked into the 1WX2 site, the active conformation of which is presented in Figure 12.

To illustrate the detailed interaction mechanism, **N18** (4-butylbenzaldehyde thiosemicarbazone), the more potent inhibitor among the 14 compounds, was selected for further analysis. In the following discussion, all descriptions refer to **N18** unless otherwise noted. Figure 13 is a stereo view of **N18** in complex with the 1WX2 site model. The structure of **N18** contains a chain of atoms (H^{21}–N^{9}–C^{10}–N^{12}–H^{22}) spatially arranged in what might be termed a 'clamp' structure. The distance between both hydrogen atoms of the clamp and the carbonyl oxygen atom of Tyr^{98} is 1.991 Å, indicating the likely presence of a pair of hydrogen bonds. Formation of these two hydrogen bonds stabilizes the position of Tyr^{98}, preventing Tyr^{98} from participating in the interaction between PO and ORF378. ORF378 is consequently unable to serve as a transporter of Cu (II) ions to the catalytic site of PO, a necessary step in PO catalytic activity. Figure 12 illustrates an additional condition in which a butyl substituent *para* to the aldehyde group interacts with surrounding amino acid residues such as Asp^{99}, Gly^{145}, Ser^{146}, and Ala^{148}, indirectly leading to a loss of function in the active center of phenoloxidase.

In addition, the inhibition kinetics study revealed that the compound **N18** was a noncompetitive inhibitor to the substrate tyrosine. One of several possible noncompetitive inhibitory mechanisms is interaction of the inhibitor with the primary amino acid residues surrounding the active site. As we expected, the molecular docking results suggest the compound **N18** actually interacts with Tyr^{98}. Hence, the interaction between the 'clamp' structure of **N18** and Tyr^{98} may be the principal mechanism of inhibition in the 1WX2 active site.

3. Conclusion

In the present study, the inhibitory activities and IC_{50} values of 57 compounds belonging to the benzaldehyde thiosemicarbazone, benzaldehyde, and benzoic acid families against phenoloxidase from *Pieris rapae* (Lepidoptera) larvae were determined, the inhibitory kinetics of 4-butylbenzaldehyde thiosemicarbazone against PO were also measured in air-saturated solutions for the oxidation of L-DOPA. The results indicated that the compound is a reversible noncompetitive inhibitor.

The most important was the CoMFA and CoMSIA 3D-QSAR analyses were applied to predict the inhibitory activity of three sets of analogs. The QSAR models gave good statistical results in terms of q^2 and r^2 values. The CoMFA model provided the most significant correlation of steric and electrostatic fields with biological activities. The effects of the steric, hydrophobic, and hydrogen bond acceptor fields around the aligned molecules on their activities were clarified by analyzing the CoMSIA contour maps. The information derived from this study provides a tool for predicting the affinity of related compounds, and for guiding further structural modification and synthesis of potent new PO inhibitors. Additionally, a set of benzaldehyde thiosemicarbazone derivatives was successfully docked into the PO site using the FlexX program, demonstrating that the inhibition mechanism of **N18** on PO may involve the 'clamp' structure of the inhibitor interacting with Tyr^{98} of the active site.

4. Materials and methods

4.1. Insects and reagents

Pieris rapae larvae were reared on Chinese cabbage *Brassica parachinesi* (Bailey) in a greenhouse at 25 ± 1 °C with a 14:10 h light:dark photoperiod. The 5th instar larvae were gathered for the experiments. L-3,4-dihydroxyphenylalanine (L-DOPA) and dimethyl sulfoxide (DMSO) were purchased from Aldrich Chemical Co. (Milwaukee, WI, USA). Sephadex G-100 was purchased from Amersham Pharmacia Biotech. (Uppsala, Sweden). The title compounds **Q1**–**Q17** and **S1**–**S18** (see the structures in Chart shown in Table 1) were purchased from Alfa Aesar (Ward Hill, MA, USA), Aldrich

Chemical Co. (Milwaukee, WI, USA) and Fluka (Milwaukee, WI, USA). Compounds **N1**–**N22** were designed and synthesized in our laboratory and purified by chromatography over silica gel. Their chemical structures were confirmed by ^{1}H NMR, mass spectroscopy, and elemental analysis. All other reagents were local products of analytical grade. The water was re-distilled and ion-free.

4.2. Inhibitory kinetics of N18 against PO

Extraction and purification of PO from *P. rapae* L., and the inhibitory kinetics of **N18** (4-butylbenzaldehyde thiosemicarbazone) against PO, was tested in vitro according to the method described previously.[13] The extent of inhibition was expressed as the chemical concentration to inhibit 50% of the enzyme activity (IC_{50}). The inhibition type was determined by the Lineweaver–Burk plot. The inhibition constant was determined by the plot of the apparent Michaelis–Menten constant versus the concentration of the inhibitor.

4.3. Data set and structures

The biological activities (expressed as IC_{50}) of the 57 compounds were converted to pIC_{50} ($-\log IC_{50}$) for construction of the 3D-QSAR models. The structures and biological activities of the compounds are contained in Table 1. The 3D-QSAR models were generated using a training set of 45 molecules, and the predictive power of the resulting models was evaluated using a test set of 12 molecules (Table 1). The test compounds were selected manually such that the data set included diverse structures and a wide range of activity. The PO crystal structure used in this research was PDB entry 1WX2, retrieved from the Protein Data Bank (PDB http://www.rcsb.org/pdb), and is the suitable PO crystal structure currently available. The structure consists of the enzyme complexed with the caddie protein ORF378.

4.4. Molecular modeling and alignment rule

Molecular modeling studies were performed using SYBYL 6.9 running on a Silicon Graphics Origin 300 server. The compounds **Q1**–**Q17**, **S1**–**S18**, and **N1**–**N22** 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 the CoMFA and CoMSIA models. A common substructure-based alignment was adopted in the present study. It attempted to align molecules to the template molecule on a common backbone, as illustrated in Figure 4. For manual alignment of the inhibitors, the structure of compound **N12** was used as a template.

4.5. 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 the 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.

4.6. 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

$$A^q_{F,k}(j) = \sum_i \omega_{\text{probe},k}\omega_{ik}\mathrm{e}^{-\alpha r^2_{iq}}$$

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; 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, hydrophobic, hydrogen bond accepting probe atom with a radius of 1 Å. A Gaussian-type distance dependence was considered 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,[16–18] an extension of multiple 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), and *F* values and standard error of estimate (SEE) values were computed according to the definitions in SYBYL and are listed in Table 2. In Table 2, $Pr^2 = 0$ 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 have not occurred by chance.

In SYBYL/QSAR, the intensity of the cross-validation process is controlled by selecting the number of groups or the number of times the cross-validation step is to

be carried out while predicting all rows (at each stage of model development). To perform an even more rigorous statistical test, several runs of cross-validation using five groups were done in which each target property value is predicted by a model based on about four-fifths or 80% of the available data. To further obtain statistical confidence limits for the analysis, bootstrapping analysis (100 runs) was performed.

A common test to check the consistency of the models is to scramble the biological data and repeat the model derivation process, thus allowing detection of possible chance correlations. After our data set was randomized, very low or negative q^2 value were observed in all of the PLS analyses.

4.7. Molecular docking using FlexX

FlexX is a fast, flexible docking method that uses an incremental construction algorithm to place ligands into an active site.[19,20] The coordinates of the phenoloxidase crystal structure (1WX2) were recovered from the PDB. 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, and hydrogen atoms were generated in standard geometry, whereas metal atoms were always kept in the protein description. We shall give further target-specific details below. In the present study, 14 compounds of **N1**–**N22** were selected and docked into the PO active site using the default FlexX parameter settings.

Acknowledgment

The present investigation was financially supported by the National Natural Science Foundation of PR China (Grant No. 30571237) for Wan-Chun LUO.

References and notes

1. Sánchez-Ferrer, A.; Rodríguez-López, J. N.; García-Cánovas, F.; García-Carmona, F. *Biochim. Biophys. Acta* **1995**, *1247*, 1.
2. Chase, M. R.; Raina, K.; Bruno, J.; Sugumaran, M. *Insect Biochem. Molec.* **2000**, *30*, 953.
3. Chen, Q. X.; Liu, X. D.; Huang, H. *Biochemistry-Moscow* **2003**, *68*, 644.
4. Oetting, W. S. *Pigm. Cell Res.* **2000**, *13*, 320.
5. Xu, Y.; Stokes, A. H.; Roskoski, R. J.; Vrana, K. E. *J. Neurosci. Res.* **1998**, *54*, 691.
6. Asanuma, M.; Miyazaki, I.; Ogawa, N. *J. Neurotox. Res.* **2003**, *5*, 165.
7. Ashida, M.; Brey, P. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 10698.
8. Matoba, Y.; Kumagai, T.; Yamamoto, A.; Yoshitsu, H.; Sugiyama, M. *J. Biol. Chem.* **2006**, *281*, 8981.
9. Wang, S. D.; Luo, W. C.; Gao, X. X. *Agr. Sci. China* **2004**, *3*, 923.
10. Wang, S. D.; Luo, W. C.; Xu, S. J.; Ding, Q. *Pestic. Biochem. Phys.* **2005**, *82*, 52.
11. Wang, S. D.; Luo, W. C.; Xiao, T.; Xue, C. B. *Biopestic. Int.* **2005**, *1*, 82.
12. Wang, X. Y.; Liu, C. Y.; Zhang, J. D.; Luo, W. C. *Insect Sci.* **2005**, *12*, 231.
13. Xue, C. B.; Luo, W. C.; Chen, Q. X.; Wang, Q.; Ke, L. N. *Insect Sci.* **2006**, *13*, 251.
14. Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
15. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.
16. Wold, S.; Albano, C.; Dunn, W. J., III; Edlund, U.; Esbensen, K.; Geladi, P.; Hellberg, S.; Johanson, E.; Lindberg, W.; Sjostrom, M. *NATO ASI Ser., Ser. C* **1984**, *138*, 17.
17. Wold, S.; Rhue, A.; Wold, H.; Dunn, W. J. I. *SIAM J. Sci. Stat. Comput.* **1984**, *5*, 735.
18. Clark, M.; Cramer, R. D., III *Quant. Struct.-Act. Relat.* **1993**, *12*, 137.
19. Rarey, M.; Kramer, B.; Lengauer, T.; Kleb, G. *J. Mol. Biol.* **1996**, *261*, 470.
20. Rarey, M.; Kramer, B.; Lengauer, T. *J. Comput. Aid. Mol. Des.* **1997**, *11*, 369.
21. Yang, G. F.; Huang, X. Q. *Curr. Pharm. Design.* **2006**, *12*, 4601.
22. Wei, D. G.; Yang, G. F.; Wan, J.; Zhan, C. G. *J. Agric. Food Chem.* **2005**, *53*, 1604.
23. Yang, G. F.; Jiang, X. H.; Yang, H. Z. *Pest Manag. Sci.* **2002**, *58*, 1063.
24. Yang, G. F.; Liu, H. Y.; Yang, H. Z. *Pestic. Sci.* **1999**, *55*, 1143.
25. Yang, G. F.; Jiang, X. H.; Ding, Y.; Yang, H. Z. *Acta Chim. Sinica* **2002**, *60*, 134.