ELSEVIER

A DFT-based QSARs study of protoporphyrinogen oxidase inhibitors: phenyl triazolinones

Li Zhang, Jian Wan and Guangfu Yang*

Key Laboratory of Pesticide & Chemical Biology (CCNU), Ministry of Education; College of Chemistry, Central China Normal University, Wuhan 430079, PR China

Abstract—The equilibrium geometries, electronic structures, and electrostatic potentials of a series of substituted phenyl triazolinones, a kind of important protoporphyrinogen oxidase (Protox) inhibitors, had been investigated using density functional theory (DFT) method at the B3LYP/6-31G(d,p) basis set. The quantum chemical descriptors, such as energy difference (ΔE) between the lowest unoccupied molecular orbital and the highest occupied molecular orbital, electrophilic and nucleophilic frontier electron density (f_i^E and f_i^N), and net atomic charge (Q_i), were computed at the same DFT level. Based on these useful quantum chemical descriptors, the quantitative structure–activity relationships was carried out and the results showed that descriptors, $Q_{C_{11}}$, $f_{N_5}^E$, $f_{C_{10}}^N$, $f_{O_6}^E$, $f_{C_{11}}^N$, and ΔE, were most likely to be responsible for the in vitro biological activity and the greenhouse pre-emergence activity of phenyl triazolinones. The descriptors accounted for 77–86% of the variation in the in vitro biological activity among the herbicidal phenyl triazolinone analogs **1**–**26** (except compounds **19** and **20**). The results of the regression analysis showed that the activity was parabolically related not only with the descriptor $f_{O_6}^E$, but also with the descriptor $f_{C_{11}}^N$. The optimum values of the terms $f_{O_6}^E$ and $f_{C_{11}}^N$ were about 11.15 and 0, respectively. Studies also showed that compound **19** exhibiting the highest in vitro activity mimicked the three-ring portion of protoporphyrinogen IX (Protogen). The present work had proved that the DFT-based quantum chemical descriptors could lead to the better correlation relationship than that the PM3-based electronic descriptors, therefore, DFT-based QSARs could be expected to help facilitate the design of additional substituted phenyl triazolinone derivatives of Protox inhibitors with good biological activity.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

Protoporphyrinogen oxidase (Protox, EC 1.3.3.4), the last common enzyme in chlorophylls biosynthesis,[1] has been identified as the target site of action for several photodynamically active porphyric herbicides of diphenyl ethers, phenyl heterocycles, and heterocyclic carboxamides.[2] These known Protox inhibitors apparently compete with protoporphyrinogen IX (Protogen) at or near the catalytic site on the enzyme, so Protogen accumulates in plastids and diffuses through the plastidial membranes into the cytosol, where it is oxidized to protoporphyrin (Proto) by a plasma membrane-bound Protox.[3] The Proto reacts with light to produce singlet oxygen, leading to lipid peroxidation and the destruc-

tion of cellular components.[4] The attributes of low application rates, good crop selectivity, low residue and environmental safety exhibited by these compounds are important characteristics for modern agrochemicals, which have led to the rapid success of Protox inhibitors as herbicidal products and attracted a worldwide research commitment.

Phenyl triazolinones are one of the most important Protox inhibitors. A series of studies on substituted phenyl triazolinones (Fig. 1) were carried out by Theodoridis.[5–8] They qualitatively proposed the optimum
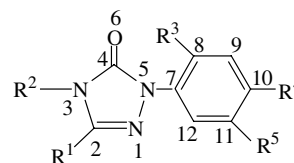
Figure 1. General structure of phenyl triazolinone compounds.

herbicidal activity at positions 2 (R^1) and 3 (R^2) of the heterocyclic ring and positions 8 (R^3) and 10 (R^4) of the phenyl ring and hypothesized that tricyclic protox inhibitors mimicked three rings of Protogen[6]. In addition, they investigated the structure–activity requirements at position 11 (R^5) of the phenyl ring using the QSAR analysis. Two physicochemical descriptors, the hydrophobicity term π and the Sterimol term B_1, were used to correlate with biological activity pI_{50} (-log concentration in mol L^{-1}) for 14 substituents at position 11 (R^5) of phenyl ring. However, this set of physicochemical descriptors failed to predict the biological activities for the substituents at position 10 (R^4). Subsequently, Reddy[9] et al. hypothesized that bicyclic Protox inhibitors (e.g., diphenyl ethers, isoxazolecarboxamides) mimicked two rings of Protogen and that tricyclic Protox inhibitors (e.g., phenoxyphenoxy triazolinones) mimicked three rings of Protogen. These earlier QSAR studies were limited to the use of the physicochemical descriptors or the electronic structure descriptors based on semiempirical molecular orbital (MO) calculations.

As well known, QSAR studies could also help to understand molecular mechanism of biological activities of environmentally important molecules including medicines and pesticides. It is essential for this purpose to obtain QSAR equations with high-quality descriptors, because the success of a QSAR model is highly dependent upon the choice of descriptors. So exploring the use of reliable descriptors, especially descriptors based on reasonably high level electronic structure calculations, could lead to significantly better QSAR results. Semiempirical molecular orbital (MO) calculations have been used to obtain electronic descriptors for many years. However, the latest development of the computer technology and software of electronic structure theory allows calculating quantum chemical descriptors at first-principles levels, such as DFT, with higher accuracy including some effective consideration of electron correlation effects. We are particularly interested in testing quantum chemical descriptors based on density functional theory (DFT), because DFT has played an important role in many research areas of modern computational chemistry.

There have been dramatic changes in computational chemistry over the last two decades. Electronic structure theory has become an effective and powerful tool for use in predicting the properties of a wide range of molecules including geometries, energetics, reactivity, and spectroscopic properties. One of the major reasons for the acceleration of the use of electronic structure theory in predicting molecular properties for larger molecules has been the development of density functional theory (DFT). The combination of relatively low computational cost with reasonable accuracy has led to the successful application of the DFT method to the prediction of a broad range of properties of molecules. So, DFT has emerged as a practical and versatile tool to obtain accurate information on molecular systems of chemical interest.[10–19] This approach, which includes the dynamic correlation effects, represents a valid alternative to the Hartree–Fock (HF) theory, or to more sophistical post-HF methods, such as those of the Møller–Plesset theory, coupled-cluster theory, or configuration interaction approach, that are highly demanding in terms of CPU time. One of the most popularly used hybrid density functionals is Becke's three-parameter hybrid exchange functional and the Lee–Yang–Parr correlation functional (B3LYP).[13,14] DFT has become a powerful computational approach to investigate the precise electronic characteristic of molecular structure, which is the key factor to determine the interactions between the receptors and the ligands. However, DFT-based quantum chemical descriptors were seldom used to carry out quantitative structure–activity relationships study.[20]

As a part of our research interest on the structural modification and molecular design of Protox inhibitors, we carried out a DFT-based quantitative structure–activity relationships (DFT/QSARs) study for both substitutes at positions 10 and 11 of the phenyl triazolinone derivatives (Fig. 1). The equilibrium geometries, electronic structures, and electrostatic potentials of a series of substituted at positions 10 and 11 phenyl triazolinones were investigated by DFT method. The quantum chemical descriptors,[21] such as energy differences between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), electrophilic or nucleophilic frontier electron densities and net atomic charges, were derived from the DFT calculations (Table 1). Based on these important quantum chemical descriptors, a quantitative structure–activity

Table 1. Quantum chemical descriptors of phenyl triazolinone compounds

No	Descriptors						No	Descriptors					
	$Q_{C_{11}}$	$f^E_{O_6}$	$f^E_{N_5}$	$f^N_{C_{10}}$	$f^N_{C_{11}}$	ΔE		$Q_{C_{11}}$	$f^E_{O_6}$	$f^E_{N_5}$	$f^N_{C_{10}}$	$f^N_{C_{11}}$	ΔE
1	0.35	14.23	21.86	23.59	0.41	5.27	14	0.03	11.26	18.00	26.52	4.50	5.24
2	0.31	13.73	21.35	23.18	0.30	5.29	15	0.40	3.63	6.24	20.47	0.61	5.05
3	0.39	13.69	21.29	23.51	0.38	5.29	16	0.42	6.86	11.79	21.76	0.28	5.12
4	0.36	7.70	13.31	22.18	0.17	5.15	17	0.37	4.53	7.58	20.10	0.60	5.09
5	0.05	15.40	23.30	24.56	0.78	5.33	18	0.38	15.80	23.60	0.41	0.07	4.31
6	0.39	11.36	17.60	23.46	0.56	5.16	19	0.37	6.25	10.79	21.13	0.26	5.09
7	0.37	14.65	22.43	24.15	0.90	5.28	20	−0.29	11.67	18.98	0.59	1.39	5.33
8	0.16	14.59	22.46	24.38	0.55	5.34	21	0.23	2.11	4.11	0.52	0.41	4.66
9	−0.10	15.77	23.77	24.43	0.43	5.35	22	−0.10	10.48	17.20	10.80	5.94	5.36
10	0.38	13.66	20.44	23.94	1.07	5.18	23	0.07	14.09	22.03	23.47	0.52	5.34
11	0.38	5.20	9.07	21.19	0.36	5.09	24	−0.24	16.66	24.77	14.60	2.42	5.54
12	0.05	14.67	22.23	23.96	0.36	5.24	25	−0.14	18.66	26.49	7.04	6.32	4.06
13	−0.11	14.06	21.46	24.10	0.39	5.23	26	−0.25	10.15	16.99	12.70	4.11	5.50

relationships (QSARs) study had been carried out for both substitutes at positions 10 and 11 of the title compounds. To our knowledge, this is the first application based on DFT/QSAR method in the pesticide field.

2. Materials and methods

2.1. Test compounds

All phenyl triazolinone compounds and their in vitro biological activity pI_{50} values and in vivo biological activity ED_{90} values used for the analysis were taken from literature[8] and listed in Table 2.

2.2. Calculated methods

The geometry optimizations were performed by the DFT[10–19] method using the B3LYP[14,15] functional. A standard valence double-zeta with polarization function 6-31G(d,p) basis set[22] was used for all kinds of atoms involved in phenyl triazolinones. Vibrational frequencies were computed at the B3LYP/6-31G(d,p) levels to characterize stationary points. Net atomic charges were derived from ChelpG[23] method that produces charges fit to the electrostatic potential at points selected. All calcul-

ations were performed using the Gaussian 03 program[24] on a Silicon Graphics Origin 300 server.
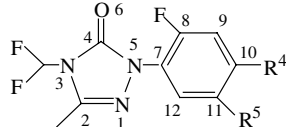
2.3. Quantum chemical descriptors

Quantum chemical descriptors (listed in Table 1) taken from DFT calculations were used to analyze variations in the herbicidal activity of compounds **1**–**26** by Hansch–Fujita method. The minimum energy conformations were selected as the bioactive conformations and used to calculate electronic descriptors such as ΔE (the difference between LUMO and HOMO orbital energy), Q_{C11} (net atomic charges of carbon atom at position 11), $f_{N_5}^{E}$ (the electrophilic frontier electron density of 5-position nitrogen atom), $f_{C_{10}}^{N}$ (the nucleophilic frontier electron density of 10-position carbon atom), $f_{C_{11}}^{N}$ (the nucleophilic frontier electron density of 11-position carbon atom), and $f_{O_6}^{E}$ (the frontier electron density of HOMO at the 6-position oxygen atom),[21] etc.

Atomic frontier electron density is defined as:

$$f_i^{E} = \sum (C_{HOMO,i})^2 \times 100$$
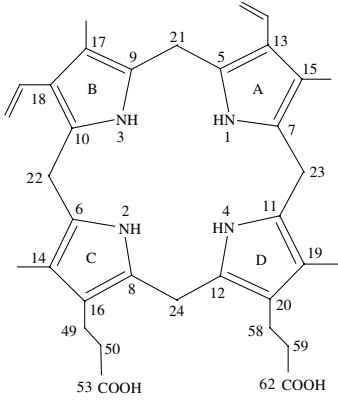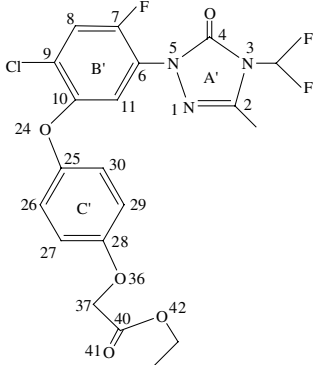
$$f_i^{N} = \sum (C_{LUMO,i})^2 \times 100$$

Table 2. Structure and herbicidal activity of phenyl triazolinone derivatives

No	Substituents		pI_{50}							log 1/ED_{90}		
	R^4	R^5	Obsd	Eq. 4	δ_1[a]	Eq. 5	δ_2[a]	Eq. 6	δ_3[a]	Obsd	Eq. 9	δ_4[a]
1	Cl	OCH₂CCH	7.60	7.43	0.17	6.96	0.64	7.36	0.24	1.52	1.07	0.45
2	Cl	OCH₂CHCH₂	7.50	7.26	0.24	6.95	0.55	7.20	0.30	0.90	0.96	−0.06
3	Cl	OCH₃	7.50	7.53	−0.03	6.97	0.53	7.43	0.07	0.90	1.13	−0.23
4	Cl	OH	7.20	6.73	0.47	7.00	0.20	6.88	0.32	0.52	0.48	0.04
5	Cl	CH₂OCH₃	7.10	6.59	0.51	6.79	0.31	6.69	0.41	0.30	0.38	−0.08
6	Cl	NHSO₂C₂H₅	7.10	7.19	−0.09	7.24	−0.14	7.24	−0.14	1.21	0.82	0.39
7	Cl	OCOCH₃	7.10	7.54	−0.44	6.91	0.19	7.47	−0.37	0.52	1.12	−0.60
8	Cl	CH₃	7.00	6.89	0.11	6.87	0.13	6.93	0.07	0.30	0.59	−0.29
9	Cl	H	6.80	6.14	0.66	6.72	0.08	6.31	0.49	0.30	0.06	0.24
10	Cl	NHSO₂CH₃	6.70	7.40	−0.70	7.12	−0.42	7.37	−0.67	1.52	0.98	0.54
11	Cl	OC₆H₅	6.60	6.51	0.09	6.66	−0.06	6.65	−0.05	0	0.24	−0.24
12	Cl	Cl	6.50	6.44	0.06	6.98	−0.48	6.61	−0.11	0	0.34	−0.34
13	Cl	Br	6.50	5.84	0.66	7.06	−0.56	6.16	0.34	0	−0.13	0.13
14	Cl	C₆H₅	6.30	6.05	0.25	6.52	−0.22	6.44	−0.14	−0.30	−0.32	0.02
15	Cl	O-(4-NHSO₂Et)Phenyl	6.60	6.39	0.21	6.34	0.26	6.53	0.07	—	—	—
16	Cl	O-(4-Methoxy)phenyl	6.70	6.83	−0.13	6.91	−0.21	6.92	−0.22	—	—	—
17	Cl	O-(4-Cl)Phenyl	6.70	6.41	0.29	6.45	0.25	6.51	0.19	—	—	—
18	Cl	O-(4-NO₂)Phenyl	6.80	6.86	−0.06	6.58	0.22	6.39	0.41	—	—	—
19	Cl	O-(4-OCH₂CO₂Et)Phenyl	9.00	6.57	2.43	6.83	2.17	6.71	2.29	—	—	—
20	O-(4-Cl)Benzyl	H	7.90	5.09	2.81	5.39	2.51	4.38	3.52	1.21	1.20	0.01
21	O-(4-Cl)Benzyl	NH₂	4.90	5.34	−0.44	5.17	−0.27	4.97	−0.07	—	—	—
22	O-(4-Cl)Benzyl	Cl	5.00	5.64	−0.64	4.67	0.33	5.30	−0.30	—	—	—
23	Br	H	6.10	6.54	−0.44	6.86	−0.76	6.61	−0.51	—	—	—
24	F	H	5.30	5.92	−0.62	5.42	−0.12	5.50	−0.20	—	—	—
25	NO2	H	5.20	5.18	0.02	5.23	−0.03	5.48	−0.28	—	—	—
26	OCH(CH₃)₂	H	5.10	5.23	−0.13	5.34	−0.24	4.97	0.13	—	—	—

[a] δ: Difference between observed and calculated values.

Table 3. Some selected torsion angles in the most stable conformation of the protoporphyrinogen IX and compound **19**

Protoporphyrinogen IX   Compound **19**

Atoms	Degree	Atoms	Degree
A–B rings		C–D rings	
N1–C5–C21–C9	30.82	N2–C8–C24–C12	34.20
C13–C5–C21–C9	−146.97	C16–C8–C24–C12	−147.18
C5–C21–C9–N3	71.84	C8–C24–C12–N4	68.98
C5–C21–C9–C17	−106.89	C8–C24–C12–C20	−107.12
B–C rings		D–A rings	
N3–C10–C22–C6	−37.43	N4–C11–C23–C7	−42.55
C18–C10–C22–C6	140.08	C19–C11–C23–C7	139.62
C10–C22–C6–N2	−63.81	C11–C23–C7–C15	117.21
C10–C22–C6–C14	115.60	C11–C23–C7–N1	−61.21
A′–B′ rings		B′–C′ rings	
N1–N5–C6–C11	42.03	C11–C10–O24–C25	−7.06
C4–N5–C6–C7	50.73	C9–C10–O24–C25	174.58
		C10–O24–C25–C30	−79.20
		C10–O24–C25–C26	105.08

where f_i^E is the electrophilic frontier electron density of atom i at HOMO and $C_{HOMO,i}$ is the coefficient of the atomic orbital X_i in the HOMO. f_i^N is the nucleophilic frontier electron density of atom i at LUMO and $C_{LUMO,i}$ is the coefficient of the atomic orbital X_i in the LUMO. f_i represents the reactivity of different atoms within a molecule (Table 3).

3. Results and discussion

3.1. QSARs analysis and model validation

Variations in the herbicidal activity of phenyl triazolinone compounds (Table 2, **1**–**26**) were analyzed using the quantum chemical descriptors listed in Table 1 and the obtained equations with best correlation coefficients were listed in Table 4.

In these and the following equations, n is the number of compounds, q^2 is the crossvalidated correlation coefficient using the Leave-One-Out (LOO) method. s_{press} is the root mean predictive error sum of squares, r^2 is the correlation coefficient, s is the standard error, and F is F-test value. The figures in parentheses under each coefficient are the 95% confidence intervals of the regression coefficient.

From Table 4, we can conclude that the correlation coefficients of these equations are not satisfied and compounds **19** and **20** are two outliers. After omitting compounds **19** and **20**, Eqs. 4–6 with significantly improved correlation coefficients were obtained as follows:

$$pI_{50} = 0.3968 + 3.3413 Q_{C_{11}} + 0.8530 \Delta E + 0.0960 f_{O_6}^E$$
$$\phantom{pI_{50} = } (0.0852) \quad\quad (0.4198) \quad\quad (0.2574) \quad\quad (0.0213)$$

$$n = 24, \quad q^2 = 0.68, \quad s_{press} = 0.496, \quad r^2 = 0.77,$$
$$s = 0.417, \quad F = 22.90 \tag{4}$$

Table 4. Correlation between some quantum chemical parameters and biological activity

No	Regression equations	n	r^2	s	F-test
1	$pI_{50} = 1.1075 + 2.4970 Q_{C_{11}} + 0.8805 \Delta E + 0.0516 f_{O_6}^E$	26	0.35	0.8239	4.01
2	$pI_{50} = 10.5197 + 0.0253 f_{C_{10}}^N - 0.0525 \left(f_{C_{11}}^N \right)^2 - 1.2679 \Delta E + 0.5640 f_{O_6}^E - 0.0267 \left(f_{O_6}^E \right)^2$	26	0.55	0.7238	4.82
3	$pI_{50} = 5.3458 + 1.9441 Q_{C_{11}} + 0.0305 f_{N_5}^E + 0.0232 f_{C_{10}}^N$	26	0.31	0.8505	3.31

$$pI_{50} = 11.1270 + 0.0673 f_{C_{10}}^{N} - 0.0409\left(f_{C_{11}}^{N}\right)^2 - 1.4226\Delta E$$
$$\phantom{pI_{50} = }{}_{(0.0845)}{}_{(0.0169)}{}_{(0.0087)}{}_{(0.3807)}$$

$$+ 0.3390 f_{O_6}^{E} - 0.0152\left(f_{O_6}^{E}\right)^2$$
$${}_{(0.1308)}{}_{(0.0063)}$$

$$n = 24, \quad q^2 = 0.64, \quad s_{press} = 0.556, \quad r^2 = 0.80,$$
$$s = 0.414, \quad F = 14.44 \tag{5}$$

$$pI_{50} = 4.1272 + 2.6437 Q_{C_{11}} + 0.0520 f_{N_5}^{E} + 0.0495 f_{C_{10}}^{N}$$
$$\phantom{pI_{50} = }{}_{(0.0673)}{}_{(0.3409)}{}_{(0.0125)}{}_{(0.0096)}$$

$$n = 24, \quad q^2 = 0.76, \quad s_{press} = 0.434, \quad r^2 = 0.86,$$
$$s = 0.330, \quad F = 40.67 \tag{6}$$

The development of these equations and the intercorrelation of variables were shown in Tables 5–10, respectively. The in vitro biological activities calculated by Eqs. 4–6 were listed in Table 2, respectively.

A stepwise regression analysis for the development of Eq. 4 in Table 5 indicates $Q_{C_{11}}$ term accounts for 44% of the biological variance. Adding ΔE and $f_{O_6}^{E}$ terms improves the explained variance to 77%. The positive coefficient of the $Q_{C_{11}}$ term indicates that the more the positive charges of the carbon at position 11, the higher the activity, which suggests that the position 11 of the phenyl ring should be occupied by electron-withdrawing substituents. The positive coefficient of the ΔE term indicates the greater the energy differences between the LUMO and the HOMO, the higher the activity. In addition, Table 5 shows that the introduction of the $f_{O_6}^{E}$ term led to a significant improvement in the correlation and its positive coefficient shows that the higher the electrophilic frontier electron density of 6-position oxygen atom, the higher the activity, which shows that the ability of the oxygen atom to donate electrons has an important effect on the herbicidal activity.

Table 5. Development of QSAR of Eq. 4

Intercept	$Q_{C_{11}}$	ΔE	$f_{O_6}^{E}$	s	r	F-test
6.0916	2.3344			0.6279	0.66	17.19
1.9955	2.5363	0.7894		0.5786	0.74	12.58
0.3968	3.3413	0.8530	0.0960	0.4174	0.88	22.90

Table 6. Correlation matrix for variables used to drive Eq. 4

	$Q_{C_{11}}$	ΔE	$f_{O_6}^{E}$
ΔE	0.0299	1.0000	
$f_{O_6}^{E}$	0.1787	0.0006	1.0000

Table 7. Development of QSAR of Eq. 5

Intercept	$f_{C_{10}}^{N}$	$\left(f_{C_{11}}^{N}\right)^2$	ΔE	$f_{O_6}^{E}$	$\left(f_{O_6}^{E}\right)^2$	s	r	F-test
5.1183	0.0711					0.6308	0.66	16.83
5.6645	0.0517	−0.0332				0.5353	0.78	16.46
10.3936	0.0825	−0.0339	−1.0344			0.4692	0.85	16.73
9.9973	0.0776	−0.0370	−0.9990	0.0276		0.4632	0.86	13.25
11.1270	0.0673	−0.0409	−1.4226	0.3390	−0.0152	0.4138	0.89	14.44

Table 8. Correlation matrix for variables used to drive Eq. 5

	$f_{C_{10}}^{N}$	$\left(f_{C_{11}}^{N}\right)^2$	ΔE	$f_{O_6}^{E}$	$\left(f_{O_6}^{E}\right)^2$
$\left(f_{C_{11}}^{N}\right)^2$	0.1537	1.0000			
ΔE	0.4439	0.0790	1.0000		
$f_{O_6}^{E}$	0.0159	0.0349	0.0006	1.0000	
$\left(f_{O_6}^{E}\right)^2$	0.0002	0.0422	0.0114	0.9596	1.0000

Table 9. Development of QSAR of Eq. 6

Intercept	$Q_{C_{11}}$	$f_{N_5}^{E}$	$f_{C_{10}}^{N}$	s	r	F-test
6.0916	2.3344			0.6279	0.66	17.19
4.6882	3.1383	0.0691		0.4925	0.82	21.34
4.1272	2.6437	0.0520	0.0495	0.3298	0.93	40.67

Table 10. Correlation matrix for variables used to drive Eq. 6

	$Q_{C_{11}}$	$f_{N_5}^{E}$	$f_{C_{10}}^{N}$
$f_{N_5}^{E}$	0.1834	1.0000	
$f_{C_{10}}^{N}$	0.0336	0.0243	1.0000

Eq. 5 indicates that the frontier electron density is a useful descriptor in determining the biological activity. The negative coefficient of the $\left(f_{O_6}^{E}\right)^2$ term indicates that variations in the activity are parabolically related to the electrophilic frontier electron density of 6-position oxygen atom. The optimum $f_{O_6}^{E}$ value according to Eq. 5 is about 11.15. The negative coefficient of the $\left(f_{C_{11}}^{N}\right)^2$ terms indicates that variations in the activity are also parabolically related to the nucleophilic frontier electron density of 11-position oxygen atom and the optimum value is about 0. In addition, Eq. 5 shows that $f_{C_{10}}^{N}$ term correlates with the biological activity positively and the greater the ability of the carbon atom to accept electrons, the higher the activity (Fig. 2).

A stepwise regression analysis for the development of Eq. 6 indicates that introducing the $f_{N_5}^{E}$ term improves the explained variance from 44% to 67%. Eq. 6 is statistically better than Eqs. 4 and 5 with larger correlation coefficient, small standard error and few descriptors. The observed versus predicted pI_{50} values according to Eq. 6 were plotted in Figure 3a. Eq. 6 shows that 5-position nitrogen atom, 10-position carbon atom and 11-position carbon atom are the very important sites, which maybe bind to the receptor.
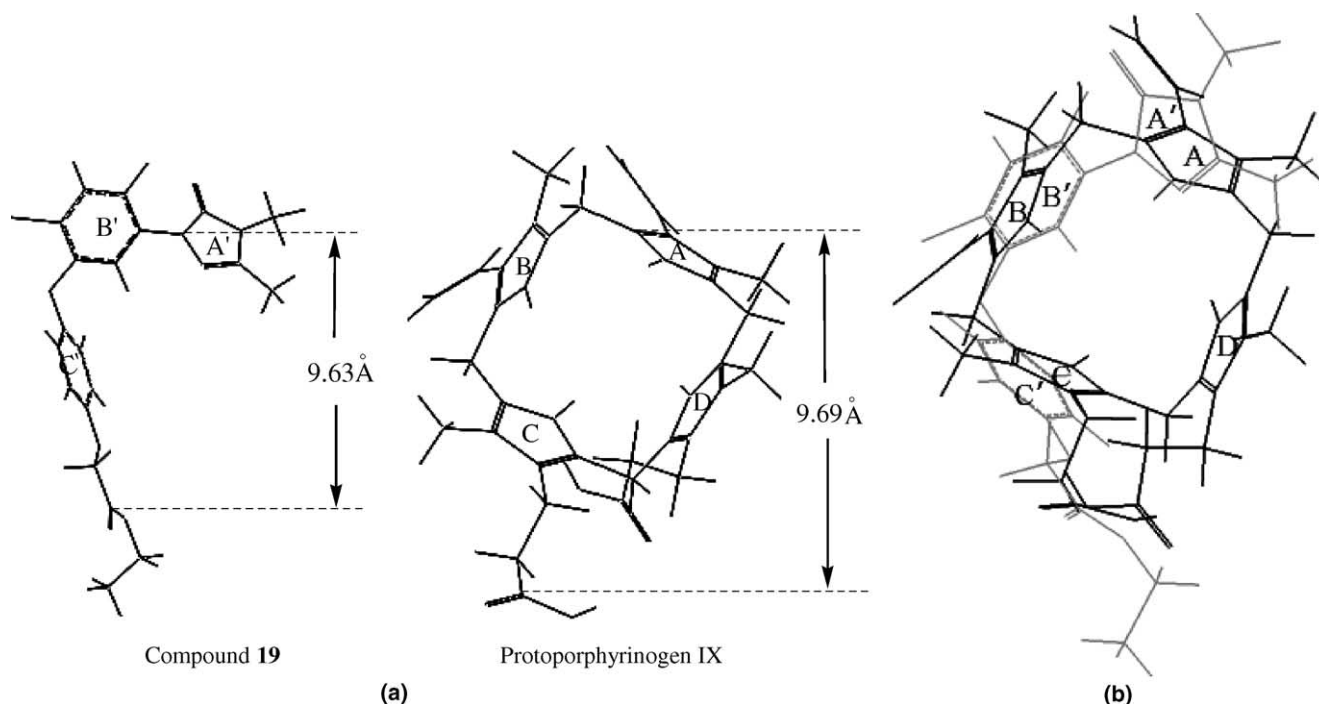
Figure 2. (a) Molecular size of compound **19** and protoporphyrinogen IX. (b) Matching of compound **19** with protoporphyrinogen IX.
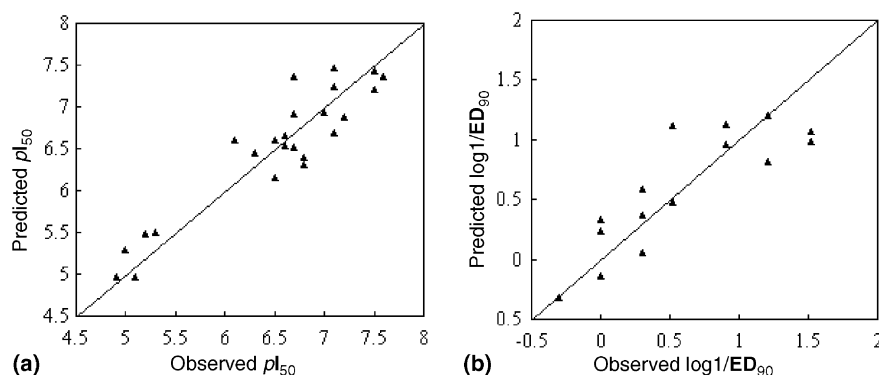

Figure 3. (a) Plot of observed (from Eq. 6) versus predicted pI_{50} and (b) plot of observed (from Eq. 9) versus observed log $1/\mathbf{ED}_{90}$ for phenyl triazolinones of interest.

Compound **19** was predicted to have a pI_{50} value of 6.4 according to Eq. 7 obtained by Theodoridis,[8] while we got the predicted pI_{50} value of 6.57, 6.83, and 6.71 according to Eqs. 4–6, respectively, whereas the observed value was found to be 9.0.

$$pI_{50} = 2.51 + 7.02B_1 - 0.18\pi^2 - 2.57B_1^2$$
$$n = 14, \quad r^2 = 0.78, \quad s = 0.216 \tag{7}$$

Compound **20** was not included in Eq. 7 for in vitro activity and Eq. 8 for in vivo activity obtained by Theodoridis.[8] We predicted the in vitro pI_{50} values of compound **20** to be 5.09, 5.39, and 4.38 according to Eqs. 4–6, respectively. For the greenhouse pre-emergence activity, we obtained Eq. 9 containing compound **20** with improved correlation compared to Eq. 8. The observed versus predicted log $1/\mathbf{ED}_{90}$ values according

to Eq. 9 were plotted in Figure 3b. According to Eq. 9, we can predict well the in vitro activity of compound **20**. Therefore, we deduced that the in vitro pI_{50} value (7.90) of compound **20** may be a wrong value.

$$\log 1/\mathbf{ED}_{90} = 0.66pI_{50} + 0.25\pi - 6.98$$
$$n = 14, \quad r^2 = 0.60, \quad s = 0.40 \tag{8}$$

$$\log 1/\mathbf{ED}_{90} = \underset{(0.0919)}{0.3237} + \underset{(0.5364)}{2.4748Q_{C_{11}}} + \underset{(0.0271)}{0.0867f_{N_5}^{E}} - \underset{(0.0193)}{0.0854f_{C_{10}}^{N}}$$
$$n = 15, \quad q^2 = 0.42, \quad s_{press} = 0.504, \quad r^2 = 0.71,$$
$$s = 0.356, \quad F = 8.95 \tag{9}$$

To compare the difference between the DFT and semi-empirical methods, the QSAR analysis of compounds

1–**26** was also carried out by using the same descriptor values calculated by using MOPAC program[25] (version 6.0) with PM3 parameterizations. Based on these descriptor values, we obtained the following correlation relationships, respectively:

$$pI_{50} = -2.9718 + 6.3141 Q_{C_{11}} + 1.1441 \Delta E - 0.8481 f^E_{O_6}$$
$$\quad\;\; (0.1373) \qquad\quad (1.7103) \qquad\quad (1.0329) \qquad\quad (6.7477)$$

$$n = 24, \quad r^2 = 0.41, \quad s = 0.672, \quad F = 4.72 \qquad (10)$$

$$pI_{50} = 30.9822 - 1.7486 f^N_{C_{10}} - 20.6866 \left(f^N_{C_{11}}\right)^2 - 3.1307 \Delta E$$
$$\quad\;\; (0.1336) \qquad\quad (2.5067) \qquad\qquad (35.1394) \qquad\qquad (1.2851)$$

$$\qquad + 108.1887 f^E_{O_6} - 928.3773 \left(f^E_{O_6}\right)^2$$
$$\qquad\qquad (28.3940) \qquad\qquad (227.4528)$$

$$n = 24, \quad r^2 = 0.50, \quad s = 0.654, \quad F = 3.61 \qquad (11)$$

$$pI_{50} = 6.1669 + 5.8871 Q_{C_{11}} + 1.0297 f^E_{N_5} + 1.5230 f^N_{C_{10}}$$
$$\quad\;\; (0.1383) \qquad\quad (1.5940) \qquad\quad (2.0813) \qquad\quad (2.3400)$$

$$n = 24, \quad r^2 = 0.41, \quad s = 0.677, \quad F = 4.56 \qquad (12)$$

A comparison of the DFT-based QSAR results (Eqs. 4–6) with the PM3-based QSAR results (Eqs. 10–12) clearly demonstrates that DFT-based quantum chemical descriptors led to the better correlation relationship than that the corresponding descriptors based on PM3 method. This shows that the QSAR analysis using DFT-based quantum chemical descriptors is more reliable than that using the corresponding descriptors based on the MOPAC-PM3 method.

Validation is a crucial aspect of any quantitative structure–activity relationship (QSAR) modeling. Most of the QSAR modeling methods implement the leave-one-out (or leave-some-out) cross-validation procedure.[26] The outcome from the cross-validation procedure is cross-validated q^2, which is used as a criterion of both robustness and predictive ability of the model. It has been shown that a cross-validated q^2 of 0.3 corresponds to a probability of chance correlation with activity of less than 0.05, hence a q^2 value of 0.3 or more is considered to be statistically significant.[27–29] Many authors consider higher q^2 (for instance, $q^2 > 0.5$)[30] as an indicator or even as the ultimate proof that the model is highly predictive. The models of Eqs. (4)–(6) have the following statistic results $q^2 = 0.68$, $s_{press} = 0.496$, $q^2 = 0.64$, $s_{press} = 0.556$, and $q^2 = 0.76$, $s_{press} = 0.434$, respectively, which suggests that these models are highly predictive and reliable for the prediction of in vitro biological activity. The model of Eq. 9 has the statistic results $q^2 = 0.42$, $s_{press} = 0.504$, which suggests the present QSAR model for in vivo biological activity is statistically significant.

3.2. Molecular shape similarity of compound 19 and Protogen

Through the comparison of the difference between observed and predicted values (Table 2) from QSAR equations 4–6, we can find that compounds **19** and **20** are outlier. The foregoing discussion indicates the pI_{50} value (7.9) of compound **20** may be a wrong value. So we specially compared the geometries of compound **19** and Protogen (Fig. 2) here. And the torsion angles of Protogen and compound **19** in the most stable geometries were listed in Table 3, respectively. The equilibrium geometry of Protogen was investigated at the same DFT level.

Table 3 listed some selected torsion angles in the most stable conformation of Protogen and compound **19**. From Table 3, we can conclude that the torsion angles of Protogen B–C rings and compound **19** B′–C′ rings are very close. Furthermore, we can find from Figure 2(a) that the distance between nitrogen atom at position 5 and carbon atom at position 40 of compound **19** is 9.63 Å, which is almost equal to that (9.69 Å) between carbon atom at position 5 and carbon atom at position 53 of Protogen. The superimposition model of compound **19** and Protogen is shown in Figure 2(b). These results suggest that compound **19** exhibited great similarity to the A, B, and C rings of Protogen, which explains reasonably why compound **19** has the highest activity.

4. Conclusion

It is the key to obtain QSARs equations with high quality, because the success of a QSARs model is highly dependent upon the choice of descriptors. So exploring reliable descriptors, especially electronic descriptors, and studying their applicability could lead to vital improvements in QSARs.

By considering all of the above facts, the present study was performed to examine the applicability of DFT-based quantum chemical descriptors in QSAR analysis for studying the biological activity of a series of Protox inhibitors. The DFT-based quantum chemical descriptors were obtained at the B3LYP/6-31G(d,p) level. It has been shown that the use of the DFT-based quantum chemical descriptors indeed led to a better QSAR equation than that obtained from the use of the corresponding descriptors evaluated at a semiempirical PM3 level.

The obtained QSAR results based on the DFT-based descriptors showed that the more the positive charges at the 11-position carbon atom of the phenyl ring, the higher the activity. The greater the ability of the 10-position carbon atom to accept electrons, the higher the activity. And the greater the ability of the 6-position oxygen atom and 5-position nitrogen atom to donate electrons, the higher the activity. Moreover, the negative coefficients of the $(f^N_{C_{11}})^2$, $(f^E_{O_6})^2$ terms indicated that variations in the activity were parabolically related to the nucleophilic frontier electron density of 11-position oxygen atom, the electrophilic frontier electron density of 6-position oxygen atom, respectively. The optimum $f^E_{O_6}$ value and optimum $f^N_{C_{11}}$ value in Eq. 5 were about 11.15 and 0, respectively. The results indicated the

triazo ring was an electron-donating site binding to the electropositive region of receptor and the phenyl ring was an electron-accepting site binding to the electronegative region of receptor. Furthermore, geometry analysis for compound **19** and Protogen suggested that the similarity between the inhibitors and a part of the natural substrate of the enzyme in terms of 3-D molecular shape was important for the molecular recognition at the active site of the Protox.

In conclusion, the excellent QSAR results for phenyl triazolinones were obtained using important quantum chemical descriptors based on DFT method. Moreover, the reason of the highest activity of compound **19** can also be explained rationally using geometry analysis and the greenhouse pre-emergence activity of compound **20** can also be predicted exactly using quantum chemical descriptors. The crossvalidation using the LOO method shows the QSARs model is reliable. Therefore, DFT-based QSARs could be expected to help facilitate the future design of additional substituted phenyl triazolinone derivatives of Protox inhibitors with good biological activity.

Acknowledgements

Supplementary data

The equilibrium geometries optimized at DFT-B3LYP/6-31G(d,p) level for all phenyl triazolinones and protoporphyrinogen IX listed in Table A. Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2004.08.046.

References and notes

1. Beale, S. I.; Weinstein, J. D. In *Biosynthesis of Heme and Chlorophylls*; Dailey, H. A., Ed.; McGraw-Hill: New York, 1990; pp 287–391.
2. Anderson, R. J.; Norris, A. E.; Hess, F. D. In *Porphyric pesticides: Chemistry, Toxicology and Pharmaceutical Applications*; Duke, S. O., Rebeiz, C. A., Eds.; ACS Symposium Series; American Chemical Society: Washington, DC, 1994; Vol. 559, pp 18–33.
3. Dayan, F. E.; Duke, S. O. In *Herbicide Activity: Toxicology, Biochemistry and Molecular Biology*; Roe, R. M., Burton, J. D., Kuhr, R. J., Eds.; I.O.S.: Amsterdam, The Netherlands, 1997, pp 11–35.
4. Gupta, I.; Tripathy, B. C. *Indian J. Biochem. Biophys.* **2000**, *37*, 498.
5. Theodoridis, G.; Baum, J. S.; Hotzman, F. W.; Manfredi, M. C.; Maravetz, L. L.; Lyga, J. W.; Tymonko, J. M.; Wilson, K. R.; Poss, K. M.; Wyle, M. J. In *Synthesis and Chemistry of Agrochemicals III*; Baker, D. R., Fenyes, J. G., Basarak, G. S., Eds.; ACS Symposium Series; Amercian Chemical Society: Washington, DC, 1992; Vol. 504, pp 134–146.
6. Theodoridis, G.; Poss, K. M.; Hotzman, F. W. In *Synthesis and Chemistry of Agrochemicals IV*; Baker, D. R., Fenyes, J. G., Basarak, G. S., Eds.; ACS Symposium Series; Amercian Chemical Society: Washington, DC, 1995; Vol. 584, pp 78–89.
7. Theodoridis, G.; Bahr, J. T.; Davidson, B. L.; Hart, S. E.; Hotzman, F. W.; Poss, K. M.; Tutt, S. F. In *Synthesis and Chemistry of Agrochemicals IV*; Baker, D. R., Fenyes, J. G., Basarak, G. S., Eds.; ACS Symposium Series; Amercian Chemical Society: Washington, DC, 1995; Vol. 584, pp 90–99.
8. Theodoridis, G. *Pestic. Sci.* **1997**, *50*, 283.
9. Reddy, K. N.; Dayan, F. E.; Duke, S. O. In *Comparative QSAR*; Devilliers, J., Ed.; Taylor and Francis, 1998, pp 197–233.
10. Par, R. G.; Yang, W.. In *Density-functional theory of atoms and molecules*; Oxford Univ. Press: Oxford, 1989.
11. Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.
12. Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
13. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
14. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
15. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
16. Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
17. Foresman, J. B.; Frisch, A. In *Exploring Chemistry with Electronic Structure Methods*; Gaussian Inc.: Pittsburgh, PA, 1996.
18. Sulpizi, M.; Folkers, G.; Rothlisberger, U.; Carloni, P.; Scapozza, L. *Quant. Struct.-Act. Relat.* **2002**, *21*, 173.
19. Bernardi, F.; Bottoni, A.; Garavelli, M. *Quant. Struct. Act. Relat.* **2002**, *21*, 128.
20. Arulmozhiraja, S.; Morita, M. *Chem. Res. Toxicol.* **2004**, *17*, 348.
21. Karelson, M.; Lobanov, V. S. *Chem. Rev.* **1996**, *96*, 1027.
22. Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. *J. Chem. Phys.* **1988**, *89*, 2193.
23. Breneman, C. M.; Wiberg, K. B. *J. Comp. Chem.* **1990**, *11*, 361.
24. Gaussian 03, Revision B.03, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T., Jr.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Pittsburgh, PA, 2003.
25. Stewart, J. J. P. MOPAC version 6.0.

26. Clark, R. D.; Sprous, D. G.; Leonard, J. M. In *Rational Approaches to Drug Design*; Höltje, H.-D., Sippl, W., Eds.; Prous Science, SA, 2001; pp 475–485.

27. Agarwal, A.; Pearson, P. P.; Taylor, E. W.; Li, H. B.; Dahlgren, T.; Herslof, M.; Yang, Y.; Lambert, G.; Nelson, D. L.; Regan, J. W.; Martin, A. R. *J. Med. Chem.* **1993**, *36*, 4006.

28. Clark, M.; Cramer, R. D. I.; Jonea, D. M.; Patterson, D. E.; Simeroth, P. E. *Tetrahedron Comput. Meth.* **1990**, *3*, 47.

29. Thomas, B. F.; ComDton, D. R.; Martin, B. R.; Seamus, S. F. *Mol. Pharmacol.* **1991**, *40*, 656.

30. Golbraikh, A.; Tropsha, A. *J. Mol. Graph. Model.* **2002**, *20*, 269.