



3D-QSAR and Docking Modeling Study of 1,3,5-Triazine Derivatives as PSII Electron Transport Inhibitor

CHUNLING ZHANG¹, SHAN CHANG^{1,*}, XUHONG TIAN¹ and YUANXIN TIAN²

¹College of Information, South China Agricultural University, Guangzhou, P.R. China

²School of Pharmaceutical Sciences, Southern Medical University, Guangzhou, P.R. China

*Corresponding authors: Fax: +86 20 85280320; Tel: +86 13763362095; E-mail: schang@scau.edu.cn; scauzcl@163.com

Received: 8 June 2013;

Accepted: 25 July 2013;

Published online: 26 December 2013;

AJC-14520

Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) were applied to a set of 38 1,3,5-triazines derivatives herbicide. The comparative molecular field analysis model yielded the leave-one-out (LOO) crossvalidated correlation coefficient (q^2) of 0.634 and a non-cross-validated correlation coefficient (r^2) of 0.954. Comparative molecular similarity indices analysis model gave $q^2 = 0.679$ and $r^2 = 0.924$. The results indicate that two models have good predictability. The binding model of 1,3,5-triazines with the D1 protein of PSII was built. The docking studies indicated the F atom on 6-fluoroalkyl could formed the H-bond with NH of the main chain and side chain of His215 on the D1 protein, which could explain fluoromethyl-substituted is very important for strong inhibition and have high activity. The models provide guidance for designing highly active compound of herbicide synthesis.

Keywords: CoMFA, CoMSIA, 1,3,5-triazine derivatives, Herbicide, Docking.

INTRODUCTION

Photosystem II (PS II) is a membrane protein complex, it captures sunlight and performs the light-induced oxidation of water to atmospheric oxygen¹. It plays an important role in photosynthesis and the growth of plant. A series of electron transfer for light-driven take place in D1 and D2 subunits of PSII, the primary electron donor P680 and the plastoquinone Q_A , forming $P680^+$. Q_A^- , the electron transferred to the final electron acceptor plastoquinone Q_B , after accepting electrons Q_B leaves the Q_B site of D1 protein and replaced by new plastoquinone PQ^2 . The Q_B site in the D1 protein is the final sites of the electron transfer in PSII and also is the target site of most herbicides³. Because of the herbicide displace plastoquinone Q_B from the Q_B site, the herbicide can not transfer electron then the photosynthetic electron transfer blocked and the plant die.

Triazine-type herbicides like atrazine, simetryne and aimazine, were introduced in the early 1950s and contributed strongly to the general practice of chemical weed control. Most triazine herbicides are of the structure 4-(alkylamino)-6-(alkylamino)-1,3,5-triazine, while at the 2-site there is a chlorine⁴. Although about 62 biological species resistant against triazines-herbicide, triazine-herbicides are still widely used and occupied 50 % of market⁵. In order to develop better herbicides, a series 2-benzylamino-4-methyl-6-fluoroalkyl

1,3,5-triazines were found to possess strong herbicidal activities and can control atrazine-resistant *Chenopodium album* effectively^{4,6,7}. Quantitative structure activity (QSAR) plays more and more role in development herbicide, Clark⁸ listed the 20 references on herbicide QSAR from 2005 to 2011, indicate the importance of QSAR to herbicide discovery. In this paper the 38 compounds of 3D-QSAR models are constructed by comparative molecular field analysis (CoMFA)⁹ and comparative molecular similarity indices analysis (CoMSIA)¹⁰, explained the relationship between molecule structure and herbicide activity. Combined with docking for study the interaction between herbicide and D1 protein the binding model of 1,3,5-triazine derivatives with the D1 protein was built. The models can provide some references for the design of new triazine herbicide and forecasting the inhibition activity of novel molecules.

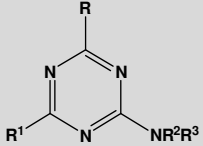
EXPERIMENTAL

A data set of 38 derivatives of 2-benzylamino-4-methyl-6-fluoroalkyl 1,3,5-triazine herbicide reported for this study^{4,6,7}. The given I_{50} is 50 % inhibition concentration for Photosynthetic Electron Transport (PET), all of I_{50} were converted to the logarithm scale using following formula. $pI_{50} = -\log I_{50}$.

The structures and Photosynthetic Electron Transport inhibitory activity of compound were listed in Table-1. The

data set was divided into a training set and a test set. Each set included a wide range of inhibitory activity. The training set which compounds marked 1-31 in Table-1 for generating 3D QSAR models and the test set of seven compounds which marked * in Table-1 randomly selected for validating the quality of the models.

TABLE-1
STRUCTURE AND PI_{50} OF 1,3,5-TRIAZINE
DERIVATIVES COMPOUNDS

					
Comp.	R	R ¹	R ²	R ³	PI_{50}
1	CF ₃	CH ₃	H	CH	3.80
2	CF ₃	CH ₃	H	C ₂ H ₅	4.69
3	CF ₃	CH ₃	C ₂ H ₅	C ₂ H ₅	3.26
4	CF ₃	CH ₃	H	C ₂ H ₁₁	6.37
5	CF ₃	CH ₃	H	C ₈ H ₁₇	6.79
6	CF ₃	CH ₃	H	C ₁₈ H ₃₇	4.27
7	CF ₃	CH ₃	H	CH ₂ C ₆ H ₅	6.85
8	CF ₃	C ₂ H ₅	H	CH ₂ C ₆ H ₅	5.64
9	CF ₃	i-C ₃ H ₇	H	CH ₂ C ₆ H ₅	4.21
10	CF ₃	CCl ₃	H	CH ₂ C ₆ H ₅	5.25
11	CHF ₂	CH ₃	H	CH ₂ C ₆ H ₅	5.69
12	C ₂ F ₅	CH ₃	H	CH ₂ C ₆ H ₅	6.46
13	n-C ₃ F ₇	CH ₃	H	CH ₂ C ₆ H ₅	5.29
14	CF ₃	CH ₃	H	CH ₂ (4-C ₆ H ₄ F)	6.91
15	CF ₃	CH ₃	H	CH ₂ (4-C ₆ H ₄ Cl)	6.98
16	CF ₃	CH ₃	H	CH ₂ (4-C ₆ H ₄ -tert-Bu)	6.75
17	CF ₃	CH ₃	CH ₃	CH ₂ C ₆ H ₅	4.01
18	CF ₃	CH ₃	H	(CH ₂) ₂ C ₆ H ₅	5.39
19	CF ₃	CH ₃	H	(CH ₂) ₃ C ₆ H ₅	6.75
20	CF ₃	CH ₃	H	CHCH ₃ (S)C ₆ H ₅	6.83
21	CF ₃	CH ₃	H	CHCH ₃ (R)C ₆ H ₅	4.55
22	CF ₃	CH ₃	H	CHCH ₃ (4-C ₆ H ₄ F)	6.32
23	CF ₃	CH ₃	H	CHCH ₃ (4-C ₆ H ₄ Cl)	6.69
24	CF ₃	CH ₃	H	CHCH ₃ (4-C ₆ H ₄ Bu-t)	6.66
25	CF ₃	CH ₃	H	CHC ₂ H ₅ (4-C ₆ H ₄ Cl)	6.79
26	CF ₃	CH ₃	H	CH ₂ (2-C ₆ H ₄ Cl)	5.95
27	CF ₃	CH ₃	H	CH ₂ (3-C ₆ H ₄ Cl)	7.21
28	CF ₃	CH ₃	H	CH ₂ (4-C ₆ H ₄ Br)	6.94
29	CF ₃	CH ₃	H	CH ₂ (4-C ₆ H ₄ CF ₃)	7.29
30	CF ₃	CH ₃	H	CH ₂ (4-C ₆ H ₄ CH ₃)	6.74
31	CF ₃	CH ₃	H	CH ₂ (4-C ₆ H ₄ OCH ₃)	6.58
32*	CH ₃	CH ₃	H	CH ₂ (4-C ₆ H ₄ Cl)	5.78
33*	CH ₂ F	CH ₃	H	CH ₂ (4-C ₆ H ₄ Cl)	6.15
34*	Cl	CH ₃	H	CH ₂ C ₆ H ₅	5.33
35*	OCH ₃	CH ₃	H	CH ₂ (4-C ₆ H ₄ Cl)	6.09
36*	C ₂ H ₅	CH ₃	H	CH ₂ (4-C ₆ H ₄ Cl)	6.45
37*	C ₃ H ₇ -i	CH ₃	H	CH ₂ C ₆ H ₅	5.65
38*	C ₄ H ₉ -t	CH ₃	H	CH ₂ C ₆ H ₅	5.99

Molecular modeling and alignment: All compound structure was sketched using sketch module of SYBYL7.3 program package¹¹. Due the highest activity, the compound 29 was selected as the template the other compound was modified based on it. Partial atomic charges were added to each atom using Gasteiger-Huckel charge¹². All compound was used the TRIPOS force field¹³ and Powell optimization method with distance dependent dielectric and conjugate gradient

method with convergence criterion of 0.05 kcal/mol to get the energy minimizations conformation.

The geometric similarity usually presents between the modeled structure and bioactive conformation for 3D QSAR. Alignment of compounds is the most important factor in obtaining a reliable model. Since the 29 compound is the most activity compound thus it used as template structure for the alignment. The reference atoms of compound 29 is 1,3,5-triazines. The molecular alignment was carried out with database Align of SYBYL 7.3. The aligned molecule of training set was shown in Fig. 1.

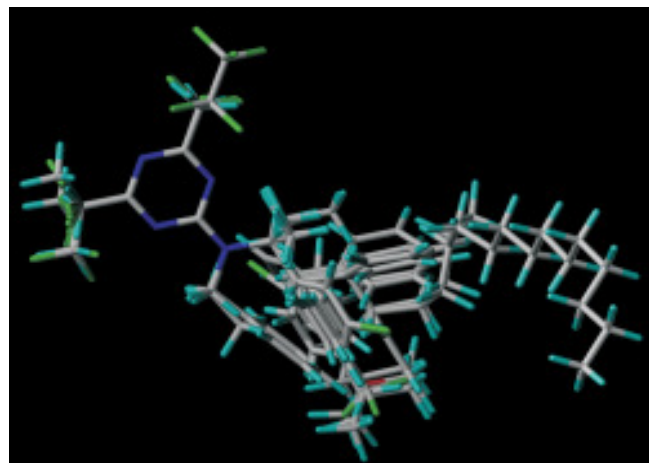


Fig. 1. 3D-view of aligned molecules (training set) based on Sybyl align database method

CoMFA and CoMSIA calculations: The steric and electrostatic CoMFA potential fields^{14,15} were calculated at each lattice interaction of a regularly spaced grid of 2.0 Å. The grid box dimensions were determined automatically in such the way which the regions' boundaries were extended beyond 4 Å in each direction from the co-ordinates of each molecule. The steric and electrostatic contributions were truncated to +30 kcal/mol. A distance dependent dielectric constant of 1 was used. The CoMFA Field Class used Tripos Standard. A *sp*³ hybridized carbon atom with +1 charge served as probe atom to calculate steric and electrostatic fields.

In addition to calculate steric and electrostatic, hydrophobic, hydrogen-bond donor and hydrogen-bond acceptor field were calculated as well. A probe with a charge of +1, hydrophobicity +1.0 and H-bond donor and acceptor properties of +1.0 was used to calculate five physicochemical descriptors. The lattice interactions of a regularly spaced grid for evaluated five physicochemical descriptors was 2.0 Å. The lattice dimensions were set 4Å as well as CoMFA. For attenuation factor a controlling the steepness of Gaussian function the standard value of 0.3 was accepted.

Molecular docking: In order to check the main factors affecting the activity from the 3D-QSAR models and study binding conformation, the highest activity compound 29 was docked into the Q_B site of D1 protein by Surflex-Dock program from Sybyl7.3. Most of the complexes structure of PSII were isolated from *Synechococcus elongates*^{2,16,17} and *Thermosynechococcus elongates*¹, the resolution from 3.8 Å to 1.9 Å. Although the D1 protein crystal structure of higher plants

has not been reported, the sequences of *Thermosynechococcus elongates* (PDB entry code: 1S5L) are similar to higher plant *Pisum sativum*, thus it can be used to mimic the higher plant in the docking study¹⁸. Before beginning docking all the water and ligands were removed and hydrogen atoms were added to the crystal structure randomly, atomic partial charges were added to the protein with Kollman-Unit charge. The protocol of Surflex-Dock can be established by ligand, residues and automatic etc. Because the ligand was not available in crystal 1S5L thus in this paper the protocol was established by residues. The Q_B binding site composed of the residues including Met214, Leu218, Ala251, Phe255 and Leu271¹⁶, thus these residues were guiding residues. Others parameters were set as default values.

RESULTS AND DISCUSSION

Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) analysis: A training set with 31 molecules was used to obtain both the conventional CoMFA and CoMSIA models. A test set with 7 molecules was used to judge the accuracy of the models. The parameters of statistical for the CoMFA and CoMSIA model were shown in Table-2. The predicted pl_{50} values and residual values of CoMFA and CoMSIA model were shown in Table-3.

Comparative molecular field analysis models were obtained by steric and electrostatic. There are 4 optimum numbers of components in PLS analysis. The correlation of a cross-validated q^2 is 0.634. The non-cross-validated PLS analysis was run again with the previously optimum number which get by the cross-validated analysis, to give r^2 of 0.954 with a standard error of estimate (SEE) of 0.265.

TABLE-2
SUMMARY OF COMFA AND COMSIA RESULTS

QSAR parameters	CoMFA	CoMSIA
$q^2_{L_{oo}}$	0.634	0.679
r^2	0.954	0.924
N	4	4
SEE	0.265	0.341
F-value	134.655	78.683
Fraction of field contribution		
Steric	0.644	0.180
Electrostatic	0.356	0.329
Hydrophobic	–	0.147
Acceptor	–	3.41e-005
Donor	–	0.344

$q^2_{L_{oo}}$, Cross-validated correlation with leave-one-out; N, optimum number of components as determined by L_{oo} study; SEE, standard error of estimation; r^2 , predictive correlation coefficient.

Comparative molecular similarity indices analysis models were obtained by using five fields of steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor. The correlation of a cross-validated q^2 is 0.679 and give 4 optimum numbers of components. The non-cross-validated PLS analysis was run again to give r^2 of 0.924 with a standard error of estimate (SEE) of 0.341.

Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA)

TABLE-3
EXPERIMENTAL AND PREDICTED COMPOUND pl_{50} AND RESIDUALS OBTAINED BY THE COMFA AND COMSIA MODELS FOR TRAINING AND TEST SET COMPOUNDS

Comp.	pl_{50}	CoMFA		CoMSIA	
		Predicted	Residual	Predicted	Residual
1	3.80	3.930	-0.1302	4.006	-0.2056
2	4.69	4.404	0.2858	4.372	0.3185
3	3.26	3.209	0.0512	3.322	-0.0616
4	6.37	6.241	0.1293	6.115	0.2549
5	6.79	6.958	-0.1681	7.046	-0.2562
6	4.27	4.402	-0.1317	4.229	0.0413
7	6.85	6.817	0.0328	6.921	-0.0710
8	5.64	5.686	-0.0460	5.526	0.1141
9	4.21	4.034	0.1762	4.241	-0.0308
10	5.25	5.563	-0.3135	5.271	-0.0208
11	5.69	6.259	-0.5688	5.870	-0.1797
12	6.46	6.307	0.1530	6.287	0.1729
13	5.29	5.500	-0.2099	5.551	-0.2613
14	6.91	6.772	0.1382	6.817	0.0927
15	6.98	6.848	0.1323	6.948	0.0321
16	6.75	7.004	-0.2545	6.832	-0.0823
17	4.01	4.168	-0.1579	4.038	-0.0284
18	5.39	5.308	0.0824	5.402	-0.0117
19	6.75	6.662	0.0883	6.637	0.1126
20	6.83	6.468	0.3618	6.748	0.0822
21	4.55	4.414	0.1361	4.422	0.1281
22	6.32	6.543	-0.2232	6.475	-0.1546
23	6.69	6.740	-0.0504	6.671	0.0193
24	6.66	6.604	0.0557	6.810	-0.1501
25	6.79	6.694	0.0963	6.706	0.0840
26	5.95	5.998	-0.0482	6.225	-0.2751
27	7.21	6.665	0.5453	6.930	0.2796
28	6.94	6.921	0.0188	7.062	-0.1218
29	7.29	7.099	0.1915	7.220	0.0702
30	6.74	6.972	-0.2320	6.737	0.0026
31	6.58	6.721	-0.1407	6.474	0.1058
32*	5.78	5.879	-0.0987	6.164	-0.3841
33*	6.15	5.738	0.4120	5.994	0.1561
34*	5.33	5.633	-0.3025	5.569	-0.2390
35*	6.09	5.884	0.2057	6.021	0.0685
36*	6.45	5.820	0.6302	6.069	0.3811
37*	5.65	5.897	-0.2467	5.923	-0.2728
38*	5.99	5.819	0.1706	5.781	0.2087

contours maps: The activity of impact which impacted by the various fields of molecular compound can be seen in CoMFA and CoMSIA contours maps. The molecule in center is compound 22. The contour maps of CoMFA are shown in Fig. 2. The electrostatic fields (A) is characterized by blue and red contour. Blue contours refer to regions where negatively charged substituents are disfavored; red contours indicate regions where negatively charged substituents are favored. The steric field is characterized by green and yellow contour which shown in Fig. 2(B). The green contours indicated sterically favored regions where increase steric will enhance activity and yellow contours indicated sterically disfavored regions where increase steric will increase activity.

In Fig. 2.(A) the blue contour map around the benzene ring of R^3 indicates that positive charge or electron deficient substitute will enhance the activity at this position. For example on the compound 22 and the compound 23 C(4) of benzene ring of R^3 is F and Cl respectively, the electronegativity of F is larger than Cl thus the activity of compound 22 is lower than

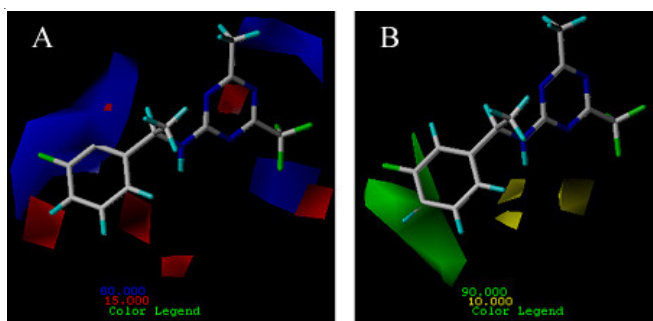


Fig. 2. Contour maps from the final CoMFA analysis in combination with all analogues of the training set. (A) Electrostatic fields: Blue contours (80 % contribution) refer to regions where negatively charged substituents are disfavored; red contours (15 % contribution) indicate regions where negatively charged substituents are favored. (B) Steric fields: Green contours (90 % contribution) refer to sterically favored regions; yellow contours (10 % contribution) indicate disfavored areas. The molecular in center is compound 22

compound 23. In Fig. 2.(B) the big green contours around the benzene ring of R^3 means steric bulks will increase activity in these area, such as compounds 24, 29 and 31 with larger steric bulks than compound 22 thus the activity of compounds 24, 29 and 31 higher than compound 22. The CoMSIA contour maps were shown in Fig. 3, compound 22 was selected as a reference molecule. The CoMSIA electrostatic and steric contours plots shown in Fig. 3.(A) and (B) are similar with those of CoMFA. Fig. 3.(C) and (D) describe the hydrophobic and donor contour map of CoMSIA model. In Fig. 3.(C) the yellow contour for hydrophobic group increase activity, white indicates the disfavor region; in Fig. 3.(D) the cyan contour for H-bond donor favor region, purple indicates the disfavor region. In Fig. 3.(C) there are white contours on R^2 and R^3 which mean more hydrophobic group here lower activity of compound. For example in the R^3 of compound 5 there are 8 carbon atoms made of carbon chain, there are 18 carbon atoms in the same region of compound 6, the activity of compound 5 is higher than compound 6. Compound 17 with CH_3 in R^2 and compound 7 with H in R^2 thus the activity compound 7 higher than compound 17.

Docking analysis: The cavity of Q_B site composed Met214, Leu218, Ala251, Phe255 and Leu271 residues, Q_B forms hydrogen bond with residue His215 and shares another hydrogen bond with Phe265 and Ser246¹⁶. The interacting model of compound 29 with D1 protein and the main residues involved in the interaction were generally depicted in Fig. 4. Compound 29, residue His215 and residue Ala263 were shown as stick, the H-bond was shown as red dot line and the residues of hydrophobic pocket were labeled. The hydrophobic pocket includes Met214, Leu271, Phe211, Phe265 and Ala263, it is similarly with the work of Zouni *et al.*¹⁷. The triazine part of compound is matched to the hydrophilic area. The 6-fluoromethyl substituent of compound 29 is CF_3 , in this model one F atom of CF_3 formed an H-bond with the main chain NH in residue His215, at the same position NH of His215 also formed an H-bond with the N of triazine, the other F atom of CF_3 formed an H-bond with the side chain NH in residue His215, which could explain 6-fluoromethyl-substituted is very important for strong inhibition and have high activity.

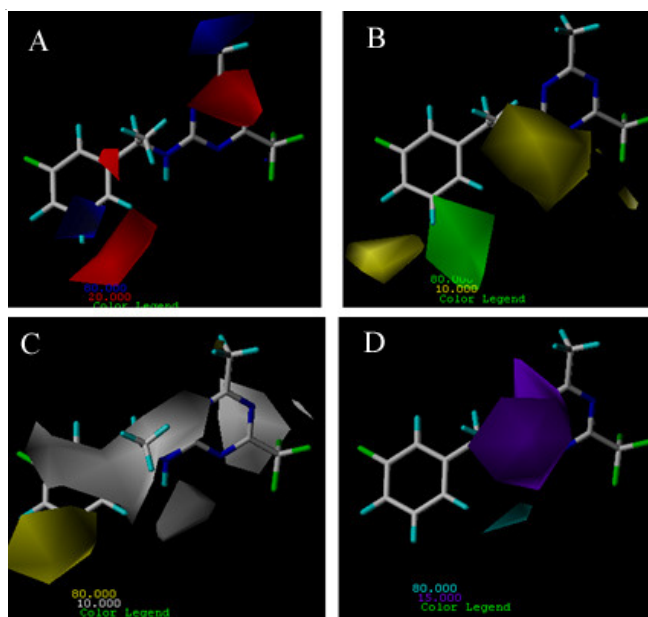


Fig. 3. CoMSIA contour maps illustrating steric, electrostatic, hydrophobic and donor features in combination with compound 22. (A) The green contour (80 % contribution) for steric favored region and yellow contour (10% contribution) calls for a reduction of this potential to improve affinity. (B) Blue contour (80 % contribution) for positive charge preferred region to improve binding affinity, red contours (20 % contribution) indicate regions where negatively charged substituents are favored. (C) The yellow contour (80 % contribution) for hydrophobic group increase activity, white contour (10 % contribution) indicates the disfavor region. (D) The cyan contour (80 % contribution) for H-bond donor favor region, purple contour (15 % contribution) indicates the disfavor region

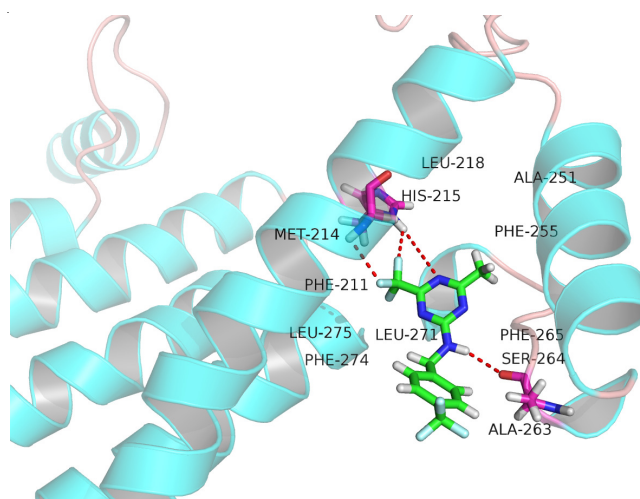


Fig.4. The binding mode between compound 29 and the D1 protein. Hydrogen bonds were labeled by red dot line, the residues of hydrophobic pocket were labeled. Compound 29, residue His215 and residue Ala263 were shown as stick

Most of biology resistant to triazine-herbicide is mutation take place in the Ser246 residue of D1 protein³. In this model R^2 formed H-bond with O of Ala263 instead of Ser264, it inferred the 2-benzylamino-4-methyl-6-fluoroalkyl 1,3,5-triazines can control atrazine-resistant effectively.

Conclusion

The 3D-QSAR analyses of 38 1, 3, 5-triazine derivatives as herbicides of Photosystem II was build up using by CoMFA and CoMSIA methods, obtaining CoMFA ($q^2 = 0.634$, $r^2 = 0.954$, $SEE = 0.265$) and CoMSIA ($q^2 = 0.679$, $r^2 = 0.924$, $SEE = 0.341$). In CoMFA the contribution of steric fields are larger than electrostatic fields. In CoMSIA the most contribution are electrostatic and H-donor. The Contour map of CoMFA and CoMSIA shown steric bulks around the benzene ring of R³ subsistent will increase activity of herbicide, positive charge or electron deficient substitute around the benzene ring of R³ subsistent will enhance the activity. Thus the 2-benzylamino substitution may be important for the herbicide activity. The docking mode indicates the 6-fluoromethyl substitution is important for inhibition.

ACKNOWLEDGEMENTS

This work were supported by the National Natural Science Foundation of China (No. 31200990) and National Science Foundation of China (No. 81001423).

REFERENCES

1. Y. Umena, K. Kawakami, J.R. Shen and N. Kamiya, *Nature*, **473**, 55 (2011).
2. B. Loll, J. Kern, W. Saenger, A. Zouni and J. Biesiadka, *Nature*, **438**, 1040 (2005).
3. W. Oettmeier, *Cell. Mol. Life Sci.*, **55**, 1255 (1999).
4. H. Kohno, A. Ohki, S. Ohki, K. Koizumi, M.E. Van den Noort, G.C. Rodrigues, J.J.S. Van Rensen and K. Wakabayashi, *Photosynth. Res.*, **65**, 115 (2000).
5. I. HEAP, International Survey of Herbicide Resistant Weeds (2002).
6. K. Koizumi, N. Kuboyama, K. Tomono, A. Tanaka, A. Ohki, H. Kohno, K. Wakabayashi and P. Böger, *Pestic. Sci.*, **55**, 642 (1999).
7. R. Okano, A. Ohki, S. Ohki, H. Kohno, J.J.S. van Rensen, P. Boger and K. Wakabayashi, *J. Biosci.*, **57**, 1009 (2002).
8. R.D. Clark, *Pest Manage. Sci.*, **68**, 513 (2012).
9. R.D. Cramer, D.E. Patterson and J.D. Bunce, *J. Am. Chem. Soc.*, **110**, 5959 (1988).
10. M. Bohm, J. Strzebecher and G. Klebe, *J. Med. Chem.* **42**, 458 (1999).
11. Sybyl, Molecular Modeling Package, version 7.3, St Louis (MO): Tripos Associates (2000).
12. J. Gasteiger and M. Marsili, *Tetrahedron*, **36**, 3219 (1980).
13. M. Clark, R.D. Cramer and N. Van Opdenbosch, *Comput. Chem.*, **10**, 982 (1989).
14. R.D. Cramer III, J.D. Bunce, D.E. Patterson and I.E. Frank, *Quant. Struct. Act. Relat.*, **7**, 18 (1988).
15. A.K. Ghose and G.M. Crippen, *J. Chem. Inform. Computer Sci.*, **27**, 21 (1987).
16. K.N. Ferreira, T.M. Iverson, K. Maghlaoui, J. Barber and S. Iwata, *Science*, **303**, 1831 (2004).
17. A. Zouni, H.T. Witt, J. Kern, P. Fromme, N. Krauss, W. Saenger and P. Orth, *Nature*, **409**, 739 (2001).
18. X.F. Han, Y.X. Liu, Y. Liu, L.H. Lai, R.Q. Huang and Q.M. Wang, *Chinese J. Chem.*, **25**, 1135 (2007).