3D-QSAR ON 1-SUBSTITUTED PHENOXYACETOXYALKYLPHOSPHONATES AND PHOSPHINATES USING COMFA AND COMSIA

Junbo He, Guoqiang Dai, Hao Peng, and Hongwu He
Key Laboratory of Pesticide and Chemical Biology, Ministry of Education and College of Chemistry, Central China Normal University, Wuhan, China

GRAPHICAL ABSTRACT

Abstract A three-dimensional quantitative structure–activity relationship (3D-QSAR) study was performed on two different chemical series, which have great herbicidal activity, employing comparative molecular field analysis (CoMFA) and comparative molecular similarity indices (CoMSIA) techniques to further investigate the structural requirement for the dicotyledonous stem inhibition. The optimal CoMFA and CoMSIA models obtained for the training set were all statistically significant with cross-validated coefficients ($q^2$) of 0.702, 0.726 and conventional coefficients ($R^2$) of 0.980, 0.963, respectively. These models were validated by a set of six molecules that were not included in the training set. The CoMFA yielded comparable models for phosphonate and phosphinate derivatives highlighting the significance of steric and electrostatic fields toward dicotyledonous stem inhibitory activity. The CoMSIA models indicated the importance of electrostatic and hydrophobic fields toward dicotyledonous stem inhibitory activity. The CoMSIA steric and electrostatic field maps are in accordance with field distribution of CoMFA maps and consistent with structure–activity relationships.

Keywords Phosphonate; phosphinates; CoMFA; CoMSIA; pyruvate dehydrogenase

INTRODUCTION

The pyruvate dehydrogenase complex (PDHe) is already known to be an important target of herbicides, and has received much more attention because it plays a key role in
catalyzing the oxidative decarboxylation of pyruvate with the production of acetyl coenzyme A (acetyl-CoA), NADH, and CO₂. The PDHc is composed of multiple copies of three enzymes: pyruvate decarboxylase (E1 subunit), dihydrolipoyl acetyl transferase (E2 subunit), a molecular protein which contains a transacetylase and so-called outer and inner lipoyl domains, E2c, E2L1, and E2L2, respectively), and dihydrolipoyl dehydrogenase (E3 subunit). Pyruvate dehydrogenase E1 component (PDHc E1, E.C. 1.2.4.1), which is the primary member of PDHc, catalyzes the first step of the multistep process, using thiamine diphosphate (ThDP) and Mg²⁺ as cofactors. Therefore, the design of inhibitors of PDHc E1 is favored as herbicides. Kluger et al. applied a promising method for experimental conformation of the involvement of specific covalent intermediates, which had been developed by Westerik and Wolfenden and by Thompson to demonstrate the existence of reactive intermediate of an enzyme-bound, covalent adduct of TPP, 2-lactyl TPP in the reaction catalyzed by the pyruvate decarboxylase (E1) enzyme of PDHc. Based on the view of biochemical design of agrochemicals, a series of acylphosphinates and acylphosphonates have been prepared as potential mechanism-based inhibitors of the enzyme by Baillie et al. These compounds were insufficient in themselves to be develop into commercial herbicides, however.

Furthermore, we designed and synthesized a series of 1-(substituted phenoxyacetoxy)-alkylphosphonate and alkylphosphinate derivatives based on PDHc, and some of them showed notable herbicidal activities. For the purpose of obtaining a predictive three-dimensional quantitative structure–activity relationship (3D-QSAR) model, we applied a comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) to investigate the various interaction fields on the bioactivities.

RESULTS AND DISCUSSION

CoMFA and CoMSIA Analysis

The title compounds (1–49 in Table S1) studied in this work were designed and synthesized to cover their potential herbicidal activity as widely as possible. The results of CoMFA and CoMSIA analysis are summarized in Table 1. The CoMFA PLS analysis yielded a high cross-validated correlation coefficient $q^2$ of 0.702 with an optimum number of nine components. This analysis was used for final non-cross-validated run, giving a correlation coefficient $R^2$ of 0.980 with a SEE of 0.090 showing a good linear correlation between the observed and predicted activities of the molecules in the training set. These statistical indexes were reasonably high, indicating that the CoMFA model might have a credible predictive ability. The steric field descriptors explain 0.493 of the variance, whereas the electrostatic field descriptors explain 0.507. Therefore, the electrostatic field had a slightly greater influence than the steric field.

The CoMSIA analysis using steric and electrostatic fields as descriptors gave a model with $q^2$ of 0.666 and $R^2$ of 0.931. The steric and electrostatic field descriptors explain variance of 0.400 and 0.600, respectively. Introducing a hydrophobic field into the CoMSIA analysis using steric and electrostatic fields resulted in a great increase in the $q^2$ and $R^2$ values ($q^2 = 0.726, R^2 = 0.963$), the steric, electrostatic, and hydrophobic fields descriptors explain variances of 0.191, 0.259, and 0.550, which indicates that the hydrophobic property of the title compounds exhibits a significant effect on the biological activity. However, further introduction of the hydrogen bond acceptor field into the CoMSIA
Table 1 Summary of Results from the CoMFA and CoMSIA Analysis

|               | CoMFA | CoMSIA (S,E) | (S,E,H) | (S,E,H,\(A\)) \(a\) |
|---------------|-------|--------------|---------|-----------------------|
| \(q^2\)      | 0.702 | 0.666        | 0.726   | 0.754                 |
| N             | 9     | 7            | 10      | 10                    |
| SEP           | 0.348 | 0.358        | 0.339   | 0.321                 |
| \(R^2\)      | 0.980 | 0.931        | 0.963   | 0.964                 |
| SEE           | 0.090 | 0.163        | 0.125   | 0.122                 |
| F value       | 182.181 | 67.316      | 82.840  | 86.296                |
| Fraction      |       |              |         |                       |
| Steric        | 0.493 | 0.400        | 0.191   | 0.126                 |
| Electrostatic | 0.507 | 0.600        | 0.259   | 0.237                 |
| Hydrophobic   | /     | /            | 0.550   | 0.422                 |
| Acceptor      | /     | /            | /       | 0.215                 |
| \(r^2(\text{bs})\) \(b\) | 0.990 | 0.958        | 0.978   | 0.978                 |
| SD            | 0.005 | 0.01         | 0.012   | 0.011                 |

\(a\) S, E, H, and A represent the steric, electrostatic, hydrophobic, and acceptor property fields, respectively.

\(b\) Results from 100 runs of bootstrapped analyses.

did not improve the statistical significance as expressed by similar \(q^2\) and \(R^2\) values. From the CoMSIA results, it may be concluded that electrostatic and hydrophobic interactions are the major factors to explain the field properties, which is the same as that described in previous work.\(^\text{16}\) To further assess the statistical ability and the robustness of the models, bootstrapping analysis (100 runs) was performed and \(r^2(\text{bs})\) of 0.990 (SD = 0.005), 0.978 (SD = 0.012) were obtained for CoMFA and CoMSIA, respectively.

The real test for the model predictability is to predict the activity of compounds which were not used in the generation of the model. Six compounds, which were not included in the training set, were selected as a test set to validate the QSAR models. The plots of the predicted versus the actual activity values for the training set and test set are shown in Figure 1. Both the CoMFA and CoMSIA models showed a good predictability on these compounds.

### CoMFA and CoMSIA Contour Maps

One of the most pleasant advantages of CoMFA and CoMSIA is that the field effect on the target property can be mapped as 3D coefficient contour plots. As shown in Figures 2–4, the coefficient contour maps provided a distinguished observation of important regions where any changes in the steric, electrostatic, and hydrophobic fields may affect the biological activity. In Figures 2–4, the contour diagrams of the field contributions (“\text{stdev*coeff}”) of different properties obtained from the CoMFA and CoMSIA analyses (using the steric, electrostatic, and hydrophobic fields) are illustrated together with exemplary ligands.

The steric contribution contour maps of CoMFA and CoMSIA are plotted in Figure 2. Green polyhedra represent a steric group that confers an increased herbicidal activity whereas yellow polyhedra represent a bulky group that results in a decreased herbicidal activity. The CoMFA sterically favorable green contours indicated obviously that bulky groups around the 2- and 4-positions of the phenyl ring of compound 40 are favorable for higher activity. However, in the CoMSIA contour, the yellow region is also located around
Figure 1  Predicted and actual activity of training set and test set using (a) CoMFA and (b) CoMSIA models.
the 4-position of the phenyl ring, suggesting that the volume of the substituent should be appropriated. This can be seen for compounds 19 and 20, in which replacing 4-chloro with 4-nitrogroup results in a decrease of activity as compared with compounds 5 and 8. From the contour of CoMFA, there is a region of yellow contour near the R^4-substituent of the training set; the occupation of this area by a bulky group will have a negative effect on the herbicidal activity as represented by compounds 12 and 16, which have lower activity than compounds 8 and 5, respectively.

The CoMSIA steric contour indicated obviously that a big space of yellow contour is located between the phosphoryl part and phenyl ring, suggesting that the steric block will decrease the herbicidal activity. Therefore, a small substituent at the phosphoryl part is favorable and will increase the activity, for example, compounds 35–40 possess higher activity. For the region of the R^3 substituent, from a large sterically favorable region and a small sterically disfavored region surrounding the molecule represented by both analyses,
Figure 3  Electrostatic contour maps from the CoMFA and CoMSIA models. (A) Electrostatic field distribution of CoMFA model, (B) electrostatic field distribution of the CoMSIA model using the steric, electrostatic, and hydrophobic field. Blue contours (80% contribution) encompass regions where an increase of positive charge will enhance activity, whereas in red contours areas (20% contribution) more negative charges are favorable for activity. Compound 40 is shown inside the field.

it can be concluded that the size of the binding site is limited and an optimal value for the steric effect is needed.

The electrostatic contour maps of CoMFA and CoMSIA were found to be plentiful and consistent with each other (Figure 3), just the CoMSIA approach provides more contiguous contour diagrams. A red contour indicates that an electronegative group will favor the activity while a blue contour will reduce the activity. In CoMFA, a great red block is found near the 2- and 4-position of the phenyl ring whereas it is located mainly at the 4-position in CoMSIA, which indicates that electronegative groups at the 4-position of the phenyl are essential for higher activity. As indicated in Figures 3A and 3B, great red regions above the same side of the oxygen atom of the carbonyl and the oxygen atom of phosphoryl group suggested that negative charge can increase the activity, and that electrostatic interaction plays an important role in the activity. Therefore, several blue regions at the oxygen atoms
of the phosphoryl group, as shown for compounds 40–49, indicate negative charge at OCH$_3$ and positive charge at CH$_3$; all of these compounds show very great activity. Therefore, it could be reasonably presumed that there is a significant electrostatic interaction between the phenyl ring and the phosphoric position with the receptor. As shown in Figures 3A and 3B, large regions of red contour located in the vicinity of the R$^3$-substituent indicated that electronegative groups at this position could help to increase the activity. This is reflected in certain compounds, such as 7, 15, 17, 18, and 46, which possess electronegative substitutes on the aromatic ring and have high activity.

The hydrophobic contour is obtained from the CoMSIA steric, electrostatic, and hydrophobic fields. The yellow contour suggests that a hydrophobic group favors the herbicidal activity whereas a white contour reduces activity. As displayed in Figure 4, a yellow area near the 4-position of the phenyl ring indicates the introduction of hydrophobic groups, which will increase the activity. This is just in the space where the red region appears in the electrostatic contour of the CoMFA, and it could be reasonable to suggest that the hydrophobic groups must be electronegative groups too. A very distinct hydrophilic site is the phosphoric position, where the big white block covers indicate that hydrophilic groups would increase the activity. For example, in compounds 35–40, the OH group increases the hydrophilic property and these compounds show high activity. From the white contour around the R$^3$ substituent, it can be seen that the hydrophilic effect makes some impact on the activity. For example, in compounds 9–12 and 47–49, the increase of hydrophobic effect results in a lower activity. Therefore, when the steric effect and electrostatic effect of the substituent of the phenyl ring is appropriate, increasing the hydrophilic property of R$^3$ substituent will increase the herbicidal activity.

CONCLUSIONS

In conclusion, CoMFA and CoMSIA were performed on 49 phosphonate and phosphinate derivatives which inhibited the dicotyledonous stem. Both models showed good prediction capabilities in terms of $q^2$ and $R^2$ values, while CoMFA model showed better predictive ability [SEE (standard error of estimate) = 0.090] than CoMSIA model.
The good correlation between experimental and predicted bioactivities for six compounds in the test set further verified the reliability of the constructed models. The present work indicates that the combined analysis of the results of the CoMFA and CoMSIA analyses enable us to obtain more comprehensive information about the structure-activity correlation. The CoMFA yielded comparable models for phosphonate and phosphinate derivatives highlighting the significance of steric and electrostatic fields towards dicotyledonous stem inhibitory activity. CoMSIA models indicated the importance of electrostatic, hydrophobic, and acceptor fields for the inhibitory activity toward dicotyledonous stem. The CoMSIA steric and electrostatic field maps are in accordance with field distribution of CoMFA maps and consistent with structure-activity relationships.

**FUNDING**

This work was supported by the National Key Technologies R & D Program of China (No.2011BAE06B03).

**SUPPLEMENTAL MATERIAL**

Supplemental data for this article can be accessed on the publisher’s website at http://dx.doi.org/10.1080/10426507.2014.931396

**REFERENCES**

1. Eto, M. *Biosci. Biotech. Biochem.* **1997**, 61, 1-11.
2. Linn, T. C.; Pelley, J. W.; Pettit, F. H.; Hucho, F.; Randall, D. D.; Reed, L. *J. Arch. Biochem. Biophys.* **1972**, 148, 327-342.
3. Linn, T. C.; Pettit, F. H.; Reed, L. *J. Proc. Natl. Acad. Sci. USA.* **1969**, 62, 234-341.
4. Dobritsch, D.; König, S.; Schneider, G.; Lu, G. *J. Biol. Chem.* **1998**, 273, 20196-20204.
5. Alvarez, F. J.; Ermer, J.; Hübner, G.; Schellenberger, A.; Schowen, R. L. *J. Am. Chem. Soc.* **1991**, 113, 8402-8409.
6. Kern, D.; Kern, G.; Neef, H.; Tittmann, K.; Killenberg-Jabs, M.; Schneider, C. W.; Hübner, G. *Science* **1997**, 275, 67-70.
7. Kluger, R.; Pike, D. C. *J. Am. Chem. Soc.* **1997**, 99, 4504-4506.
8. Westerik, J. O.; Wolfende, R. *J. Biol. Chem.* **1972**, 247, 8195-8197.
9. Westerilk, J. O.; Wolfenden, R. *J. Biol. Chem.* **1974**, 249, 6351-6353.
10. Thompson, R. C.; Blout, E. R. *Biochemistry* **1973**, 12, 44-47.
11. Baillie, A. C.; Wright, K.; Wright, B. J.; Earnshaw, C.G. *Pestic. Biochem. Physiol.* **1988**, 30, 103-112.
12. He, H. W.; Liu, Z. J.; Wang, J. *Chin. J. Appl. Chem.* **1998**, 3, 88-90.
13. Wang, J.; Chen, X. Y.; Liu, X. F.; Wang, S. Q.; He, H. W. *Huaxue Shi Ji* **1999**, 21, 301-303.
14. Lu, A. H.; He, H. W.; Wang, J.; Wan, S. Q.; Liu, Z. J. *J. Centr. China Normal University (Nat. Sci.)* **1999**, 33, 543-546.
15. He, H. W.; Liu, Z. J.; Li, M. Q.; Lu, A. H.; Hu, L. M.; Yan, G. CN 1151162 C, **2004**.
16. Peng, H.; Wang, T.; Xie, P.; Chen, T.; He, H. W.; Wan, J. *J. Agric. Food. Chem.* **2007**, 55, 1871-1880.
17. Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1998**, 110, 5959-5967.
18. Klebe, G.; Abrahma, U.; Mietzner, T. *J. Med. Chem.* **1994**, 37, 4130-4146.