http://dx.doi.org/10.5012/bkcs.2013.34.10.2909

The Structure-Based Three-Dimensional Pharmacophore Models for *Arabidopsis thaliana* HPPD inhibitors as Herbicide

Jae Eun Cho, Jun Tae Kim, Eunae Kim,† Young Kwan Ko,† and Nam Sook Kang*

Graduate School of New Drug Discovery and Development, Chungnam National University, Daejeon 305-764, Korea
**E-mail: nskang@cnu.ac.kr*
†Division of Convergence Chemistry, Korea Research Institute of Chemical Technology, Yuseong-gu, Daejeon 305-600, Korea
Received June 24, 2013, Accepted July 5, 2013

p-Hydroxyphenylpyruvate dioxygenase (HPPD) is a potent herbicide target that is in current use. In this study, we developed a predictive pharmacophore model that uses known HPPD inhibitors based on a theoretically constructed HPPD homology model. The pharmacophore model derived from the three-dimensional (3D) structure of a target protein provides helpful information for analyzing protein-ligand interactions, leading to further improvement of the ligand binding affinity.

Key Words : HPPD, Herbicide, Pharmacophore, Molecular docking

Introduction

The success rate for herbicidal discovery in agricultural research has been decreasing in recent years. Thus, there is an urgent need for innovation in the area of herbicide discovery. Accordingly, a paramount goal for modern herbicide research is to discover a new potent target that severely disrupts a plant's metabolism by inhibiting essential enzymes for growth. The successful design of a novel herbicide depends on selectivity and safety issues. Selective herbicide targets are usually single enzymes in plants with a conventional mode of action. By uniquely binding to the active site of the target protein, the inhibition activity should affect unwanted weeds and thus protect agricultural products. Regarding the safety issue, the herbicide should follow strict toxicity regulations to mitigate any harmful effects on humans.[1]

p-Hydroxyphenylpyruvate dioxygenase (HPPD; EC 1.13.11.27), a non-heme Fe(II)-dependent enzyme, is a new target for bleaching herbicides. HPPD catalyzes the conversion of 4-hydroxyphenyl pyruvate to homogentisic acid and carbon dioxide in the presence of oxygen and ferrous ion. Because homogentisic acid is a key precursor of plastoquinone and vitamin E, which are essential elements for the light-dependent reaction of photosynthesis, the inhibition of HPPD leads to photodynamic bleaching of the foliage. For safety reasons, most herbicides have a low toxic effect on mammals because the targeted biological pathway exists only in plants, such as photosynthesis, essential amino-acids biosynthesis, or chlorophyll biosynthesis. HPPD, on the other hand, shares the metabolic pathways between mammals and plants. However, HPPD is an effective herbicide target that has low mammalian toxicity.[2-4]

Pharmacophore model is widely used as a powerful tool for hit to lead optimization in the field of medicinal chemistry. Essentially, the initial discovery process of agrochemistry is not much different from that of medicinal chemistry. In the field of agrochemistry, *in vivo* screening requires a considerable amount of effort and time (~1 month). Therefore, if we make sufficient use of Pharmacophore model, we will discover active agro-compounds faster. Several publications have reported the herbicidal activity of HPPD inhibitors through Pharmacophore studies.[5,6] However, the known studies only explored analogous compounds to initiate the Pharmacophore study.

In this research, we developed the structure-based pharmacophore model to develop HPPD inhibitor. First, the 3D-structure of *Arabidopsis thaliana* (*At*) HPPD was described *via* a theoretical method. By utilizing the information of the *At*HPPD homology model structure, we constructed a pharmacophore model. Furthermore, the power of our pharmacophore model was validated using various classes of *At*HPPD inhibitors, including triketone, diketone, NTBC, sulcotrione and leptospermone.

Experimental Section

Workflow. Three-dimensional (3D) structure of *At*HPPD protein was constructed in order to apply the structure-based drug design (SBDD) method. 3D-pharmacophore models which created from the receptor structure were used as queries for a virtual screen of in-house compounds. The screening results are analyzed by using the computational tools.

Structure Preparation. To apply the structure-based drug design method using current knowledge of protein and drug interactions, a three-dimensional (3D) protein structure is necessary.[7] Although the known protein crystal structural information of HPPD from an experiment was exist such as 1TFZ,[8] 1TG5,[8] 1SP9[9] and 1SQD[8] in the Protein Data Bank, residues around binding site were truncated for all known HPPD crystal structures. However, the crystal structure of *Streptomyces avermitilis* HPPD (1T47[10]) does not have the truncated region anywhere. Since the conformation of the truncated residues strongly affect the binding mode of input

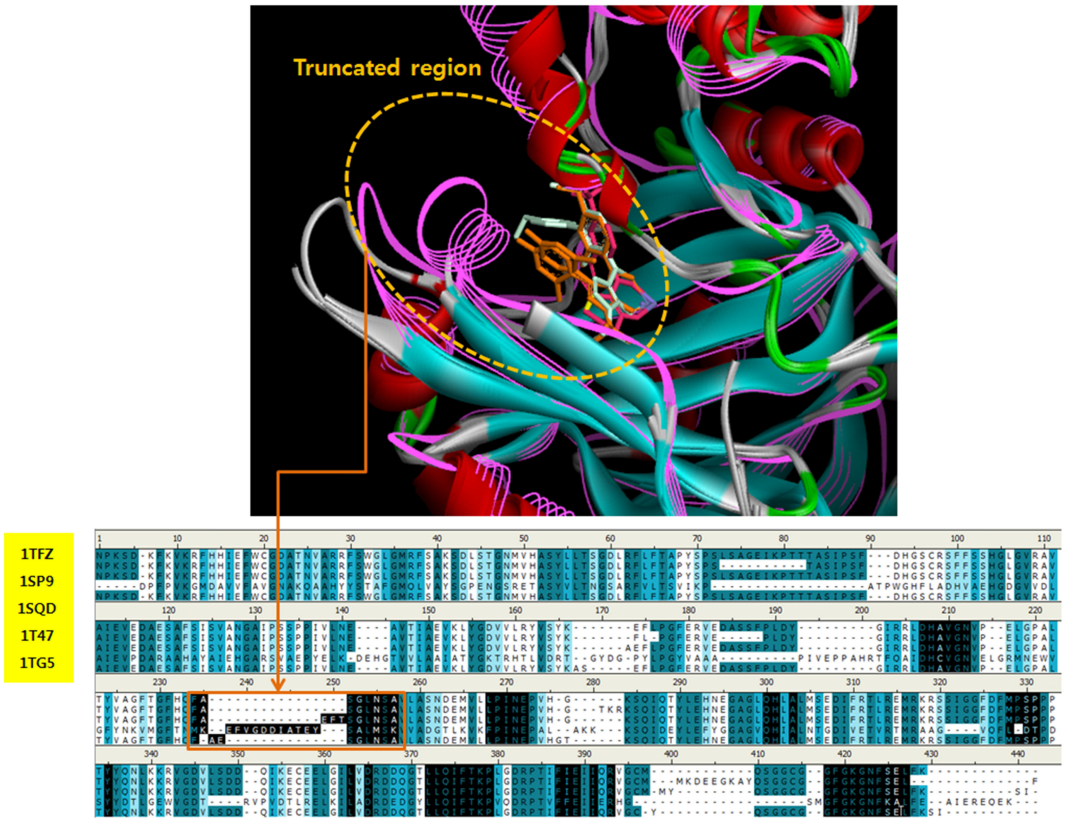


Figure 1. The superimposed five different protein crystal structures originated from different species but structurally similar to *Arabidopsis thaliana* HPPD. The crystal structure of *Streptomyces avermitilis* HPPD (1T47) does not have truncated region within 7 Å of binding site unlike the other *Arabidopsis thaliana* HPPD structures.

compounds, *Streptomyces avermitilis* HPPD (1T47) was selected as a template structure for developing a homology model of *At*HPPD as well as the amino acid sequence of *At*HPPD (Accession code: Q9SK87). To obtain the 3D-structure of *At*HPPD, we adopted a hierarchical protein structure modeling approach based on secondary-structure enhanced Profile-Profile threading Alignment (PPA) and iterative implementation by the Threading ASSEmbly Refinement (TASSER) program.[11] We obtained five candidate models for the three-dimensional structure of *At*HPPD and then performed molecular dynamics simulations with the use of the CHARMM force field (version 27.0)[12] with default parameters interfaced with Accelrys Discovery Studio 3.5. In order to identify binding sites, we collected the known crystal structures of homologous proteins with *At*HPPD as templates from the RCSB Protein Data Bank (PDB). Five different protein crystal structures originating from different species but structurally similar to *At*HPPD were selected and superimposed as shown in Figure 1 (PDB id: 1TFZ, 1TG5, 1SP9, 1SQD, 1T47). Based on the information of all overlapped known crystal structure we defined the binding site.

Pharmacophore Model Generation. The receptor structure developed by homology modeling was used for pharmacophore model generation to facilitate a novel scaffold development for *At*HPPD. To build the pharmacophore model for screening, first of all, possible interaction points from active site of *At*HPPD protein were generated using the 'Interaction Generation' protocol implemented in Discovery Studio software package version 3.5. The active crystal ligand of 1TF7 was set as reference compound for generating interaction points. The parameters for both 'Density of Lipophilic Sites' and 'Density of Polar Sites' were defined as 25. The lists of features of the minimum and maximum values were the following manner: H-bond acceptor (Hba) 0 and 5, H-bond donor (Hbd) 0 and 5, Hydrophobic (Hy) 0 and 5 Aromatic ring (Ra) 0 and 5.

Result and Discussion

Pharmacophore Model. As a result of the all possible interaction point calculation, total 255 features were generated. Green, magenta, cyan feature represents hydrogen bonding acceptor, hydrogen bonding donor and hydrophobic respectively as shown in Figure 2. Among various interaction points of the active site, we manually selected two hydrogen acceptor features that involves in Fe(II)-ion chelation in the active site of the *At*HPPD. In addition, aromatic ring feature, the key position to define the π-π stacking interaction with Phe 381 was selected. Furthermore, hydrogen acceptor feature to define the area of the solvent accessibility region and a hydrophobic feature to define the interaction with Phe 392 and Met 335 were selected as depicted in Figure 3. Exclu-

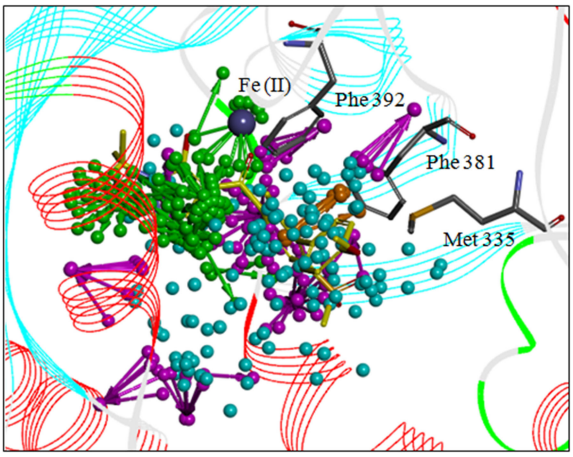


Figure 2. Possible interaction points in the active site of *At*HPPD homology model. Green, magenta, cyan orange feature represents hydrogen bonding acceptor, hydrogen bonding donor, hydrophobic and aromatic ring respectively.

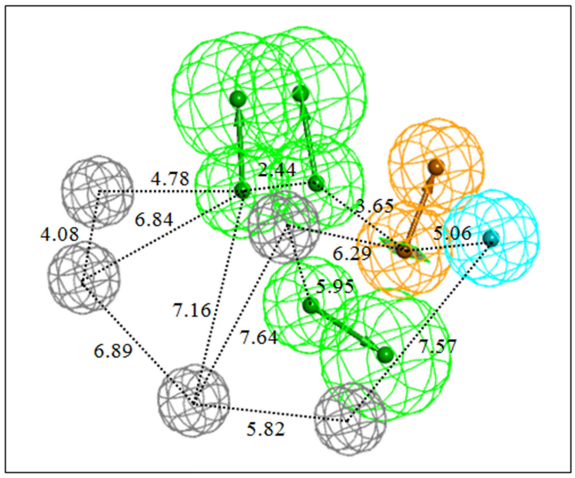


Figure 3. Manually constructed pharmacophore features. Green, orange, cyan and gray features and its spheres represent hydrogen bonding acceptor, aromatic ring and excluded area respectively.

sion volume spheres were added to the structure based models onto coordinates defined by protein side chain atoms to characterize inaccessible areas for any potential ligand. Further model refinement such as addition of shape constraints was carried out according to the shape of the protein binding site. Green, orange, cyan and gray features and its spheres shown in Figure 3 represent hydrogen bonding acceptor, aromatic ring, hydrophobic and excluded areas, respectively.

Table 1. The validation set compounds

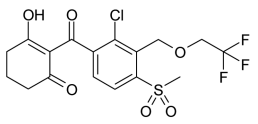
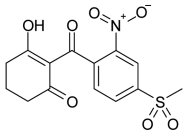
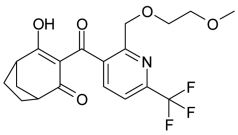
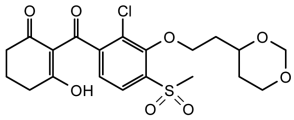
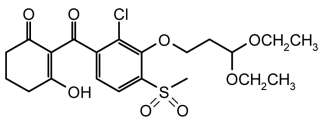
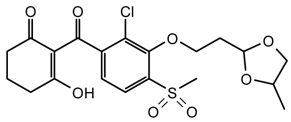
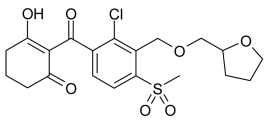
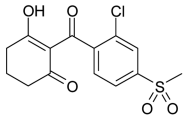
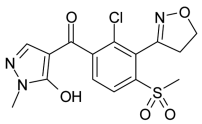
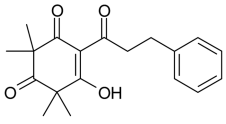
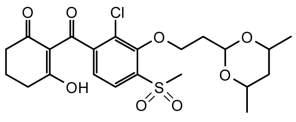
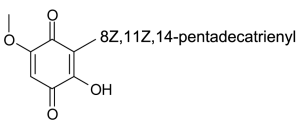
Compound	Structure	pIC_{50}	Compound	Structure	pIC_{50}
1		8.0	**11**		8.0
2		8.0	**12**		8.0
3		7.7	**13**		8.0
4		8.0	**14**		8.0
5		8.0	**15**		6.7
6		8.0	**16**		6.4

Table 1. The validation set compounds (continued)

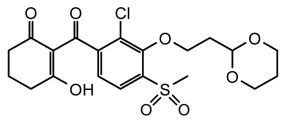
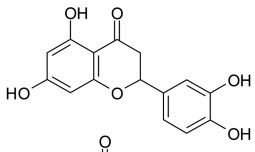
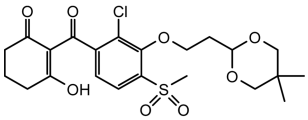
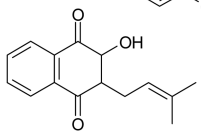
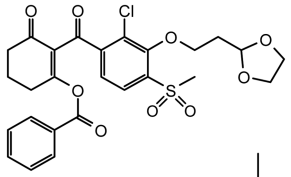
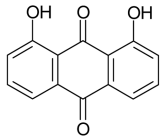
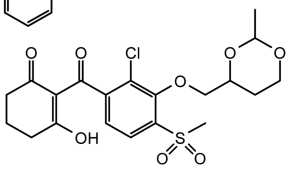
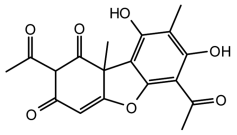
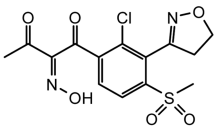
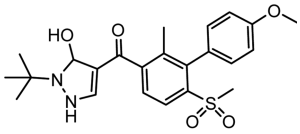
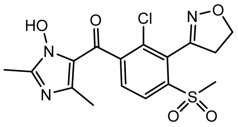
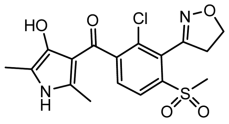
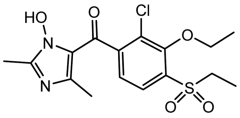
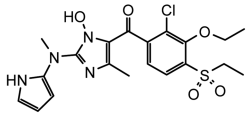
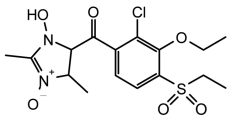
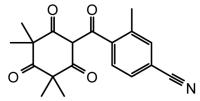
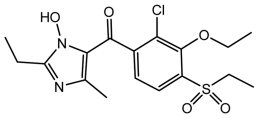
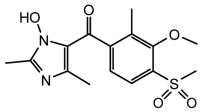
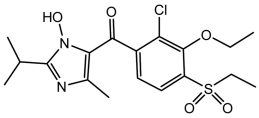
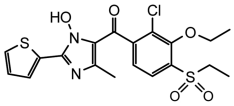
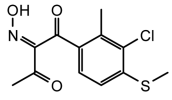
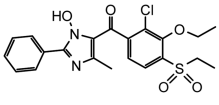
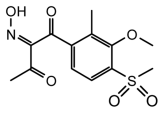
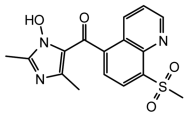
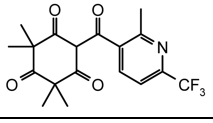
Compound	Structure	pIC_{50}	Compound	Structure	pIC_{50}
7		8.0	**17**		4.0
8		8.0	**18**		4.0
9		7.7	**19**		4.0
10		8.0	**20**		7.1
21		4.9	30		8.2
22		7.4	**31**		6.7
23		7.3	**32**		4.7
24		6.7	**33**		10.3
25		7.1	**34**		4.7
26		7.4	**35**		4.7
27		4.7	**36**		4.7
28		4.7	**37**		4.7
29		10.2			

Table 2. The result of virtual screening

Compound	pIC_{50}	Fit Value
1	8.0	4.1
2	8.0	3.9
3	7.7	3.2
4	8.0	3.8
5	8.0	3.8
6	8.0	3.4
7	8.0	3.6
8	8.0	3.7
9	7.7	3.1
10	8.0	3.6
11	8.0	3.5
12	8.0	3.4
13	8.0	3.4
14	8.0	3.3
15	6.7	2.1
16	6.4	1.8
17	4.0	0.4
18	4.0	0.1
19	4.0	0.3
20	7.1	2.3
21	4.9	1.8
22	7.4	3.9
23	7.3	3.7
24	6.7	2.4
25	7.1	3.5
26	7.4	3.8
27	4.7	1.9
28	4.7	1.2
29	10.2	3.5
30	8.2	3.2
31	6.7	2.5
32	4.7	0.1
33	10.3	3.4
34	4.7	0.8
35	4.7	1.7
36	4.7	1.6
37	4.7	0.1

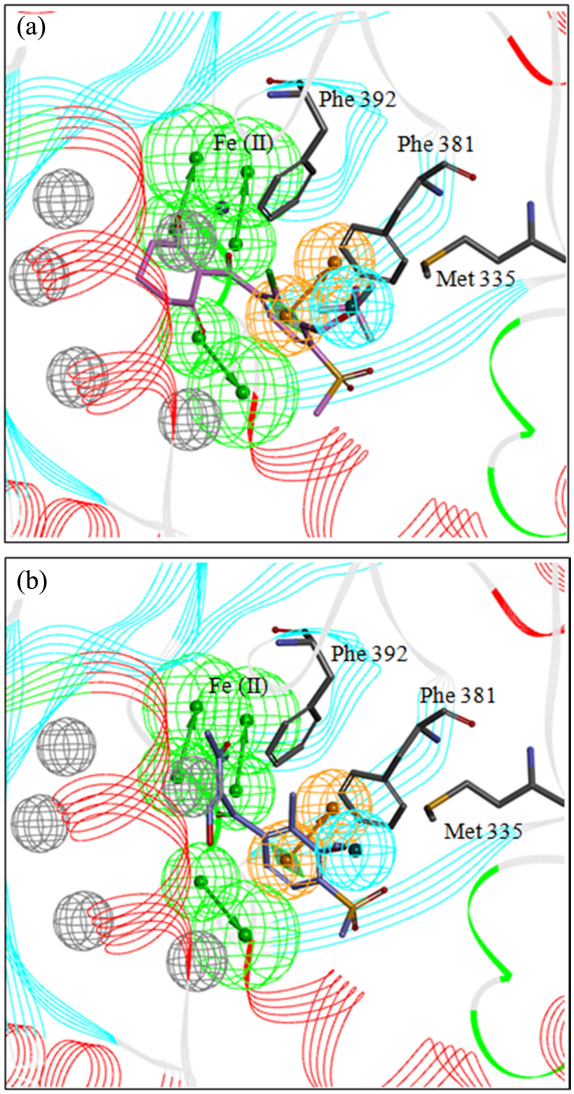


Figure 4. The binding mode of active compound (a) and inactive compound (b). Compound **1** (pIC_{50} = 8.0) is nicely mapped with our model illustrate in (a). In (b), compound **28** (pIC_{50} = 4.7) is overlapped with our model.

Pharmacophore Model Validation. The validation data set of *At*HPPD inhibitors were chosen from the several published patents and reported research papers.[13-16] The total 37 compounds of data set shown in Table 1 were used as the validation set. In order to validate the prediction power of our receptor based pharmacophore model, 'Screening Library' protocol implemented in Discovery Studio 3.5 (Accerlys, Inc.) was utilized. For the advanced setting, each conformation of the input ligands was set to be flexible for better fit of the pharmacophore. Significantly high correlation coefficient (r^2 = 0.77) between the biological activity and the fit value was obtained from the validation experiment, and we reported the fit values for training set as shown in Table 2.

Binding Mode. The active ligand (compound **1**) and inactive ligand (compound **28**) are mapped with our model as shown in Figure 4. For the active compound (pIC_{50} = 8.0), all features were nicely mapped. However, the Fe(II) interaction site does not mapped with the hydrogen bond acceptor features for the inactive compound (pIC_{50} = 4.7) because of high flexibility of the compound.

ROC (Receiver Operating Characteristic) Curve. To analyze enrichment based on a set of poses generated by a 'Screening Library' method, 'Calculate ROC Curve' protocol was used. The output ligands data from the virtual screening were used directly for input. The curve was created by plotting true positive rate for x-axis against false positive rate on y. The accuracy of the test is reported by measuring the area under the curve (AUC). The result of our model represented with the excellent AUC score of 0.97.

Hit Rate. We also used a specialized plotting method to analyze the predictive power of our model. The 'Hit Rate Plot' enables to visualize the raking of the input compounds

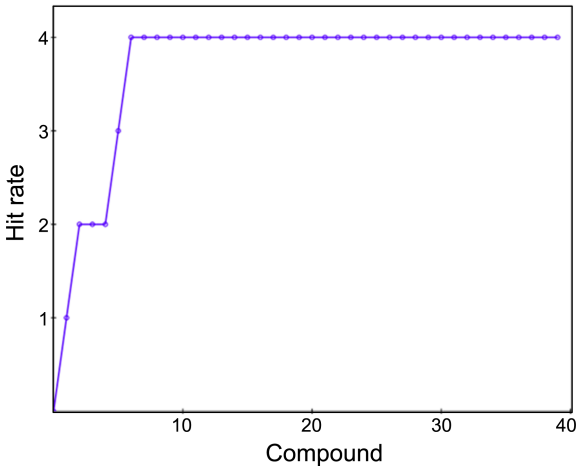


Figure 5. Hit rate plot of our model. Each circle represents the input compounds.

with respect to a given scoring function. To generate a 'Hit Rate Plot', we used the 37 output compounds from the virtual screening to input data. The most active ligand (compound **33**) among the input compounds was set to control ligand. The resulted plot shows the number of hits encountered on the y-axis plotted against the ligand rank on the x-axis yields the Hit Rate Plot (Figure 5).

Conclusion

The potent herbicide target *Arabidopsis thaliana* HPPD has been thoroughly studied to protect agricultural products. Various researchers have attempted to discover potent inhibitors for *At*HPPD and have created a predictive model for an analogous HPPD inhibitor originating from diverse organisms. However, we used various classes of *At*HPPD inhibitors to overcome the limitations for a prediction.

In this study, a predictive 3D-pharmacophore model from a receptor-ligand complex was generated to provide insight for further study of the herbicide. A highly predictive screening result was obtained from the validation process. The subsequent high ROC score of 0.97 and the hit rate proves the predictive power of our 3D-pharmacophore model. This model can be used to analyze known compound data and can be used for scaffold hopping. We will design the new moiety interacting with iron and Phe 381 based our pharmacophore model to improve the biological activity. Thus, we expect that the proposed model will provide useful information for analyzing protein-ligand interactions that can lead to a further improvement of ligand binding affinity levels.

References

1. Beaudegnies, R.; Edmunds, A. J.; Fraser, T. E.; Hall, R. G.; Hawkes, T. R.; Mitchell, G.; Schaetzer, J.; Wendeborn. S.; Wibley, J. *Bioorg. Med. Chem.* **2009**, *17*, 4134.
2. Raspail, C.; Graindorge, M.; Moreau, Y.; Crouzy, S.; Lefèbvre, B.; Robin, A. Y.; Dumas, R.; Matringe, M. *J. Biol. Chem.* **2011**, *286*, 26061.
3. Shaner, D. L. *Pest. Manag. Sci.* **2004**, *60*, 17.
4. Matringe, M.; Sailland, A.; Pelissier, B.; Rolland, A.; Zink, O. *Pest. Manag. Sci.* **2005**, *61*, 269.
5. Huang, M.; Yang, D. Y.; Shang, Z.; Zou, J.; Yu, Q. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2271.
6. Dayan, F. E.; Singh, N.; McCurdy, C. R.; Godfrey, C. A.; Larsen, L.; Weavers, R. T.; Van Klink, J. W.; Perry, N. B. *J. Agric. Food Chem.* **2009**, *57*, 5194.
7. Krieger, E.; Nabuurs, S. B.; Vriend, G. *Structural Bioinformatics* **2003**, *25*, 507.
8. Yang, C.; Pflugrath, J. W.; Camper, D. L.; Foster, M. L.; Pernich, D. J.; Walsh, T. A. *Biochemistry* **2004**, *43*, 10414.
9. Fritze, I. M.; Linden, L.; Freigang, J.; Auerbach, G.; Huber, R.; Steinbacher, S. *Plant Physiol.* **2004**, *134*, 1388.
10. Brownlee, J.; Johnson-Winters, K.; Harrison, D. H. T.; Moran, G. R. *Biochemistry* **2004**, *43*, 6370.
11. Zhang, Y. *Proteins* **2007**, *69*,108.
12. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J. *J. Comp. Chem.* **1983**, *4*, 187.
13. Meazza, G.; Scheffler, B. E.; Tellez, M. R.; Rimando, A. M.; Romagni, J. G.; Duke, S. O.; Nanayakkara, D.; Khan, I. A.; Abourashed, E. A.; Dayan, F. E. *Phytochemistry* **2002**, *60*, 281.
14. Dayan, F. E.; Duke, S. O.; Sauldubois, A.; Singh, N.; McCurdy, C.; Cantrell, C. *Phytochemistry* **2007**, *68*, 2004.
15. Witschel, M. *Bioorg. Med. Chem.* **2009**, *17*, 4221.
16. Nakamura, Y.; Sano, M. WO/2005/118530.