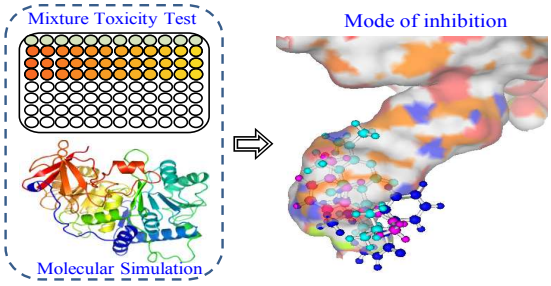
RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances



Molecular simulation techniques are used to identify the mode of inhibition of chemical at ligand-receptor level.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

Paper

Predicting the mixture effects of three pesticides by integrating the molecular simulation with the concentration addition

Fu Chen,[a] Shu-Shen Liu,* [a,b] Xin-Tian Duan [a] , and Qian-Fen Xiao [a,b]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

DOI: 10.1039/b000000x

It is a ticklish but fundamental problem in risk assessment to predict the combined toxicity of several chemicals by methods except experiment, especially the binary mixtures, in order to save time and experimental cost. Though several models have been developed, making choice among them would be confused by the variety of the mode of inhibition (MOI). For choosing a reference model appropriately, we propose an *in silico* procedure which employs molecular simulation techniques to identify MOI, which highlights the binding pattern of a small molecule to biomacromolecule. Specifically, the method is verified by the experimental study that fifteen binary mixtures of three pesticides, baygon, metacrate, and velpar, inhibit firefly luciferase bioluminescence. The results reveal that the pesticides share the same binding site at the bottom of luciferin pocket and combined toxicities could be predicted by the concentration addition model, which enables us to identify the MOI using molecular simulation techniques. In addition, there is a linear relationship between the mixture binding free energy (ΔG_{mix}) calculated from ΔGs of the components and the median effective concentrations (EC_{50}) of the mixture.

[a] *Key Laboratory of Yangtze River Water Environment, Ministry of Education, College of Environmental Science and Engineering, Tongji University, Shanghai 200092, PR China. Tel: 86-021-65982767, E-mail: ssliuhl@263.net*

[b] *State Key Laboratory of Pollution Control and Resource Reuse, College of Environmental Science and Engineering, Tongji University, Shanghai 200092, PR China*

† Electronic Supplementary Information (ESI) available:

Introduction

It is impractical to test the toxicity of every mixture in the environment. Therefore, several prediction models had been proposed, among which there are two widely available models, the concentration addition model (CA) and the independent action model (IA). [1-3] The CA[4] which is optimal for the mixture where the components have the similar mechanisms of action (MOA). Whereas the IA[5] which is suitable for the combined toxicity generated from ones having dissimilar MOAs. However, in practice, obtaining the MOAs information of a chemical against various biological targets is difficult even experimentally, which means it is a dilemma to put the models into practice.[6, 7] To get the MOA information, an existing feasible solution is quantitative structure-activity relationship (QSAR) model which could discriminate different MOAs by using chemicals' structure descriptors.[8, 9] Nonetheless, limitations of the model, deriving from limited applicability domain and small training set, restrain the utilization of QSAR. Obviously, a more feasible method is essential. Based on the viewpoint presented by some researchers that the chemicals having different MOAs are toxic in different ways due to different interactions at biomolecular level.[8, 10] we attempt to introduce computer simulation study into the process where a visible 3D pattern of small chemical molecules bound to the specific protein molecular target is simulated.

The firefly luciferase extracted from *Photinus pyralis* is a widespread bioluminescence indicator used in PubChem.[11, 12] In biochemical assays, firefly luciferase catalyzes the oxidation of D-luciferin (LH_2) in the presence of oxygen, ATP, and Mg^{2+} to form oxyluciferin and yield light simultaneously.[13] The luciferase inhibitors were divided into two primary competition categories, the luciferin and adenylate competitive inhibitors, according to the previous researches focusing on the structure-activity relationships of luciferase inhibitors,[14, 15] which further verify the existence of two binding pockets in luciferase protein.[15, 16] Later, some cocrystal structures of ligands bound to luciferase revolved by experiments and modeling confirm the pockets. [12, 17] Modeling results of substrate competitive inhibitors, quinoline and benzthiazole, exposed to Japanese firefly luciferase, reveal that the quinoline occupies approximately the same space where the purine of AMP locates, whereas the benzthiazole occupies nearly identical site as luciferin does.[16] In addition, the substitution of each extends into other binding region despite of the major binding pocket.[16] Auld[18] firstly explored the X-ray cocrystal structure of luciferase bound to Ataluren (PTC124, an analogue of luciferin) and substrate ATP being determined to 2.0 Å resolution (PDB: 3IES), which makes it available to have a

deeper insight into inhibition mechanism.

The purposes of this paper are, (1) to analyse the MOI of three pesticides (baygon, metacrate, and velpar) to luciferase by molecular simulation; (2) to determine the toxicities of three single pesticide and fifteen binary mixtures experimentally; (3) to choose an appropriate reference model to predict the toxicities of binary mixtures according to the results of (1). Specifically, for the first purpose, four steps are included, molecular docking, molecular dynamics simulation (MD), Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) energy calculations and MM/GBSA free energy decomposition analysis[19]. To achieve (2), the microplate toxicity analysis (MTA)[20] is chosen, following the procedures of the mixture composition design, concentration-response data sampling, and CA model validation. The results from (1) and (2) indicate that the toxicities of the mixtures could be well predicted by the CA model on the situations where the pesticide components share the same binding pattern to luciferase.

Methods

Structure of firefly luciferase

Generally, the firefly luciferase structure consists of a large N-terminal domain (residues 1-436), a small C-terminal (residues 440-551), a short hinge and three flexible peptides (residues 437-439).[21] The active site of the luciferase is mainly located in the N-terminal domain near the hinge region.[11, 22] In our molecular simulation system, we employed the 2.0 Å resolution cocrystal structure from PDB (Enter code of 3IES) to fit the structure of luciferase. The 3IES is a complex of PTC124-AMP bound to the *Photinus pyralis* luciferase. The structure of the luciferase, which is called as LUC in the next context, in 3IES complex only contains the 436 residues in N-terminal. The LUC has two active binding pockets: one is called ATP pocket where hydrogen bonds forms between AMP and Gly339, Gly316, Thr343, His245, Asp422 separately and the other one is called LH_2 pocket surrounded by Ala222, Phe227, Phe247, Ala313, Ala348, and Ile351 where there is no hydrogen bond but only hydrophobic interactions between PTC124 and LUC.[18] The two important binding pockets, the ATP pocket and LH_2 pocket,[23] were shown in Fig. 1 drawn by PyMol 0.99rc6 [24] and LigPlot+ [25].

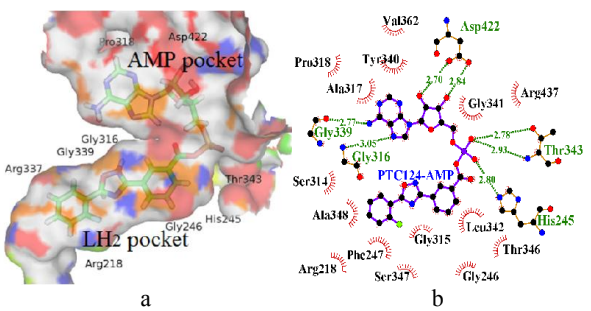


Fig. 1. (a) The binding pocket of firefly luciferase, in 3IES. (b) The hydrogen bond interaction patterns and hydrophobic contacts between the PTC124-AMP and the side chain and backbone of binding residues in 3IES, where hydrogen bond contact was expressed in green line and hydrophobic contact groups was expressed in eyelash shape.

Molecular docking and molecular dynamics simulation

The receptor structure for molecular docking was obtained by removing all water molecules as well as the ligand from the 3IES in which the missing residues were constructed by Discovery Studio 2.5 (Accelrys Software, Inc., San Diego, CA)[26]. The ligand files for pesticides were read in PyRx 0.8, all the hydrogens and Gasteiger[27] charges were added and non-polar hydrogens were merged. The structure of pesticide-LUC was obtained by semi-flexible docking approach using AutoDockVina.[28] LUC was used as the receptor, considered as if it is fully rigid. The pesticide ligand was flexible. For the docking calculations, a box of size 13.44 × 17.60 × 17.68 Å was used, centered at the geometric center of PTC124-AMP structure. The exhaustiveness value (exhaustiveness of finding the global minimum) was changed to 25 (default is 8), and the program was allowed to generate 10 binding modes (default is 9).[29] Explicit random seed is used for Genetic Algorithm. The maximum energy difference between the best binding mode and the worst one displayed is 3 kcal/mol. The docking pose with the lowest negative score value (highest binding affinity) was chosen as the initial conformation for MD.

MD was performed by using the AMBER 12 software package.[30] Three pesticide-LUC complexes (the generalized AMBER force field (GAFF)[31] for pesticide ligand and the ff99SB force field[32] for LUC protein) were solvated by the TIP3P waters, with a minimum distance of 8.5 Å from the complex surface.[33] All MDs were conducted by the standard procedure, which is comprised of energy-minimization, gradual heating of the systems, and isothermal isobaric ensemble (NPT) molecular dynamics.[34, 35] The ptraj module in AMBER was used to analyze root mean-square displacements (RMSD) and root mean-square fluctuation (RMSF). All systems were equilibrated at 4 ns, and the MDs were prolonged for another 4 ns. One hundred snapshots of the simulated structures within the last 1 ns with a step of 10 ps were sampled. More details for molecular docking and MD are given in the Supplementary Information.

Binding free energy and its components

The binding free energies (ΔG_is) of pesticides with LUC were calculated using the 100 snapshots of each complex (every 10 ps) generated from the last 1ns MD trajectories. The enthalpy (ΔH) was calculated by using MM/GBSA procedure[19, 36] in AMBER 12. The conformational entropy (translation, rotation, and vibration) upon ligand binding, $T\Delta S$, was calculated using normal-mode analysis[37] with the nmode program in AMBER 12. More details are given in the Supplementary Information.

Furthermore, the ΔG was decomposed to a single residue using MM/GBSA method. This decomposition was performed only for molecular mechanics and solvation energies but not for entropies.[38] The MM/GBSA free energy decomposition[39] procedure was used to calculate pesticide-residue pair energy (ΔG_r) between each pesticide and each individual residue.[40]

The ΔG_{mix} of a mixture was calculated from the ΔG_is of pesticide components and the p_i of the component in the mixture,

$$\Delta G_{mix} = \sum_{i=1}^{n} p_i \times \Delta G_i \qquad (1)$$

where p_i is defined as a ratio of concentration of the ith component in the mixture to the mixture concentration, ΔG_{mix} is

the assumed binding free energy of a mixture.

Determination of toxicity

The chemicals include ATP-Na_2 (Sigma-Aldrich, St. Louis, MO, ≥ 98.0% purity), QuantiLum recombinant luciferase (Promega, Madison, WI, ≥ 95% purity), endotoxin-free D-luciferin (Promega, Madison, WI, ≥ 98.5% purity), which were separately stored in the glycylglycine buffer (pH 7.8, consisting of 50 mmol/L glycylglycine, 1 mmol/L $MgSO_4$, 0.5 mmol/L EDTA, and 10 mmol/L DTT). The final optimal conditions used in our luciferase luminescence inhibition toxicity test were as follows: the luciferase concentration of 1×10^{-8} mol/L, the luciferin concentration of 1.6×10^{-5} mol/L, the ATP concentration of 1×10^{-5} mol/L, pH 7.8, and 15 min exposures at 22 ± 1 °C.

Baygon (BAY) and metacrate (MET) are methyl carbamate insecticides, and velpar (VEL) is a triazine herbicide. The concentration-response datas of single pesticides and the mixtures were determined by the MTA[20]. The relative light unit (RLU) was determined on Synergy™ 2 multi-mode microplate reader (BioTek instruments Inc., USA) with a 96-well white flat-bottom microplate (Corning, USA).

All pesticides were dissolved in 1% (*V/V*) DMSO and stored at room temperature. Twelve concentration series of pesticides and their mixtures in three parallels and twelve controls were arranged in a microplate. First, 100 μL 1% (*V/V*) DMSO was added to twelve wells of the first row in the microplate as controls. Second, 100 μL solutions of pesticide or their mixtures of twelve gradient concentrations derived by dilution factors *f* were added to twelve column wells from the second to the fourth row as treatments. Then, 50 μL ATP solution, 50 μL luciferin solution, and 50 μL luciferase solution were successively added into each well to make the final test volume be 250 μL. The RLUs of various treatments and controls were determined after exposure of 15 min. Each microplate test was repeated at least twice.

The toxicity was expressed as a percentage inhibition (*E* or *x*) and the observed concentration-response data were fitted to Weibull function, called concentration response curve fitting (CRC). As a quantitative measure of the uncertainty, the 95% observation-based confidence intervals were also determined.[41] From the OCIs, the uncertainty of some concentration such as EC_{50} was also determined by linear interpolation. More details for toxicity testing procedure are given in the Supplementary Information.

Mixture design and additivity validation

The direct equipartition ray design (EquRay) procedure[42] was employed to rationally design the concentration ranges of pesticides in some representative binary mixtures. Based on the EC_{50} of signal pesticide (Table S1), 15 binary mixture rays were designed by EquRay. The p_is of various pesticides in the mixtures were listed in Table 1.

The CA model[1, 43] was employed to predict the combined toxicity based on dose addition assumption. If the observed toxicity is consistent with that predicted, then the combined toxicity has the additivity or non-interaction. The CA can be written as follows[1, 43]:

$$EC_{x,mix} = \left(\sum_{i=1}^{n} p_i / EC_{x,i} \right)^{-1} \quad (2)$$

where $EC_{x,mix}$ is the concentration of the mixture that cause *x* combined toxicity, *n* is the number of mixture components, $EC_{x,i}$ is the concentration of the *i*th component causing *x* toxicity when applied individually, and p_i is the concentration ratio of the *i*th component in the mixture.

Table 1. The mixture ratios (p_i), model parameters (α and β), statistics (RMSE and R^2), EC_{50} (10^{-3} mol/L), lower and upper limits of 95% confidence intervals of EC_{50}, and binding free energy values (ΔG_{mix}, kcal/mol) of binary mixture rays

Ray	p_1	α	β	RMSE	R^2	EC_{50} (lower, upper)	ΔG_{mix}	EC_{50}[a]
BAY-MET system ($p_1 = p_{BAY}$)								
R1	0.731	6.25	2.36	0.010	0.999	1.572 (1.461, 1.723)	-11.0	1.653
R2	0.521	5.56	2.18	0.013	0.997	1.911 (1.718, 2.171)	-9.50	1.901
R3	0.353	5.66	2.23	0.022	0.993	1.984 (1.679, 2.409)	-8.27	2.123
R4	0.214	5.36	2.16	0.012	0.998	2.232 (2.041, 2.498)	-7.26	2.322
R5	0.098	4.97	2.08	0.016	0.996	2.718 (2.369, 3.172)	-6.41	2.502
MET-VEL system ($p_1 = p_{MET}$)								
R1	0.328	7.39	2.74	0.019	0.997	1.476 (1.314, 1.676)	-14.2	1.487
R2	0.550	6.43	2.47	0.023	0.994	1.771 (1.512, 2.099)	-11.4	1.808
R3	0.710	6.43	2.54	0.015	0.997	2.109 (1.908, 2.359)	-9.38	2.081
R4	0.830	5.22	2.12	0.009	0.999	2.316 (2.171, 2.516)	-7.85	2.311
R5	0.920	4.91	2.03	0.013	0.998	2.516 (2.276, 2.830)	-6.66	2.522
VEL-BAY system ($p_1 = p_{VEL}$)								
R1	0.790	7.63	2.75	0.029	0.992	1.238 (1.019, 1.490)	-17.2	1.209
R2	0.600	6.79	2.47	0.029	0.990	1.267 (1.030, 1.569)	-16.2	1.272
R3	0.429	6.67	2.46	0.021	0.995	1.379 (1.194, 1.607)	-15.3	1.332
R4	0.273	6.25	2.31	0.018	0.996	1.368 (1.199, 1.581)	-14.4	1.394
R5	0.131	6.28	2.34	0.019	0.995	1.446 (1.265, 1.688)	-13.7	1.443

[a]: The values of the EC_{50}s estimated from linear models (Eqs. 3 - 5)

Results and discussion

Pesticide binding pattern in active LUC

As the results of docking, for each pesticide a set of 10 plausible ligand poses was gained and shown in Fig. S1. According to the scores, the best eight poses of BAY are located at the bottom of LH_2 pocket (score = -5.9), while the other two poses are at the entrance (score = -5.1). For MET, all poses are at the bottom. Almost every pose of VEL is located at bottom (the best score = -8.5) except one at the entrance (score = -6.1). In other word, all clusters of poses have an obvious inclination that all the three pesticides tend to bind at the bottom of LH_2 pocket. The best scoring pose was chose as the initial configuration for MD.

Molecular dynamics simulations for pesticide-LUC were run and equilibrated after 4 ns equilibrium phase according to the convergences of RMSD between the trajectory structures and the first snapshot structure in 1th ns trajectory (Fig. S2). The RMSD of the protein backbone of BAY-LUC stabilized around 1.29 Å from 5 to 8 ns, and the standard deviation is 0.07 Å. The convergence of RMSD is 1.56 ± 0.09 Å for MET-LUC and 1.58 ± 0.09 Å for VEL-LUC. For each complex, the lowest potential energy conformation taken from the 100 snapshots of the last 1 ns simulation was treated as binding pattern of pesticide with LUC.

The crucial residues surrounding the pesticides were shown in Fig. S3.

The hydrogen bonds were calculated according to the following two criteria: (1) a proton donor-acceptor distance ≤ 3.5 Å, and (2) a donor-H-acceptor bond angle ≥ 120°. For BAY and MET, a hydrogen bond between the -COO- group of the pesticide and the -NH2 group of Arg218 is observed in the lowest energy conformation. For VEL, the NH group of Arg218 donates a H to the O of carbonyl group of VEL forming a stable hydrogen bond (3.50 Å), and another strong hydrogen bond (2.70 Å) forms between another carbonyl group of VEL and the H atom of the NH group of Arg337.

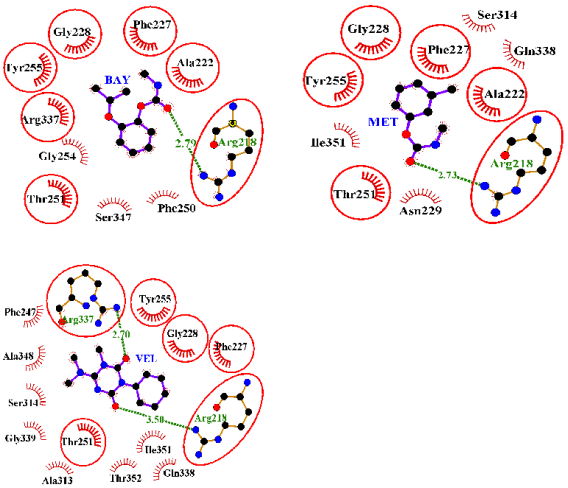


Fig. 2 Schematic view of the pesticide-LUC interactions of lowest energy conformation, where the residues enclosed by the red circles and ellipses at least appear twice in three systems.

Interaction mechanism of pesticide-LUC systems

The calculated ΔG_is and their components were presented in Table S2. The bigger absolute value of ΔE or ΔG (negative value) means the stronger binding affinity.

If the polar solvation free energy $\Delta G_{ele,sol}$ is taken into consideration, the value of polar interaction contributions ($\Delta G_{ELE} = \Delta E_{ele} + \Delta G_{ele,sol}$) turns to be an unfavorable factor for binding three pesticides (9.0, 6.9, and 4.7 kcal/mol for BAY-, MET-, and VEL-LUC, respectively), which is to a large extent determined by the desolvation energy.[44] The similar phenomenon that contribution from ligand-protein polar interactions could not compensate for the large desolvation penalty has been found in other researches.[40, 45] Furthermore, ΔE_{vdw} is found to have a similar favorable contribution in all systems, and the nonpolar energy ($\Delta G_{VMD} = \Delta E_{vdw} + \Delta G_{nonpolar,sol}$) term considering the nonpolar solvation term shows more favorable contribution (-42.0, -30.6, and -44.2 kcal/mol for BAY-, MET-, and VEL-LUC, respectively).

Considering the summation of the energy components, the calculated ΔH_is are -33.1, -23.7, and -39.5 kcal/mol for the BAY-LUC, MET-LUC, and VEL-LUC, respectively. After introduction of the $T\Delta S$, the ΔG_is for BAY-LUC, MET-LUC and VEL-LUC become -13.0, -5.7 and -18.4 kcal/mol, respectively.

The thermodynamic parameters (ΔH and ΔS) are the main evidence for confirming the binding force.[46, 47] Ross and Subramanian[48] have summed up the thermodynamic laws to determine the types of binding with various interactions. In this study, $\Delta H < 0$ and $\Delta S < 0$ means that van der Waals and hydrogen bond interactions play the main roles in the binding reaction. On the one hand, for each pesticide, either the van der Waals interaction or the nonpolar part of solvation play a crucial role in binding affinity. On the other hand, it is more convincing that the hydrogen bond interaction is responsible for providing directional constraints for the binding progress.

Per pesticide-residue interactions

The ΔG_rs of 25 residues which have the highest energy contributions for each pesticide are summarized in Fig. S4. In addition, the three series of ΔG_{ELE} and ΔG_{VDW} energy terms of each residue were also shown in Fig. 3,

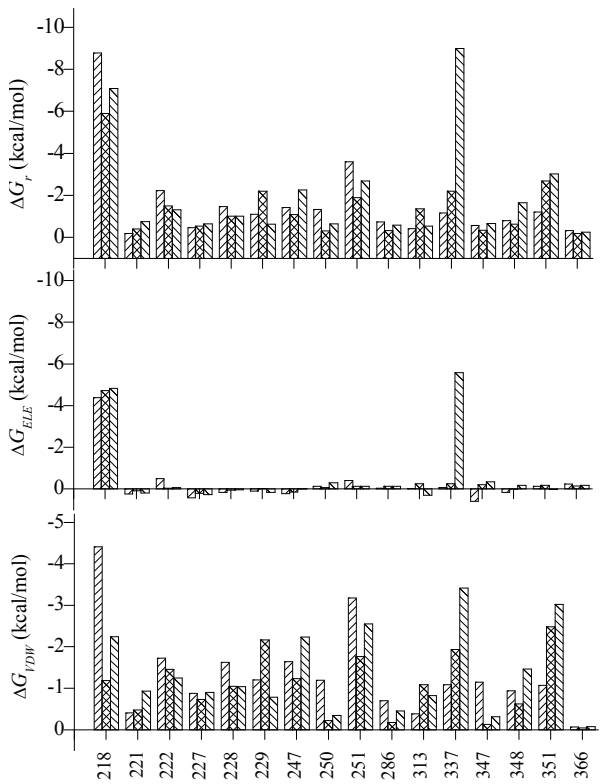


Fig. 3 (a) The binding free energy ΔG_r of the residues located in the top twenty-five high energy contribution in all systems, (b) Energy contribution of the ΔG_{VDW} ($\Delta E_{vdw} + \Delta G_{nonpolar,sol}$), and (c) Energy contribution of ΔG_{ELE} ($\Delta E_{ele} + \Delta G_{ele,sol}$) interaction terms for the residues of the LUC with the pesticides where //// refers to BAY-LUC, ⨯⨯⨯ to MET-LUC, and \\\\ to VEL-LUC

Among three systems, there is no significant distinction in energy contributions for His221, Phe227, Gly228, Leu286, Ser347, and Val366 respectively. However, for Arg337, the energy contribution is much larger in VEL-LUC than that in the other two systems (Fig. S4), which is a consequence of a strong hydrogen bond formed between the O atom of VEL and the H atom of Arg337 (Fig. 2).

Experimental validation of additivity

Alignment of the BAY-LUC, MET-LUC, and VEL-LUC shows that three pesticides occupy the same binding site that is located at the bottom of luciferin pocket (Fig. S3). The hydrophobic pocket enclosed by Asn229, Phe247, Thr251, Ala313 and Ile351 offers favourable van der Waals contacts. In addition, the off-centered parallel displaced π-π stacking interactions between the aromatic rings of BAY and Arg218, VEL and Arg218, as well as MET and Arg337 lead to a high contribution of the vdW energy. From the analysis of binding free energy decomposition, the Arg218 and Arg337 are the crucial residues for three pesticides bound to LUC.

Therefore, a profound hypothesis was proposed that if mixture components bind to the same target site and exhibit the same binding pattern, they are supposed to have the same MOI and the combined toxicity could be predicted by the CA model. The hypothesis was sufficiently verified by the following mixture toxicity testing. The CRCs information of BAY, MET, VEL, and the fifteen binary mixtures determined by the MTA were shown in Table S1 and Table 1. The corresponding CRCs were shown in Fig. 4.

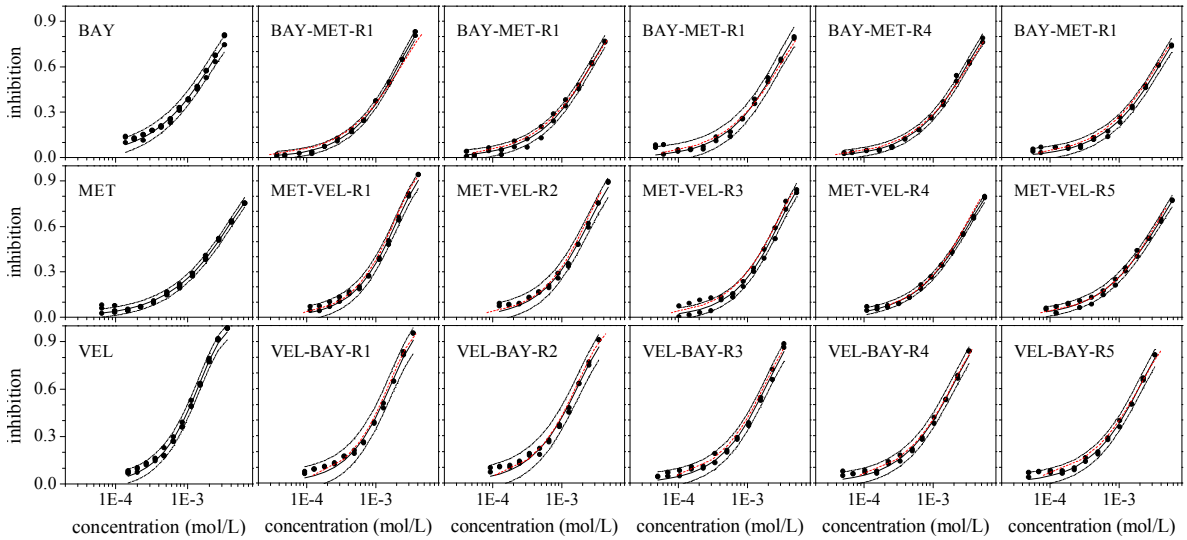


Fig. 4 The concentration-response curves (CRCs) of three chemicals and 15 mixture rays where the scattered points, solid lines, short dash lines (in red), and dash dot lines (in black) represent the experimental values, CRCs fitted, CRCs predicted by CA models, and 95% confidence intervals of observed effect, respectively

The toxicities of 15 binary mixtures perfectly follow the CA model on the whole concentration-response curves. All CRCs predicted by the CA models are located in the 95% confidence intervals, which indicates that there is no interaction between mixture components and thus the CA is suitable to predict the mixture toxicity.

Predicting combined toxicity from ΔG_{mix} of a mixture

In other researches, a linear relationship had been reported between the ΔG and the toxicity of individual pesticides.[49, 50] Taking the EC_{50} value as a toxic index (the lower value of EC_{50} means the higher toxicity), the toxicity order is VEL > BAY > MET in this study (Table S1), which is consistent to ΔG_i. Also, between the ΔG_i of three pesticide-LUC and the inhibition toxicity (lg(EC50)) a linear relationship was observed (Table S2).

More importantly, among the mixture with specific p_i we also found a linear relationship between the ΔG_{mix} and $lgEC_{50}$. Taking Eq. (1) into consideration, all the $lgEC_{50}$s of the pesticides and their mixtures with five series p_i can be expressed as a linear function of the ΔG_i of single pesticide to LUC,

$$\lg EC_{50,BAY-MET} = 0.0390 \times \left(p_{BAY} \times \Delta G_{BAY} + p_{MET} \times \Delta G_{MET}\right) - 2.3511 \quad \text{n=7, } R^2\text{=0.9483} \tag{3}$$

$$\lg EC_{50,MET-VEL} = 0.0307 \times \left(p_{MET} \times \Delta G_{MET} + p_{VEL} \times \Delta G_{VEL}\right) - 2.3943 \quad \text{n=7, } R^2\text{=0.9992} \tag{4}$$

$$\lg EC_{50,VEL-BAY} = 0.0221 \times \left(p_{VEL} \times \Delta G_{VEL} + p_{BAY} \times \Delta G_{BAY}\right) - 2.5383 \quad \text{n=7, } R^2\text{=0.9417} \tag{5}$$

where n is the number of toxicity experiments. The high coefficient of determination (R^2) indicates that the models can be applied to predict the $lgEC_{50}$ value of various mixtures only based on the ΔG_{mix} obtained from Eq. (1). From Fig. 5, all the points were rather uniformly distributed around the regression line of $lgEC_{50}$ *vs*. ΔG.

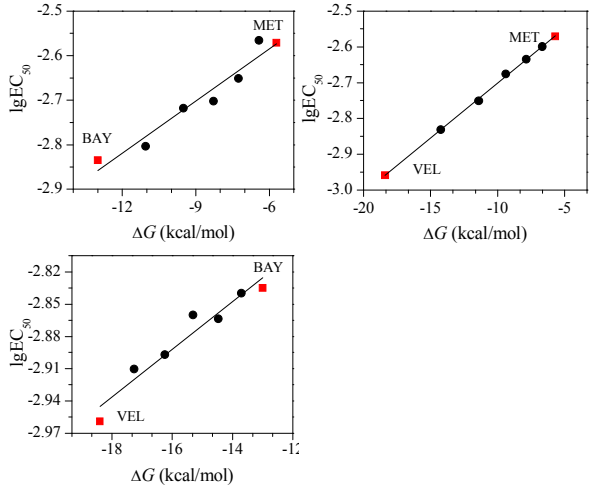


Fig. 5 Plot of experimental toxicity ($lgEC_{50}$) vs. binding free energies from the MM/GBSA using AMBER where ■ (in red) refers to the single pesticide and ● (in black) to the binary mixtures

Conclusions

We presented an *in silico* method to identify the MOI based on ligand-receptor interactions between specific pesticide molecules and luciferase biomacromolecule. The toxicity test results show that all the toxicities of the mixtures combined from the three pesticides which are classified into the same MOI model due to the same binding sites, can be predict perfectly by using the CA model. In addition, the linear models based on the ΔG_{mix} are developed to predict the EC_{50}s of the fifteen binary mixtures composed of different p_i.

Acknowledgments

The authors are especially grateful to the National Natural Science Foundation of China (21377097) and the Specialized Research Fund for the Doctoral Program of Higher Education (No. 20120072110052) for their financial support.

References

1. T. Backhaus and M. Faust, *Environ. Sci. Technol.*, 2012, **46**, 2564-2573.
2. N. Cedergreen, A. M. Christensen, A. Kamper, P. Kudsk, S. K. Mathiassen, J. C. Streibig and H. Sorensen, *Environ. Toxicol. Chem.*, 2008, **27**, 1621-1632.
3. C. A. Laetz, D. H. Baldwin, T. K. Collier, V. Hebert, J. D. Stark and N. L. Scholz, *Environ. Health Perspect.*, 2009, **117**, 348-353.
4. S. Loewe and H. Muischnek, *Naunyn-Schmiedebergs Arch. Exp. Pathol. Pharmakol.*, 1926, **114**, 313-326.
5. C. I. Bliss, *Ann. Appl. Biol.*, 1939, **26**, 585-615.
6. L.-T. Qin, S.-S. Liu, J. Zhang and Q.-F. Xiao, *Toxicology*, 2011, **280**, 164-172.
7. N. Dom, M. Penninck, D. Knapen and R. Blust, *Chemosphere*, 2012, **87**, 742-749.
8. P. L. Piotrowski, B. G. Sumpter, H. V. Malling, J. S. Wassom, P. Y. Lu, R. A. Brothers, G. A. Sega, S. A. Martin and M. Parang, *J. Chem. Inf. Model.*, 2007, **47**, 676-685.
9. S. Spycher, E. Pellegrini and J. Gasteiger, *J. Chem. Inf. Model.*, 2005, **45**, 200-208.
10. S. Brosche and T. Backhaus, *Aquat. Toxicol.*, 2010, **99**, 457-465.
11. N. Thorne, J. Inglese and D. S. Auld, *Chem. Biol.*, 2010, **17**, 646-657.
12. N. Thorne, M. Shen, W. A. Lea, A. Simeonov, S. Lovell, D. S. Auld and J. Inglese, *Chem. Biol.*, 2012, **19**, 1060-1072.
13. L. P. da Silva and J. da Silva, *Chemphyschem*, 2012, **13**, 2257-2262.
14. D. S. Auld, N. Thorne, D.-T. Nguyen and J. Inglese, *ACS Chem. Biol.*, 2008, **3**, 463-470.
15. D. S. Auld, Y.-Q. Zhang, N. T. Southall, G. Rai, M. Landsman, J. MacLure, D. Langevin, C. J. Thomas, C. P. Austin and J. Inglese, *J. Med. Chem.*, 2009, **52**, 1450-1458.
16. D. S. Auld, N. T. Southall, A. Jadhav, R. L. Johnson, D. J. Diller, A. Simeonov, C. P. Austin and J. Inglese, *J. Med. Chem.*, 2008, **51**, 2372-2386.
17. P. G. Cruz, D. S. Auld, P. J. Schultz, S. Lovell, K. P. Battaile, R. MacArthur, M. Shen, G. Tamayo-Castillo, J. Inglese and D. H. Sherman, *Chem. Biol.*, 2011, **18**, 1442-1452.
18. D. S. Auld, S. Lovell, N. Thorne, W. A. Lea, D. J. Maloney, M. Shen, G. Rai, K. P. Battaile, C. J. Thomas, A. Simeonov, R. P. Hanzlik and J. Inglese, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 4878-4883.
19. G. Rastelli, A. Del Rio, G. Degliesposti and M. Sgobba, *J. Comput. Chem.*, 2010, **31**, 797-810.
20. H.-L. Ge, S.-S. Liu, X.-W. Zhu, H.-L. Liu and L.-J. Wang, *Environ. Sci. Technol.*, 2011, **45**, 1623-1629.
21. E. Conti, N. P. Franks and P. Brick, *Structure*, 1996, **4**, 287-298.
22. Y. Mao, *Protein Eng., Des. Sel.*, 2011, **24**, 341-349.
23. A. Szarecka, Y. Xu and P. Tang, *Biophys. J.*, 2007, **93**, 1895-1905.
24. A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren, *Acta Crystallogr. Sect. D: Biol. Crystallogr.*, 1998, **54**, 905-921.
25. R. A. Laskowski and M. B. Swindells, *J. Chem. Inf. Model.*, 2011, **51**, 2778-2786.
26. S. Shanmugasundararaj, S. Lehle, H. I. Yamodo, S. S. Husain, C. Tseng, K. Nguyen, G. H. Addona and K. W. Miller, *PLoS One*, 2012, **7**.
27. J. Gasteiger and M. Marsili, *Tetrahedron*, 1980, **36**, 3219-3228.
28. O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455-461.
29. L. K. Wolf, *Chem. Eng. News*, 2009, **87**.
30. R. Salomon-Ferrer, D. A. Case and R. C. Walker, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2013, **3**, 198-210.
31. J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
32. V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, **65**, 712-725.
33. E. Pauwels, D. Claeys, J. C. Martins, M. Waroquier, G. Bifulco, V. Van Speybroeck and A. Madder, *RSC Adv.*, 2013, **3**, 3925-3938.
34. D. Zanuy, B. Teixeira-Dias, L. J. del Valle, J. Poater, M. Sola and C. Aleman, *RSC Adv.*, 2013, **3**, 2639-2649.
35. F. Chen, S. S. Liu and X. T. Duan, *Acta Chim. Sinica*, 2013, **71**, 1035-1040.
36. P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. H. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, *Acc. Chem. Res.*, 2000, **33**, 889-897.
37. J. Kottalam and D. A. Case, *Biopolymers*, 1990, **29**, 1409-1421.
38. Y. Yang, J. Qin, H. Liu and X. Yao, *J. Chem. Inf. Model.*, 2011, **51**, 680-692.
39. V. Zoete, M. Meuwly and M. Karplus, *Proteins: Struct., Funct., Bioinf.*, 2005, **61**, 79-93.
40. Y.-X. Shao, P. Zhao, Z. Li, M. Liu, P. Liu, M. Huang and H.-B. Luo, *Eur. Biophys. J. Biophys. Lett.*, 2012, **41**, 297-306.
41. X. W. Zhu, S. S. Liu, H. L. Ge and Y. Liu, *Water Res.*, 2009, **43**, 1731-1739.
42. R.-N. Dou, S.-S. Liu, L.-Y. Mo, H.-L. Liu and F.-C. Deng, *Environ. Sci. Pollut. Res.*, 2011, **18**, 734-742.
43. E. Silva, N. Rajapakse, M. Scholze, T. Backhaus, S. Ermler and A. Kortenkamp, *Toxicol. Sci.*, 2011, **122**, 383-394.
44. W. Khuntawee, T. Rungrotmongkol and S. Hannongbua, *J. Chem. Inf. Model.*, 2012, **52**, 76-83.
45. J. Guo, X. Wang, H. Sun, H. Liu and X. Yao, *J. Mol. Model.*, 2012, **18**, 1421-1430.
46. X. Y. Xie, W. J. Lu and X. G. Chen, *J. Hazard. Mater.*, 2013, **248**, 347-354.
47. Y. Teng, R. T. Liu, C. Li, Q. Xia and P. J. Zhang, *J. Hazard. Mater.*, 2011, **190**, 574-581.
48. P. D. Ross and S. Subramanian, *Biochemistry*, 1981, **20**, 3096-3102.
49. D. Liang, J. Gao, Y. Cheng, W. Cui, H. Zhang and M. Ji, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 6630-6635.

50. H. Zhang, Y. Yao, H. Yang, X. Wang, Z. Kang, Y. Li, G. Li and Y. Wang, *Insect Biochem. Mol. Biol.*, 2012, **42**, 583-595.

Appendix A: Supplementary Information

Predicting the mixture effects of three pesticides by integrating the molecular simulation with the concentration addition

Fu Chen [a], Shu-Shen Liu [a,b,*], Xin-Tian Duan [a], and Qian-Fen Xiao [a,b]

[a] Key Laboratory of Yangtze River Water Environment, Ministry of Education, College of Environmental Science and Engineering, Tongji University, Shanghai 200092, PR China

[b] State Key Laboratory of Pollution Control and Resource Reuse, College of Environmental Science and Engineering, Tongji University, Shanghai 200092, PR China

[*] Corresponding author: Prof. Shu-Shen Liu

College of Environmental Science and Engineering, Tongji University,

1239 Siping Road, Shanghai 200092, P. R. China

Tel: 86-021-65982767

E-mail address: ssliuhl@263.net (S.-S. Liu)

Sequential docking protocol

Sequential docking protocol was carried out to predict the locations of two ligands within a given binding site. In the first step of the sequential docking, pesticide A was individually docked into binding site with the same parameters described in Molecular docking and molecular dynamics simulation section. In the second step, pesticide B was docked into the complex of LUC with bound pesticide A, whose docking box expanded into 16×20×20 Å. The last step was selecting the pose of pesticide B with the most negative score. The molecular modelling of A-B-LUC was built. The results were shown in Fig. S6. As can be seen from Fig. S6, the former BAY occupies the LH_2 pocket, while the latter pesticide (BAY, MET, or VEL) just occupies the pocket of AMP that is adjacent to LH_2 pocket but separate. This phenomenon is found in other ternary complex. Furthermore, comparing the two complexes having the same second pesticide, it is clear that the RMSD of second pesticide is small. It is deduced that the former bound ligand does not influence the binding site of the second ligand.

Molecular dynamics procedure

The structural parameters of pesticide ligand and LUC protein were prepared by using antechamber module and tleap module in AMBER 12 software package. The atom types and atomic charges were assigned by the AM1-BCC[1] model in the antechamber module and the missing bonded parameters were gained from the generalized AMBER force field (GAFF)[2, 3] for the pesticide ligand. The ff99SB force field[4] was selected for the LUC protein. The tleap module was used to correct all missing hydrogen atoms of both the protein and pesticide. The pesticide-LUC complex was solvated by the TIP3P waters, with a minimum distance of 8.5 Å from the complex surface. Thus, the total number of atoms were 42672, 42665, and 42680 for the BAY-LUC, MET-LUC, and VEL-LUC systems, respectively. All three systems having a total charge of -7 were neutralized by adding seven Na^+ ions.

MD was conducted by the standard procedure, which comprises energy-minimization, gradual heating of the systems, and isothermal isobaric ensemble (NPT) molecular dynamics. To reduce bad steric interaction, the solvated system was minimized by three steps[3]: The steepest descent (2000 steps) and conjugate gradient (1000 steps) minimizations were firstly carried out on the water molecules alone, then on the backbone of complex and water system, and finally on the whole system. The equilibration phase and production phase were obtained by using the PMEMD[5] module in AMBER. The periodic boundary with NPT (constant particle, pressure, and temperature) ensemble at 1 atm was applied. The SHAKE method was used to provide an integration time step of 2 fs while keeping all bonds to the hydrogen atoms rigid. The cutoff distance for the long-range van der Waals was set as 10 Å. The long-range Coulombic interactions were handled using the particle mesh Ewald (PME) method. The whole system was gradually heated from 0 to 300 K in 50 ps and subsequently simulated at 300 K for equilibration and production phases. The whole system was equilibrated for 4 ns, and the MDs were continued for another 4 ns. One hundred snapshots of the simulated structures within the last 1 ns with a step of 10 ps were sampled.

Simulation system was monitored through the convergences of energy and temperature[3]. To verify the system stability and measure protein dynamics, the ptraj module was used to analyze the root mean-square displacements (RMSD) between the trajectory structures and the first snapshot structure in 1th ns trajectory.

Binding free energy calculation

The binding free energy (ΔG_i) of single pesticide bound to LUC was computed from the free energy of the complex ($G_{complex}$), protein ($G_{protein}$), and ligand (G_{ligand}),

$$\Delta G_i = G_{complex} - \left(G_{protein} + G_{ligand} \right) \quad (1)$$

On the other hand, ΔG_i can be expressed as follows,

$$\Delta G_i = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{sol} - T\Delta S \quad (2)$$

where ΔH is the enthalpy, ΔE_{MM} is the molecular mechanics free energy, ΔG_{sol} is the solvation free energy, and $T\Delta S$ represents the entropy term.

The molecular mechanics free energy was calculated as follows:

$$\Delta E_{MM} = \Delta E_{ele} + \Delta E_{vdw} \quad (3)$$

where ΔE_{ele} and ΔE_{vdw} represent the Coulomb and van der Waals interactions, respectively.

The solvation free energy was composed of polar and nonpolar components:

$$\Delta G_{sol} = \Delta G_{ele,sol} + \Delta G_{nonpolar,sol} \quad (4)$$

where $\Delta G_{ele,sol}$ is the polar contribution to solvation and $\Delta G_{nonpolar,sol}$ is the nonpolar solvation term. The $\Delta G_{ele,sol}$ was calculated by using the generalized Born (GB) model. The $\Delta G_{nonpolar,sol}$ was computed based on solvent accessible surface area (SASA). The bigger absolute value of ΔG negative value means the more binding ability.

Entropy contributions ($T\Delta S$) arising from changes in the translational, rotational and vibrational degrees of freedom was

calculated using normal-mode analysis[6] by the nmode program in AMBER 12. The normal-mode analysis is high computationally demanding, so that $T\Delta S$ was averaged over only 20 snapshots of the last 1 ns MD trajectory.[7, 8] Before the calculation, each snapshot was optimized for 50000 steps using a distance-dependent dielectric of $4r_{ij}$ (r_{ij} is the distance between atoms i and j) until the root-mean-square deviation of the gradient vector was less than 0.0001 kcal mol^{-1} Å^{-2}.[8]

Ligand-residue interaction decomposition

The interaction between each residue of luciferase and pesticide were computed using the MM/GBSA approach applied in AMBER 12. This decomposition was performed only for molecular mechanics and solvation energies but not for entropies.[9] The binding free energy of each complex was decomposed per residue including four terms: van der Waals contribution (ΔE_{vdW}), electrostatic contribution (ΔE_{ele}), polar solvation contribution ($\Delta G_{ele,sol}$), and nonpolar solvation contribution ($\Delta G_{nonpol,sol}$):

$$\Delta G_r = \Delta E_{vdW} + \Delta E_{ele} + \Delta G_{ele,sol} + \Delta G_{nonpol,sol} \quad (5)$$

where ΔE_{ele} and ΔE_{vdw} represent the Coulomb and van der Waals interactions, respectively. $\Delta G_{ele,sol}$ is the polar contribution to solvation and $\Delta G_{nonpolar,sol}$ is the nonpolar solvation term. ΔG_r is pesticide-residue pair energy between each pesticide and each individual residue

The microplate toxicity analysis procedure.

The toxicity of a pesticide or mixture was expressed as a percentage inhibition (E or x), which was calculated as follows:

$$E = x = \left(1 - \frac{L}{L_0}\right) \times 100\% \quad (6)$$

where L_0 is an average RLU of the controls (12 parallels), L is an average RLU of the treatments (3 parallels).

To quantitatively describe the toxicities of various concentrations, the observed concentration-response datas were fitted to nonlinear function, Weibull, called concentration response curve fitting (CRC)[10],

$$E = 1 - \exp\left(-\exp\left(\alpha + \beta \log_{10}(c)\right)\right) \quad (7)$$

where α and β are the parameters to be estimated, c is the concentration of test pesticide or mixture. The regression analysis was performed using nonlinear least-squares fit. The coefficient of determination (R^2) and the root-mean-square error (RMSE) were essential to evaluate the fitting. As a quantitative measure of the uncertainty, the 95% observation-based confidence interval was also determined.[11]

References

1. M. Han and J. Z. H. Zhang, *J. Chem. Inf. Model.*, 2010, **50**, 136-145.
2. J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
3. W. Khuntawee, T. Rungrotmongkol and S. Hannongbua, *J. Chem. Inf. Model.*, 2012, **52**, 76-83.
4. V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, **65**, 712-725.
5. D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J. Comput. Chem.*, 2005, **26**, 1668-1688.
6. J. Kottalam and D. A. Case, *Biopolymers*, 1990, **29**, 1409-1421.
7. R. Cao, Y. Jin and D. Xu, *J. Phys. Chem. B*, 2012, **116**, 6087-6096.
8. H. Zhang, Y. Yao, H. Yang, X. Wang, Z. Kang, Y. Li, G. Li and Y. Wang, *Insect Biochem. Mol. Biol.*, 2012, **42**, 583-595.
9. Y. Yang, J. Qin, H. Liu and X. Yao, *J. Chem. Inf. Model.*, 2011, **51**, 680-692.
10. M. Scholze, W. Boedeker, M. Faust, T. Backhaus, R. Altenburger and L. H. Grimme, *Environ. Toxicol. Chem.*, 2001, **20**, 448-457.
11. X. W. Zhu, S. S. Liu, H. L. Ge and Y. Liu, *Water Res.*, 2009, **43**, 1731-1739.

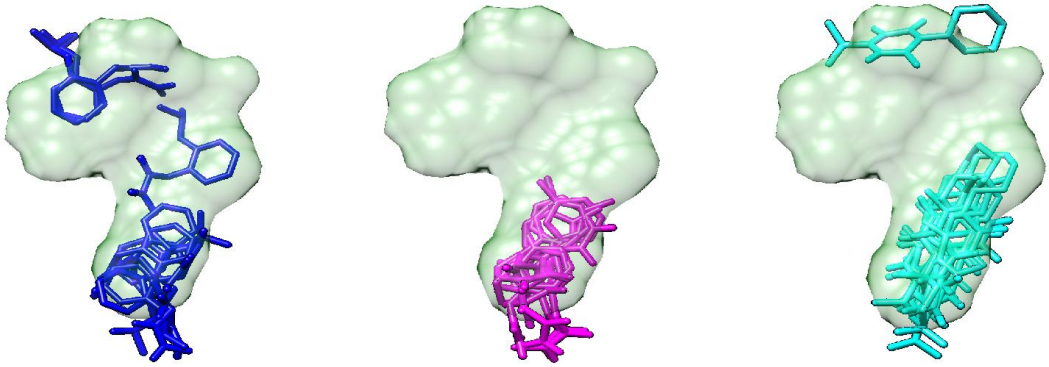
Table S1. The fitted concentration-response curve model parameters (α and β), statistics (RMSE and R^2), EC_{50} (10^{-3} mol/L), and binding free energy (ΔG_i, kcal/mol) of three pesticides

Chemicals	α	β	RMSE	R^2	EC_{50}	ΔG_i	EC_{50}[a]
BAY	5.36	2.02	0.019	0.993	1.463	-13.0	1.442
MET	4.75	1.99	0.011	0.998	2.692	-5.7	2.682
VEL	8.57	3.02	0.017	0.997	1.100	-18.4	1.110

[a]: the average value of the EC_{50}s estimated from two linear models (Eqs. 3 - 5)

Table S2. Binding free energy (kcal/mol) and its components for the three pesticide-LUC systems

Component	BAY-LUC	MET-LUC	VEL-LUC
ΔE_{vdw}	-37.6	-26.7	-39.2
ΔE_{ele}	-11.3	-13.6	-29.0
ΔE_{MM}	-48.9	-40.3	-68.2
$\Delta G_{ele,sol}$	20.3	20.5	33.7
$\Delta G_{nonpolar,sol}$	-4.4	-3.9	-5.0
ΔG_{sol}	15.8	16.6	28.7
$\Delta E_{vdw}+\Delta G_{nonpolar,sol}$	-42.0	-30.6	-44.2
$\Delta E_{ele}+\Delta G_{ele,sol}$	9.0	6.9	4.7
ΔH_i	-33.1	-23.7	-39.5
$T\Delta S_{Translational}$	-12.4	-12.2	-12.6
$T\Delta S_{Rotational}$	-9.6	-9.2	-9.8
$T\Delta S_{Vibrational}$	1.9	3.4	1.4
$T\Delta S$	-20.1	-18.0	-21.1
ΔG_i	-13.0	-5.7	-18.4
EC_{50}(mol/L)	1.463E-03	2.692E-03	1.100E-03



BAY-LUC MET-LUC VEL-LUC

Fig. S1. The 10 best scoring poses of pesticides bound to LUC.

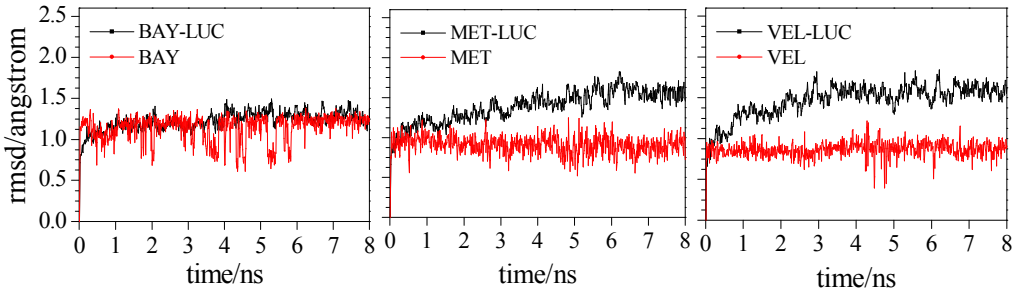


Fig. S2. Plots of root mean-square deviation (RMSD) vs. time for the simulated systems of BAY-LUC, MET- LUC, and VEL- LUC

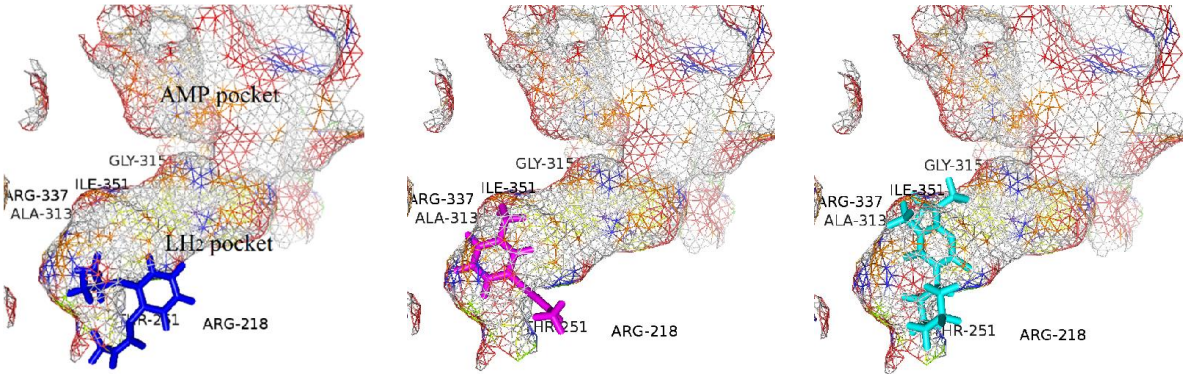


Fig. S3. The binding sites of three pesticides in the LUC pocket where the stick shown in blue, magenta, and cyan represent BAY, MET, and VEL, respectively.

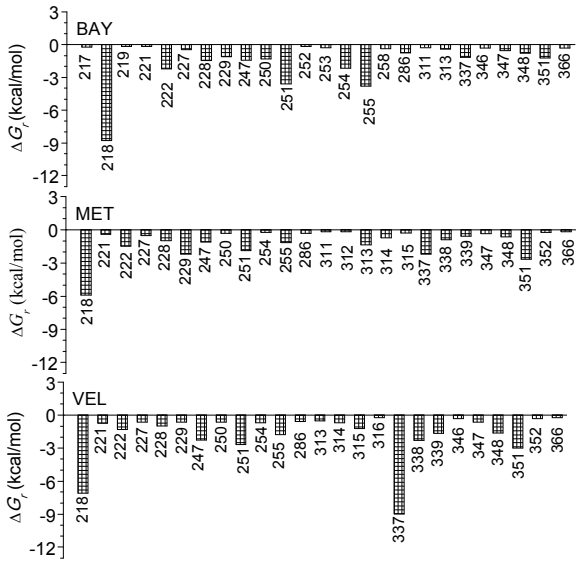


Fig. S4. Decomposition of the binding free energy on a residue-based (ΔG_r) for key residues in the pesticide-LUC.

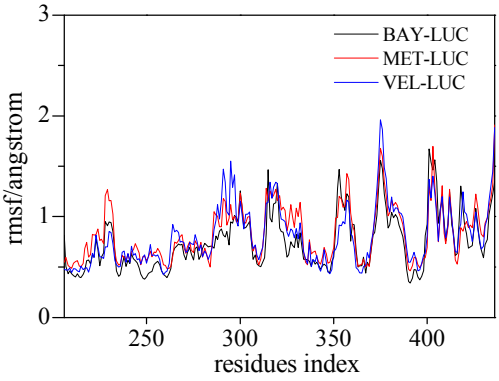


Fig. S5. Plots of root mean-square fluctuation (RMSF) vs. residue index for the simulated systems.

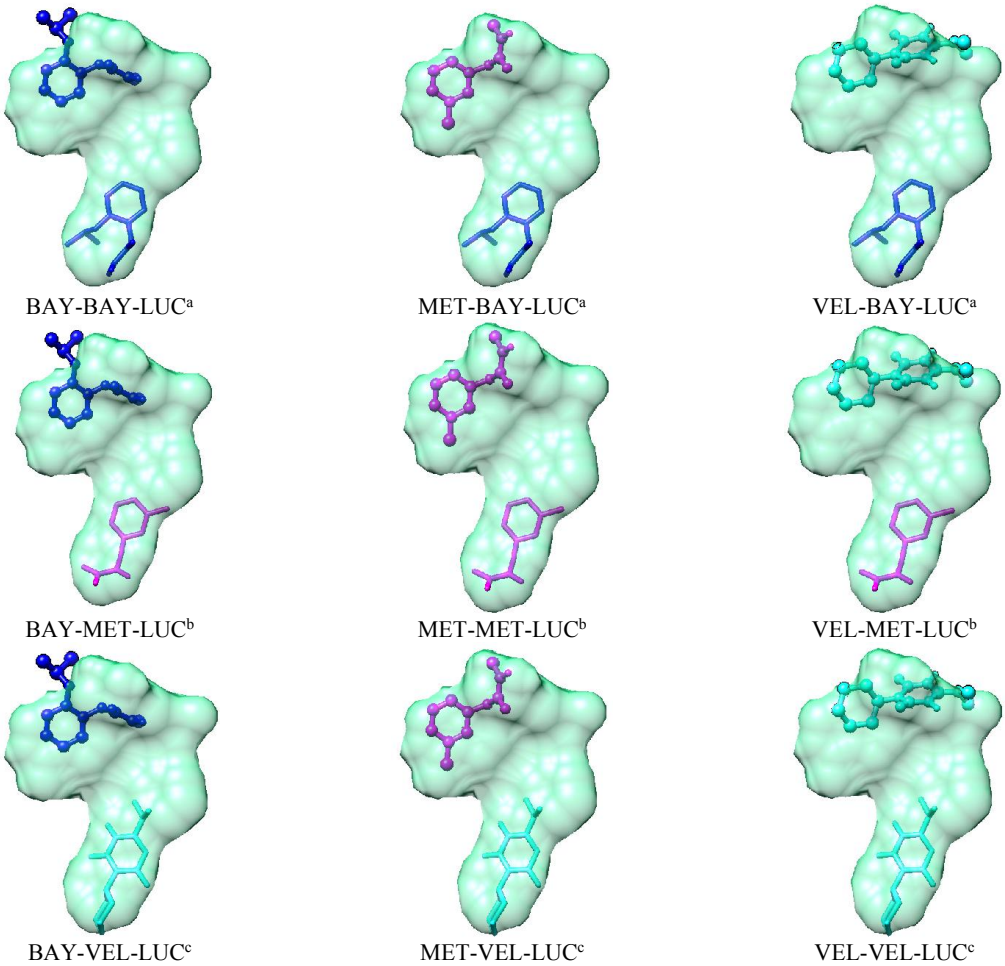


Fig. S6. sequential simulations. The former pesticide is shown as stick models, the latter pesticide is shown as ball and stick models. [a] sequential docking based on BAY-LUC, [b] sequential docking based on MET-LUC, [c] sequential docking based on VEL-LUC.