Environmental Toxicology and Pharmacology 40 (2015) 886–894

Contents lists available at ScienceDirect

Environmental Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/etap

QSAR analyses of organophosphates for insecticidal activity and its *in-silico* validation using molecular docking study

Ravi Ranjan Kumar Niraj, Vandana Saini, Ajit Kumar*

Toxicology and Computational Biology Group, Centre for Bioinformatics, M.D. University, Rohtak 124001, Haryana, India

ARTICLE INFO

Article history:
Received 3 July 2015
Received in revised form
26 September 2015
Accepted 28 September 2015

Keywords:
Combinatorial library
CoMFA
EPI Suite
Homology modelling

ABSTRACT

The present work was carried out to design and develop novel QSAR models using 2D-QSAR and 3D-QSAR with CoMFA methodology for prediction of insecticidal activity of organophosphate (OP) molecules. The models were validated on an entirely different external dataset of in-house generated combinatorial library of OPs, by completely different computational approach of molecular docking against the target AChE protein of *Musca domestica*. The dock scores were observed to be in good correlation with 2D-QSAR and 3D-QSAR with CoMFA predicted activities and had the correlation coefficients (r^2) of −0.62 and −0.63, respectively. The activities predicted by 2D-QSAR and 3D-QSAR with CoMFA were also observed to be highly correlated with $r^2 = 0.82$. Also, the combinatorial library molecules were screened for toxicity in non-target organisms and degradability using USEPA-EPI Suite. The work was first step towards computer aided design and development of novel OP pesticide candidates with good insecticidal property but lower toxicity in non-targeted organisms and having biodegradation potential.

1. Introduction

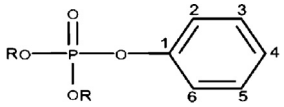
Organophosphate (OP) compounds are among the most widely used pest control agents. The usage of OPs has increased dramatically since the worldwide ban or restriction on persistent organochlorines like DDT and BHC, during 1970s. Only 1% of about 4 million tonnes of chemical pesticides, applied annually, reaches the target pest, while the rest affect non-target organisms, causing environmental and health menace (Gavrilescu, 2005). The insecticidal activity of OPs and their toxicity to non-target organisms including mammals is majorly attributed to their ability to bind, phosphorylate and inactivate the enzyme acetylcholine esterase, thus making the later unavailable for normal catalysis at neural junctions, *i.e.*, the hydrolysis of neurotransmitter acetylcholine (ACh) (Fukuto, 1990; Richardson, 1995; Wilson and Tisdell, 2001; Čolović et al., 2013; Makhaeva et al., 2014). The resistance in target organisms for pesticides is another cause of concern, as it leads to elevated applications by the end users (Fukuto, 1990; Zalom et al., 2005). Several cases of insecticide resistance to OPs in target pests have been reported (Molina and Figueroa, 2009; Osta et al., 2012; Temeyer et al., 2014; Zhao et al., 2014). These facts have always paved the way for design and development of new pesticides, including OPs.

Computational chemistry has always been in the central stage of modern days molecule design including drugs and pesticides, as they provide a cheaper and rapid alternative to *in vitro* and *in vivo* assays, so as to follow the '3Rs' policy of reducing, refining and replacing animal testing for primitive studies (Price and Watkins, 2003). Quantitative Structure Activity Relationship (QSAR) methods are the forerunners of computational tools in drug and pesticide designing (Saini and Kumar, 2014a). QSAR analysis results in generation of statistical models that are regarded as very efficient tool in both drug discovery and environmental toxicology for prediction and classification of biological activities of new and untested compounds (Perkins et al., 2003). Here, we report about QSAR studies with 2D- and CoMFA insights for designing new non-sulfur aromatic OP molecules having good insecticidal activity, lower toxicity in non-target organisms and biodegradation potential. The work also involved generation of combinatorial library of OP molecules analogous to the molecules selected for QSAR studies. These molecules were then evaluated for their insecticidal activity using generated QSAR models and for their toxicity and degradability using EPI suite (USEPA, 2013). The present study also reports about a novel method to validate, *in-silico*, the generated QSAR models using molecular docking simulation as reported earlier by us (Saini and Kumar, 2014b).

* Corresponding author.
E-mail address: akumar.cbt.mdu@gmail.com (A. Kumar).

http://dx.doi.org/10.1016/j.etap.2015.09.021
1382-6689/

Table 1
Data set of OP analogues taken for QSAR analyses showing functional groups at two different substitution points. Number prefix at point X refers to position 3 (Meta) or position 4 (Para) of template.



Molecule ID	Functional group at substitution point		Experimental pLD_{50} against *Musca nebulo*	Predicted pLD_{50} (2D-QSAR)	Predicted pLD_{50} (3D-QSAR)
	R	X			
OP 01	CH_3	H	2.75	2.68	2.58
OP 02	CH_3	3-CH_3	2	2.67	1.98
OP 03	CH_3	4-CH_3	1.99	2.68	1.96
OP 04	CH_3	4-OCH_3	2	1.98	2.24
OP 05	CH_3	3-Cl	2.1	2.78	2.84
OP 06	CH_3	4-Cl	2.6	2.78	2.47
OP 07	CH_3	3-Br	4	4.11	3.38
OP 08	CH_3	4-Br	3.53	4.11	3.32
OP 09	CH_3	3-CN	4.99	4.49	5.02
OP 10	CH_3	4-CN	4.84	4.49	4.85
OP 11	CH_3	3-NO_2	4.9	5.09	4.91
OP 12	CH_3	4-NO_2	5.1	4.99	4.97
OP 13	C_2H_5	H	3.2	2.68	3.26
OP 14	C_2H_5	4-CH_3	3	2.68	2.71
OP 15	C_2H_5	3-Cl	3.8	2.76	3.99
OP 16	C_2H_5	4-Cl	3.72	2.76	3.99
OP 17	C_2H_5	3-Br	4.11	3.93	2.99
OP 18	C_2H_5	4-Br	4.06	3.93	4.01
OP 19	C_2H_5	3-CN	5	4.49	5.04
OP 20	C_2H_5	4-CN	5.1	4.5	4.95
OP 21	C_2H_5	3-NO_2	5.1	5.09	4.99
OP 22	C_2H_5	4-NO_2	5.2	4.97	4.91
OP 23	C_2H_5	2,4-Cl	4.3	3.14	4.35
OP 24	C_2H_5	2,5-Cl	4.1	3.13	4.45
OP 25	C_4H_9	H	2.5	2.68	2.31
OP 26	C_4H_9	3-CH_3	2	2.67	1.98
OP 27	C_4H_9	4-CH_3	2.1	2.678	2.17
OP 28	C_4H_9	4-OCH_3	2.1	1.88	3.95
OP 29	C_4H_9	3-Cl	2.8	2.75	2.72
OP 30	C_4H_9	4-Cl	2.5	2.74	2.32
OP 31	C_4H_9	4-Br	2.95	3.68	2.87
OP 32	C_4H_9	3-CN	4	4.49	4.62
OP 33	C_4H_9	4-CN	4.01	4.5	4.36
OP 34	C_4H_9	3-NO_2	4.21	5.09	3.81
OP 35	C_4H_9	4-NO_2	4.38	5.00	4.62

2. Materials and methods

2.1. Generation of 2D-QSAR

For the generation of QSAR models for insecticidal activity of OP pesticide, the activity data was collected and analyzed after extensive literature search. A total of 35 OPs analogues were taken for the QSAR study and represented a set of molecules, reported earlier (Gandhe et al., 1990), with different functional groups at two positions -R and -X on the basic templates, along with their corresponding acute toxicity data (log LD_{50}) of these compounds for Housefly (*Musca nebulo* L.) as mentioned in Table 1. The biological activity of selected compounds covered a range of more than 3 log units (pLD_{50} = 1.99–5.2) and hence very suitable for QSAR studies. These data were used as a dependent variable of QSAR Pro module of Vlife MD Suite. A total of 339 descriptors were calculated for each molecule that belonged to physico-chemical type descriptor (239) and atom type count descriptor (100). The descriptors were manually deleted from further analysis if their values were either zero or uniform for all the molecules. Descriptors having less variability were also deleted by 'Remove invariable columns' tab of 2D-QSAR in Vlife QSAR to reach a final descriptor count of 144. The curated dataset was then subjected to 2D-QSAR model generation. Out of 35 molecules of dataset, 24 molecules were taken as training set while 11 molecules (OP01, OP03, OP12, OP16, OP19, OP24, OP25, OP29, OP30, OP31 and OP34) were selected as test set. The training set (70%) selection was done on random basis. The 2D-QSAR model was generated by using different default methods of QSAR Pro of Vlife MDS namely "stepwise forward" and "multiple regressions" methods. The number of variables in final equation was screened from 3 (minimum) to 6 (maximum). The model generated with 4 variables in final equation was observed to be most statistically significant and hence considered for generating QSAR equation. Finally QSAR model with highest q^2 and r^2 were selected keeping in view optimum values of other significant statistical parameters *viz*. pred_r^2, *F*-test value, Y-randomization *Z*-score (Tropsha et al., 2003). The variable selection and model building method wizard summary used for 2D-QSAR study was as per Table 2.

2.2. Generation of 3D-QSAR with CoMFA

For CoMFA studies, same set of 35 OPs, as used for 2D-QSAR analysis, was selected (Table 1) while SYBYL-X 2.1 package (SYBYL-X) was used for generation of models. The entire set of molecules used for CoMFA analysis were aligned using a common sub-structure because the results of CoMFA are highly sensitive to the alignment rules, orientation of the aligned molecules, lattice shifting step size and probe atom type (Saini and Kumar, 2014b). The 'database align' module of SYBYL was used for the purpose using the energy minimized structure of most bioactive OP molecule (OP22) as template