CODEN IEJMAT *Internet Electronic Journal of Molecular Design* **2005**, *4*, 355–366 ISSN 1538–6414
BioChem Press http://www.biochempress.com

Inter*net* Electronic Journal of
Molecular Design

May 2005, Volume 4, Number 5, Pages 355–366

Editor: Ovidiu Ivanciuc

Proceedings of the Internet Electronic Conference of Molecular Design 2004
IECMD 2004, November 29 – December 12, 2004

A QSAR Approach to Physico–Chemical Data for Organophosphates with Special Focus on Known and Potential Nerve Agents

Lars Carlsen

Awareness Center, Hyldeholm 4, Veddelev, DK–4000 Roskilde, Denmark

Received: October 20, 2004; Revised: January 26, 2005; Accepted: March 6, 2005; Published: May 31, 2005

Citation of the article:
L. Carlsen, A QSAR Approach to Physico–Chemical Data for Organophosphates with Special Focus on Known and Potential Nerve Agents, *Internet Electron. J. Mol. Des.* **2005**, *4*, 355–366, http://www.biochempress.com.

Inter*net* Electronic Journal of Molecular Design

BioChem Press
http://www.biochempress.com

A QSAR Approach to Physico–Chemical Data for Organophosphates with Special Focus on Known and Potential Nerve Agents [#]

Lars Carlsen *

Awareness Center, Hyldeholm 4, Veddelev, DK–4000 Roskilde, Denmark

Received: October 20, 2004; Revised: January 26, 2005; Accepted: March 6, 2005; Published: May 31, 2005

Internet Electron. J. Mol. Des. **2005**, *4* (5), 355–366

Abstract

Motivation. Nerve agents like Sarin, Tabun, Soman, VX, Amiton etc. are highly toxic organophosphates (OPs) that exert their toxic effect by inhibition of acetylcholine esterase. These compounds have received considerable interest due their inherent nature as weapons of mass destruction. Since these compounds have been developed for military purposes, data is typically classified material and thus only scarcely available. QSAR modeling is an obvious possibility in order to remedy the lack of data availability. However, a wide variety of structurally related OP insecticides are well known and well characterized. "Noise–deficient" QSARs, *i.e.*, a QSAR model where the natural variation in both the experimental data and the primary models data has been suppressed in a subsequent modeling step, for physico–chemical properties of nerve agents are based on the use of the EPI Suite, a general QSAR model from the US EPA. Partial order ranking is an important tool to establish an identity for nerve agents relative to well–known OP insecticides. The development of a simple QSAR model for toxicological properties was unsuccessful.

Method. The results described in the paper are obtained using QSAR modeling based on the EPI Suite in comparison with partial order ranking. The concept of "noise deficient" QSARs is introduced.

Results. "Noise deficient" QSARs can be obtained using EPI Suite generated data in combination with experimental data for the test set, the data subsequently being applied in the ranking exercise. In the present study it is shown that to a certain extent selected insecticides may act as substitutes for nerve agents in preliminary experimental studies.

Conclusions. The paper suggests that experimentally well–characterized compounds may be selected as substitutes for highly toxic compounds for preliminary experimental studies of the environmental behavior of the latter.

Keywords. Noise–deficient QSARs; partial order ranking; Hasse diagrams; organo–phosphates; nerve agents.

1 INTRODUCTION

Organophosphates (OPs) are in general toxic substances that exert their toxic effect by inhibition of acetylcholinesterase. The so–called nerve agents like Sarin, Tabun, Soman, and VX have

[#] Presented in part at the Internet Electronic Conference of Molecular Design 2004, IECMD 2004.

* Correspondence author; E–mail: LC@AwarenessCenter.dk.

received considerable interest due their potential use as weapons of mass destruction.

According to the 'Convention for the Prohibition of the Development, Production [1], Stockpiling and Use of Chemical Weapons and their Destruction' major emphasis is given to declaration and destruction of existing stockpiles of chemical weapons as well as of chemical weapons production facilities. The Convention covers both the destruction of chemical weapons stockpiles and destruction or conversion of chemical weapons production facilities. In both cases there may be a significant risk to environment and human health. Since these compounds have been developed for military purposes data on these compounds are often regarded as classified. However, an excellent review on the sources, fate and toxicity of chemical warfare agent degradation products is available [2]. It should be noted hat this review partly is based on papers and reports that not necessarily is easily obtained.

When assessing environmental or human health effects of these compounds, QSAR modeling can to a certain extent remedy the apparent lack of data. Thus, in the present study physico–chemical properties of nerve agents have been estimated using QSAR models based on the EPI Suite [3]. It should be emphasized that the environmental processes of these substances are no different from other substances. However, for environmental studies the extreme toxicity of the nerve agents obviously must be considered.

Within the frame of the present study, it has not been possible to derive a simple QSAR model for OP toxicity. However, various studies have been devoted to models to the toxicity of OPs [4–7].

The present paper emphasizes the development of "noise–deficient" QSARs that will lead to physico–chemical end–points that subsequently can be used as descriptors in a partial order ranking of OPs with focus on selected nerve agents. The main objective is to find experimentally well–characterized compounds, exhibiting significantly lower toxicity than the nerve agents that can be used as substitutes in experimental studies of the environmental behavior of the nerve agents. Obviously, unique structural elements of the nerve agents may well cause significant differences in the environmental fate of the nerve agents and the OP insecticides. Nevertheless, some insecticides may mimic the nerve agents by, *e.g.*, exhibiting identical volatilization behavior, while others may display similar rates of biodegradation. Based on a partial order ranking, taking several parameters into account simultaneously, it appears possible to give the single nerve agent an identity by comparing to structurally related OP insecticides. These substitutes will, based on an overall viewpoint exhibit analogous environmental characteristics as the nerve agents, and thus be models. Thus, such substitutes may well be used for preliminary experimental studies on the environmental fate of nerve agents, and subsequently constitute the basis of selection of a limited number of necessary experimental studies applied the actual nerve agents in order to perform risk assessment, *e.g.*, in relation to the demilitarization activities.

2 MATERIALS AND METHODS

In the present study the end–points are generated through QSAR modeling, the EPI Suite being the primary tool [3]. The EPI Suite comprises a variety of submodules to estimate various physico–chemical parameters as well as ECOSAR to derive toxicity parameters. The models are based on a group contribution approach.

The EPI generated (logarithmic) values for water solubility (log *Sol*), octanol–water partitioning (log K_{OW}), vapor pressure (log *VP*) and Henry's Law constants (log *HLC*) are further treated, *i.e.*, new linear "noise–deficient" QSAR models are built by estimating the relationships between the EPI generated data and available experimental data for up to 65 OP insecticides, the general formula for the end–points, D_i, to be used being:

$$D_i = a_i \times D_{EPI} + b_i \tag{1}$$

D_{EPI} being the EPI generated end–point value and a_i and b_i being constants. The log K_{OW} values generated in this way are subsequently used to generate log *BCF* values according to the Connell formula [8].

$$\log BCF = 6.9\times10^{-3}\times(\log K_{ow})^4 - 1.85\times10^{-1}\times(\log K_{ow})^3 + 1.55\times(\log K_{ow})^2 - 4.18\times\log K_{ow} + 4.72 \tag{2}$$

The model was somewhat modified. Thus, a linear decrease of log *BCF* with log K_{OW} was assumed in the range $1 < \log K_{OW} < 2.33$, the log $BCF = 0.5$ for $\log K_{OW} \leq 1$, the latter value being in accordance with BCFWin [3]. Subsequently data for missing OP insecticides and the nerve agents are calculated based on these formula and the appropriate EPI generated data.

Due to the lack of experimental data for the test set compounds with regard to (logarithmic water–organic carbon partitioning (log K_{OC}) and the ultimate biodegradation potential (*BDP3*), the same procedure was not applicable to these two end–points. Thus, data log K_{OC} and *BDP3* are used as estimated by the appropriate modules in the EPI Suite.

The partial order ranking of the compounds included in this study were made using the WHasse software [9] using the above described "noise–deficient" QSAR generated end–point as descriptors, *i.e.*, log Sol, log K_{OW}, log *VP* and Henry Law constants as generated by the EPI Suite, (log *HLCe*), and by the bond estimation method, (log *HLCb*), organic carbon–water partitioning coefficients (log K_{OC}), and bioconcentration factors (log *BCF*), biodegradation potentials for ultimate degradation (*BDP3*), respectively.

The theory of partial order ranking is presented elsewhere, *e.g.*, [10] and application in relation to QSAR is presented previously [11–14]. In brief, Partial Order Ranking is a simple principle, which a priori includes "≤" as the only mathematical relation. If a system is considered, which can be described by a series of descriptors p_i, a given compound A, characterized by the descriptors p_i(A) can be compared to another compound B, characterized by the descriptors p_i(B), through

comparison of the single descriptors, respectively. Thus, compound A will be ranked higher than compound B, *i.e*., B ≤ A, if at least one descriptor for A is higher than the corresponding descriptor for B and no descriptor for A is lower than the corresponding descriptor for B. If, on the other hand, p_i(A)>p_i(B) for descriptor *i* and p_j(A)< p_j(B) for descriptor *j*, A and B will be denoted incomparable. In mathematical terms this can be expressed as

$$B \leq A \Leftrightarrow p_i(B) \leq p_i(A) \text{ for all } i \quad (3)$$

Obviously, if all descriptors for A are equal to the corresponding descriptors for B, *i.e*., p_i(B) = p_i(A) for all *i*, the two compounds will have identical rank and will be considered as equivalent. It further follows that if A ≤ B and B ≤ C then A ≤ C. If no rank can be established between A and B these compounds are denoted as incomparable, *i.e*. they cannot be assigned a mutual order.

In partial order ranking (in contrast to standard multidimensional statistical analysis) neither assumptions about linearity nor any assumptions about distribution properties are made. In this way the partial order ranking can be considered as a non–parametric method. Thus, there is no preference among the descriptors. However, due to the simple mathematics outlined above, it is obvious that the method *a priori* is rather sensitive to noise, since even minor fluctuations in the descriptor values may lead to non–comparability or reversed ordering. The graphical representation of the partial ordering is often given in a so–called Hasse diagram. In practice the partial order rankings are done using the WHasse software [9]. The generation of the average rank of the single compounds in the Hasse diagram is obtained applying the simple empirical relation recently reported by Brüggemann *et al*. [15]. The average rank of a specific compound, c_i, can be obtained by the simple relation

$$Rk_{av} = (N+1) - (S+1)\times(N+1)/(N+1-U) \quad (4)$$

where N is the number of elements in the diagram, *S* the number of successors to c_i and *U* the number of elements being incomparable to c_i [15].

Nerve Agents **OP Insecticides**

Generalized formula for the Nerve agents and for OP insecticides

2.1 Chemical Data

Even though data on the nerve agents in practice often are unavailable, or at the best scarce (*vide supra*), a wide variety of structurally related compounds are well known and well characterized, *i.e*., OP insecticides such as parathion, malation, diazinon, etc. with respect to physico–chemical and toxicological characteristics [16].

Table 1. FADINAD No., CAS No. and name for the 65 OPs used as tested

FADINAD	CAS No.	Name	log *Sol*	log K_{OW}	VP, Pa	HLC, atm m^3/mol
61	64249–01–0	Anilofos	1.134	3.81		
69	35575–96–3	Azamethiphos	3.041	1.05	5.0129E–06	1.43E–11
70	2642–71–9	Azinphos ethyl	1.021	3.4	0.00031997	9.95E–08
71	86–50–0	Azinphos methyl	1.320	2.75	0.00021332	2.39E–08
126	2104–96–3	Bromophos	–0.523	5.21	0.01706522	2.05E–04
127	4824–78–6	Bromophos ethyl	–0.357	6.15		
139	36335–67–8	Butamifos	0.792	4.62	0.08399286	4.45E–05
150	95465–99–9	Cadusapfos	2.394	3.9	0.1199898	1.29E–06
194	470–90–6	Chlorfenvinphos	2.093	3.81	0.00053329	1.53E–08
201	24934–91–6	Chlormephos	1.778		7.599354	2.93E–04
216	2921–88–2	Chlorpyriphos	0.049	4.96	0.0026931	2.93E–06
217	5598–13–0	Chlorpyriphos methyl	0.678	4.31	0.00559952	3.75E–06
249	56–72–4	Coumaphos	0.176	4.13	1.2932E–05	3.09E–08
261	13067–93–1	Cyanofenphos	–0.222	4.29		
262	2636–26–2	Cyanophos	1.663	2.71	0.10505774	5.48E–06
296	10311–84–9	Dialifos	–0.745	4.69	8.266E–06	1.78E–07
300	333–41–5	Diazinon	1.602	3.81	0.01201231	1.13E–07
312	62–73–7	Dichlorvos	3.903	1.47	2.1064876	5.74E–07
319	141–66–2	Dicrotophos	6.000	–0.49	0.02133152	5.03E–11
359	3811–49–2	Dioxabenzofos	1.763	2.67		
372	5131–24–8	Ditalimfos	2.124	3.48		
390	17109–49–8	Edifenphos	1.748	3.48	3.5997E–05	7.60E–10
415	13194–48–4	Ethoprophos	2.875	3.59	0.05066236	1.62E–07
427	38260–54–7	Etrimfos	1.602		0.01066576	4.62E–07
434	22224–92–6	Fenamiphos	2.517	3.23	0.00013332	1.21E–09
524	66767–39–3	Fonofos	1.196	3.94	0.04506284	6.98E–06
533	83733–82–8	Fosmethilan				
534	98886–44–3	Fosthiazate	3.993	1.68	0.00055995	1.74E–10
558	23560–59–0	Heptonophos	3.398	2.32	0.17065216	1.92E–07
591	26087–47–8	Iprobenfos	2.602	3.34	0.00539954	3.84E–08
594	42509–80–8	Isazofos	1.839	3.82	0.01159901	2.39E–07
598	25311–71–1	Isofenphos	1.342	4.12	0.00039997	6.17E–08
613	18181–70–9	Jodfenphos	–1.000	5.51	0.00010999	4.48E–06
629	121–75–5	Malathion	2.155	2.36	0.00045063	4.89E–09
648	950–10–7	Mephosfolan	1.756	1.04		
664	62610–77–9	Methacrifos	2.602		0.1599864	9.48E–07
665	10265–92–6	Methamidophos	6.000	–0.8	0.00470627	
694	7786–34–7	Mevinphos	5.778	0.13	0.01706522	6.39E–11
705	6923–22–4	Monocrotophos	6.000	–0.2	0.00029064	
771	56–38–2	Parathion	1.041	3.83	0.00089059	2.98E–07
772	298–00–0	Parathion methyl	1.576	2.86	0.00046663	1.00E–07
795	2310–17–0	Phosalone	0.484	4.38		
796	36519–00–3	Phosdiphen	–0.155			
797	947–02–4	Phosfolan	2.813			
798	732–11–6	Phosmet	1.387	2.78	6.5328E–05	8.38E–09
799	13171–21–6	Phosphamidon	6.000	0.79	0.00219981	
810	24151–93–7	Piperophos	1.398	4.04	3.1997E–05	4.47E–09
814	23505–41–1	Pirimiphos ethyl	0.599	4.85	0.03866338	3.21E–05
815	29232–93–7	Pirimiphos methyl	0.934	4.2	0.00199983	7.01E–07
836	41198–08–7	Profenofos	1.447	4.68	0.00011999	2.21E–08
849	7292–16–2	Propafos	2.097	3.67	0.0011999	2.88E–08
853	31218–83–4	Propetamphos	2.041		0.0019065	4.81E–08
864	34643–46–4	Prothiofos	–1.155	5.67		3.01E–05
869	77458–01–6	Pyraclofos	1.519	3.77	1.5999E–06	1.73E–10
872	13457–18–6	Pyrazophos	0.623	3.8		
888	13593–03–8	Quinolphos	1.342	4.44	0.00034664	5.73E–08

Table 1. (Continued)

FADINAD	CAS No.	Name	log *Sol*	log K_{OW}	VP, Pa	HLC, atm m^3/mol
956	3689–24–5	Sulfotep	1.477	3.99	0.01399881	2.40E–06
959	35400–43–2	Sulphofos	–0.509	5.48	0.00015999	1.64E–06
974	3383–96–8	Temephos	–0.569	5.96		
975	107–49–3	TEPP	6.000			
977	13071–79–9	Terbufos	0.705	4.48	0.04266304	2.40E–05
981	22248–79–9	Tetrachlorvinphos	1.041	3.53	5.5995E–06	1.84E–09
1005	57018–04–9	Tolclofos methyl	0.041	4.56	0.05732846	1.54E–04
1015	24017–47–8	Triazophos	1.591	3.34	0.00038663	4.84E–08
1021	52–68–6	Trichlorfon /Chlorophos	5.079	0.51	0.00103991	1.70E–11

In the present paper up to 65 OP insecticides have been used as test set. Dependent of the single endpoint reduced numbers of OPs may be used, simply reflecting the limitations in experimental data (*cf*. empty entries in Table 1). The OP insecticides used are summarized in Table 1. Details on the 16 known and potential nerve agents are given in Table 2.

Table 2. Details on the 16 known and potential nerve agents

Type	CAS No.	Name/code	R	R'	Z
G–agent	77–81–6	Tabun (GA)	$–N(CH_3)_2$	$–C_2H_5$	–CN
G–agent	107–44–8	Sarin (GB)	$–CH_3$	$–CH(CH_3)_2$	–F
G–agent	329–99–7	Cyclosarin (GF)	$–CH_3$	$–c–C_6H_{11}$	–F
G–agent	96–64–0	Soman (GD)	$–CH_3$	$–CH(CH_3)C(CH_3)_3$	–F
V–agent	50782–69–9	VX	$–CH_3$	$–C_2H_5$	$–SCH_2CH_2(N(CH(CH_3)_2)_2$
V–agent	159939–87–4	R–VX	$–CH_3$	$–CH_2CH(CH_3)_2$	$–SCH_2CH_2(N(C_2H_5)_2$
V–agent		C–VX	$–CH_3$	$–C_4H_9$	$–SCH_2CH_2(N(C_2H_5)_2$
VX degr. product		EA2192	$–CH_3$	–H	$–SCH_2CH_2(N(CH(CH_3)_2)_2$
RVX degr. product		R/C "EA2192"	$–CH_3$	–H	$–SCH_2CH_2(N(C_2H_5)_2$
Possible V–agent	78–53–5	VG (Amiton)	$–OC_2H_5$	$–C_2H_5$	$–SCH_2CH_2(N(C_2H_5)_2$
Possible V–agent		Amiton methyl	$–OC_2H_5$	$–C_2H_5$	$–SCH_2CH_2(N(CH_3)_2$
Possible V–agent		Vx (EDMM)	$–CH_3$	$–C_2H_5$	$–SCH_2CH_2(N(CH_3)_2$
Possible V–agent		VM	$–CH_3$	$–C_2H_5$	$–SCH_2CH_2(N(C_2H_5)_2$
Possible V–agent		VE	$–C_2H_5$	$–C_2H_5$	$–SCH_2CH_2(N(C_2H_5)_2$
Possible V–agent		VS	$–C_2H_5$	$–C_2H_5$	$–SCH_2CH_2(N(CH(CH_3)_2)_2$
Possible V–agent		S12	$–CH_3$	$–c–C_5H_9$	$–SCH_2CH_2(N(CH(CH_3)_2)_2$

2.2 Computer Software

EPI Suite: The EPI Suite is available from the EPA web–site at http://www.epa.gov/oppt/exposure/docs/episuitedl.htm. The reference manual for the EPI Suite, P2 Manual 6–00.pdf' can be found and downloaded at http://www.epa.gov/pbt/framwork.htm [3]. WHASSE is described in Brüggemann *et al*. [9]. The software may be obtained by contacting Dr. R. Brüggemann, Institute of Freshwater Ecology and Inland Fisheries, Berlin.

3 RESULTS AND DISCUSSION

The biodegradation potential of the OPs were assessed using the BioWin module of the EPI Suite [3]. In the cases of *BDP3* (ultimate biodegradation) predicted values in the ranges 5.0–4.0, 4.0–3.0, 3.0–2.0, 2.0–1.0 and <1.0 indicate that biodegradation will take place within hours, days,

weeks, months or longer than months, respectively. Chemicals with *BDP3* in the interval of 1.75 to 2 are associated with a medium persistence potential, and *BDP3* smaller than 1.75 were assigned a high persistence potential [13]. In Figure 1 *BDP3* data for the 81 compounds under investigation are summarized. The compounds included in the study are presented in Table 1 (the OP insecticides) and Table 2 (the OP nerve agents).

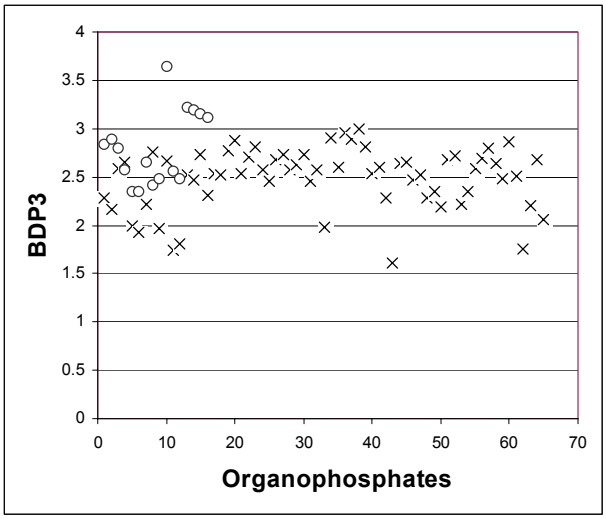


Figure 1. Ultimate biodegradation potential (*BDP3*)of 65 OP insecticides (x) and 16 known or potential nerve agents (o) as derived by BioWin.

Obviously, it is predicted that the biodegradation of the nerve agents are relative fast and as such this should not constitute a problem. However, it should be remembered that if the necessary biological activity is not present significantly longer persistence times might prevail. Thus, half–lives of VX in marine waters and in rivers of up to 1–2 years have been observed [17]. In Figures 2–5 the EPI–based modified QSARs for solubility, octanol–water partitioning, vapor pressure and Henry's Law constants are visualized.

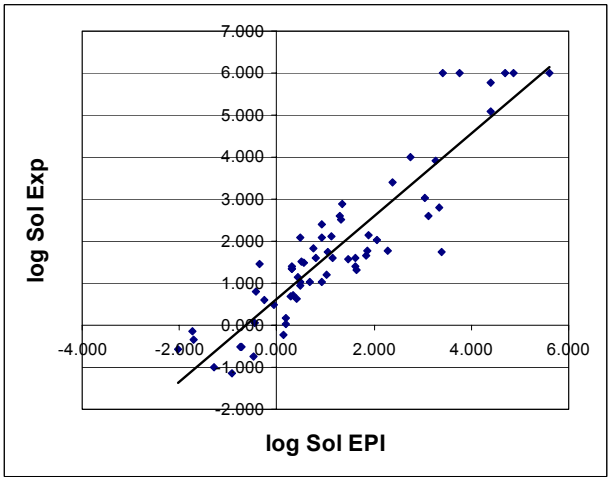


Figure 2. Visualization of the EPI–based modified QSAR modeling of log *Sol* based on 64 OP insecticides.

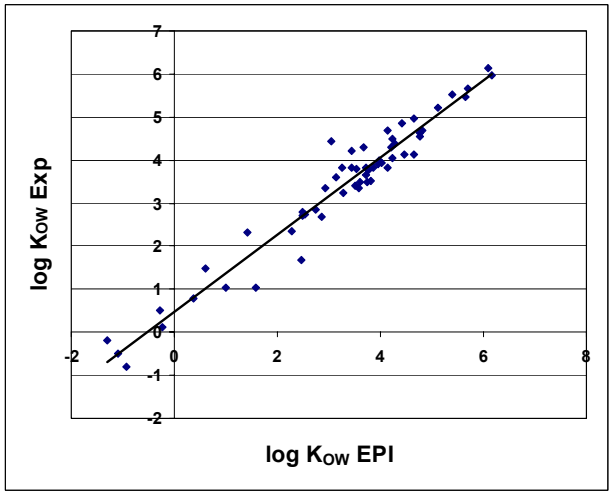


Figure 3. Visualization of the EPI–based modified QSAR modeling of log K_{OW} based on 53 OP insecticides.

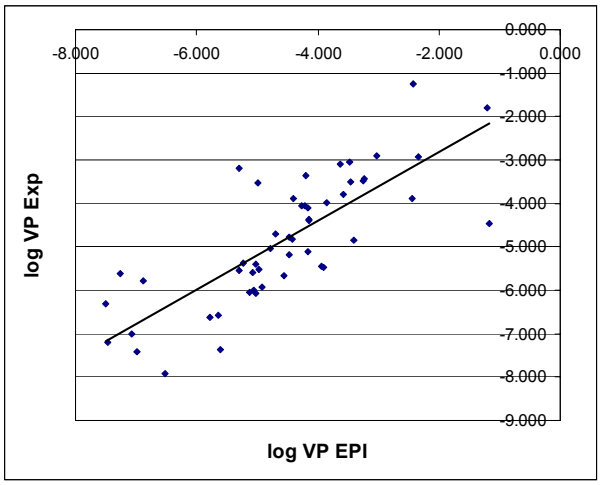
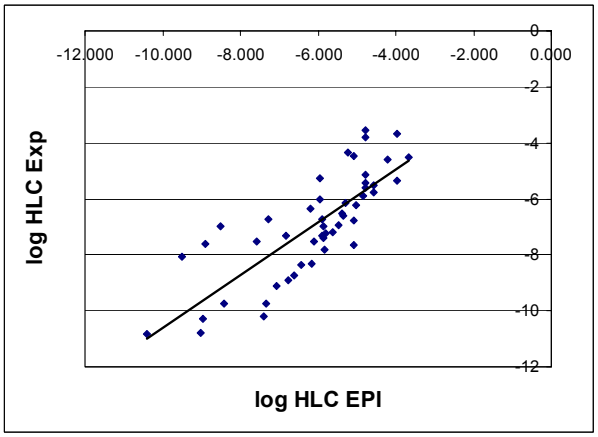


Figure 4. Visualization of the EPI–based modified QSAR modeling of log *VP* based on 51 OP insecticides.



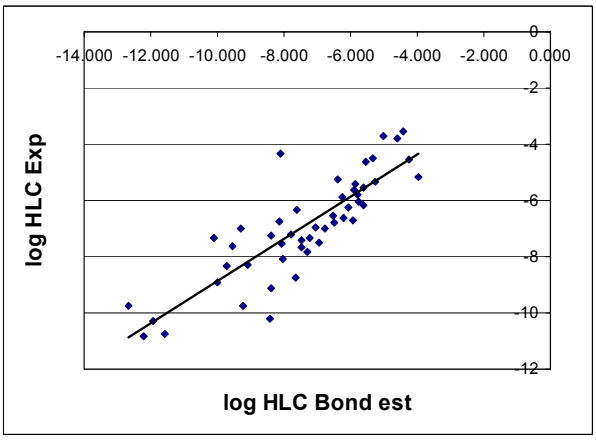


Figure 5. Visualization of the EPI–based modified QSAR modeling of log HLC_e and log HLC_b based on 49 OP insecticides.

The corresponding models are

$$\log Sol = 0.983 \times \log Sol(EPI) + 0.625;\ n = 64,\ r^2 = 0.830 \quad (5)$$

$$\log K_{OW} = 0.894 \times \log K_{OW}(EPI) + 0.487;\ n = 53,\ r^2 = 0.947 \quad (6)$$

$$\log VP = 0.793 \times \log VP(EPI) - 1.229;\ n = 51,\ r^2 = 0.612 \quad (7)$$

$$\log HLC_e = 0.946 \times \log HLCe(EPI) - 1.168;\ r^2 = 0.636 \quad (8)$$

$$\log HLC_b = 0.751 \times \log HLCb(EPI) - 1.371;\ r^2 = 0.727 \quad (9)$$

Although data for nerve agents are scarce some data have been retrieved. These may serve as a validation set for the above models. In Table 3 available experimental data for the "classical" nerve agents, *i.e.* Tabun, Sarin, Cyclosarin, Soman and VX are given together with the corresponding values derived based on the above models. Generally a satisfactory agreement can be observed. However, it appears that in the case of vapor pressure there are more significant disagreements. This is due to the fact that the model for estimating vapor pressures is rather poor, as reflected in the regression coefficient r^2 = 0.612. In some case, *i.e.*, for substances exhibiting very low vapor pressures, such as the V–agents, experimental difficulties in the estimation of the vapor pressures may in this case play a role.

Table 3. Experimental data (solubilities, octanol–water partitioning, vapor pressures, Henry's Law constants) compared to model generated data derived by "noise–deficient" QSARs based on the EPI Suite)

log *Sol* exp	log *Sol* Eq. (5)	log K_{OW} exp	log K_{OW} Eq. (6)	log *VP* exp	log *VP* Eq. (7)	log *HLC* exp	log *HLCe* Eq. (8)	log *HLCb* Eq. (9)
4.99	5.05	0.38	0.74	–1.15	–2.29			
6.00	5.21	0.30	0.70	0.46	–0.70			
4.30	3.94		1.92	–1.22	–1.99			
4.32	3.78	1.78	1.92	–0.40	–1.54	–5.34	–5.17	–4.88
4.48	4.07	2.09	2.33	–3.15	–3.38	–8.09	–7.47	–7.93

Table 4. Physico–chemical data (solubilities, octanol–water partitioning, vapor pressures, Henry's Law constants based on EPI data and bond estimation, respectively, and bioconcentration factors) for known and potential nerve agents as derived by "noise–deficient" QSARs based on the EPI Suite

	log *Sol*	log K_{OW}	log *VP*	log *HLCe*	log *HLCb*	log *BCF*	*BDP3*
Tabun	5.05	0.74	–2.29	–5.97	–6.97	0.50	2.84
Sarin	5.21	0.70	–0.70	–6.94	–7.44	0.50	2.89
Cyclosarin	3.94	1.92	–1.99	–6.74	–6.80	1.02	2.80
Soman	3.78	1.92	–1.54	–4.64	–4.57	1.02	2.58
VX	4.07	2.33	–3.38	–8.11	–10.88	1.25	2.35
R–VX	4.03	2.40	–4.14	–8.13	–10.32	1.29	2.35
C–VX	3.96	2.46	–4.41	–6.65	–7.68	1.32	2.65
EA2192	4.70	1.85	–5.42	–5.69	–5.81	0.98	2.42
R/C "EA2192 [a]	5.55	1.09	–5.28	–5.49	–5.73	0.55	2.48
VG (Amiton)	4.38	2.01	–4.06	–6.17	–7.91	1.07	3.64
Amiton methyl	5.36	1.13	–3.83	–9.46	–8.62	0.57	2.57
Vx (EDMM)	5.91	0.71	–3.03	–5.15	–5.55	0.50	2.48
VM	4.94	1.59	–3.78	–7.41	–7.13	0.83	3.22
VE	4.45	2.02	–4.11	–5.69	–4.82	1.08	3.19
VS	3.60	2.77	–4.22	–7.63	–8.97	1.51	3.15
S12	2.78	3.48	–3.41	–9.71	–10.07	2.16	3.12

[a] Analog to EA2192, however, derived from Russian or Chinese VX

Subsequently the above given models have been applied in estimating physico–chemical data for a broader range of known as well as potential nerve agents. In Table 4 these data are collected together with the corresponding logarithmic bioconcentration factors as derived through the Connell formula (*vide supra*) applying the log K_{OW} values obtained using the model given in Eq. (5).

The model–generated end–points may subsequently be used as descriptors in ranking the 65 OP insecticides together with the 16 known potential nerve agents. Thus, as in total 81 compounds are included in the subsequent ranking procedure, the resulting Hasse diagrams may seem somewhat confusing. In Figure 6 the Hasse diagram disclosing the mutual ranking of the compounds due to their combined PB (Persistent and Bioaccumulating) characteristics, *i.e.*, bringing simultaneously the *BDP3* and log *BCF* into play.

A priori, compounds located on the same level in the Hasse diagram are assumed to be close in their overall characteristics based on the set of descriptors used. On this basis, in the above example (Figure 6) the highly toxic EDMM (LD_{50} (rats, acute, oral) = 0.121 mg/kg) may be substituted by, *e.g.*, compounds No. 71 (Azinphos Methyl; LC_{50} = 4 mg/kg), 312 (Dichlorvos; LC_{50} = 56 mg/kg), 591 (Iprobenfos; LC_{50} = 490 mg/kg) or 648 (Mephosfolan; LC_{50} = 8.9 mg/kg).

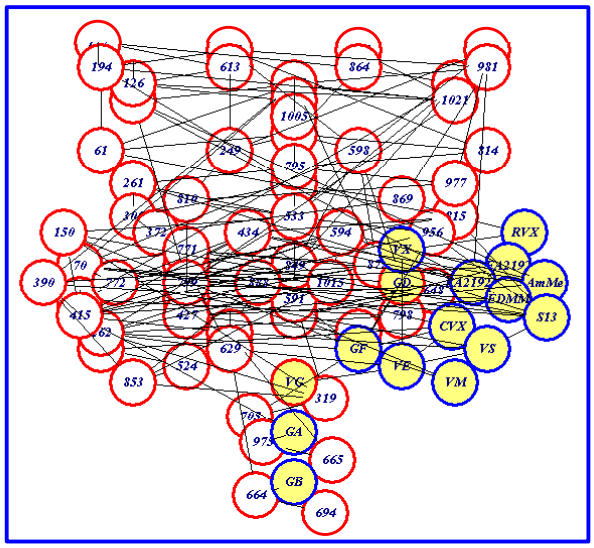


Figure 6. Hasse diagram displaying the PB characteristics of the 65 OP insecticides and 16 nerve agent (hatched), The numbers corresponds to the numbering of the OP insecticides in the FADINAP database.

However, a further analysis is necessary to disclose how close these compounds actually are. For this analysis we have chosen the concept of average rank [15,18]. Thus, it is assumed that if the average ranks, Rk_{av}, of two compounds are close, the two compounds will on an average basis display similar characteristics as being determined by the set of descriptors applied. In Table 3 the average ranks for the four compounds are compared.

From the values for the average rank (Table 5), that the four possible substitutes for EDMM located on the same level in the Hasse diagram based on average ranks apparently can be regarded as being rather close. Thus, taking the actual toxicities, as expressed through the LC_{50} values into account Dichlorvos appear as the optimal choice as substitute for EDMM in studies where the PB characteristics of the compounds is important, the toxicity associated with the experiments being decreased by a factor close to 500.

Table 5. Average ranks for the PB characteristics as determined by log *BCF* and the biodegradation potential for EDMM, Azinphos Methyl, Dichlorphos, Iprobenfos and Mephosfolan (the compound ID refers to the FADINAP database [16])

Compound name	Compounds ID (Table 1 and 2)	LC_{50} (mg/kg) Rat, Acute, Oral	Average Rank RK_{av}
EDDM	EDDM	0.121	55.9
Azinphos Methyl	71	4	51.1
Dichlorvos	312	56	55.9
Iprobenfos	591	490	50.6
Mephosfolan	648	8.9	57.9

Similar analyses for other descriptor combinations can obviously be carried out analogously. This is, however, outside the scope of the present paper and is reported elsewhere [19]. The Chemical Weapons Convention [1] obviously covers all the nerve agents. However, as a curiosity it can be noted that the nerve agents, despite their extreme toxicity based on their PB characteristics

(*cf*. Figure 1 and Table 2) would not immediately qualify to falling into the group of chemicals that requires special authorization in the coming European system for registration, evaluation and authorization of chemicals, REACH, as these compounds would neither be classified as persistent nor as bioaccumulating [20]. This is further substantiated by the placement of all the 16 known or potential nerve agents in the lower half of the Hasse diagram reflecting the combined PB characteristics (Figure 6). Thus, in cases of extremely toxic compounds, not being covered by, *e.g*., the Chemical Weapons Convention, it might in addition to the PBT and vPvB (very Persistent and very Bioaccumulating) criteria be relevant to introduce a vT (very Toxic) criteria as well.

4 CONCLUSIONS

The present study has demonstrated that "noise–deficient" QSARs can be generated using the EPI Suite as the modeling onset. Subsequently, the generated physico–chemical end–points can be used as descriptors in a partial order based ranking giving compounds where experimental data are not available an identity by comparing to a test set of experimentally well–characterized, structurally similar compounds.

On this background it has been suggested that experimentally well–characterized compounds may be selected as substitutes for highly toxic compounds, as the nerve agent. Hence, this procedure allows that the environmental behavior of the latter may be studied experimentally using compounds that from an overall viewpoint exhibit analogous environmental characteristics, however, without exhibiting the extreme toxicity.

5 REFERENCES

[1] Convention on the prohibition of the development, stockpiling and use of chemical weapons and their destruction: http://www.opcw.org/html/db/cwc/eng/cwc_frameset.html.
[2] N. B. Munro, S. S. Talmage, G. D Griffin, L. C. Waters, A. P. Watson, J. F. King, and V. Hauschild, The sources, fate and toxicity of chemical warfare agent degradation products, *Environ*. *Health Perspect*. **1999**, *107*, 933–974.
[3] Pollution Prevention (P2) Framework, EPA–758–B–00–001; may be obtained through the link 'P2 Manual 6–00.pdf' found at http://www.epa.gov/pbt/framwork.htm, US EPA.
[4] J. de Bruijn and J. Hermens, Inhibition of acetylesterase and acute toxicity of organophosphorous compounds to fish: a preliminary structure–activity analysis, *Aquatic Tox*. **1993**, *24*, 257–274.
[5] H. J. M. Verhaar, L. Eriksson, M. Sjöström, G. Schüürmann, W. Seinen, and J. L. M. Hermens, Modelling the toxicity of organophosphates: a comparison of the multiple linear regression and PLS regression methods, *Quant*. *Struct*.*–Act*. *Relat*. **1994**, *13*, 133–143.
[6] D. V. Eldred and P. C. Jurs, Prediction of acute mammalian toxicity of organophosphorous pesticide compounds from molecular structure, *SAR QSAR Environ*. *Res*. **1999**, *10*, 75–99.
[7] A. K. Singh, QSAR for the organophosphate–induced inhibition and 'aging' of the enzyme neuropathy target esterase (NTE), *SAR QSAR Environ*.*Res*. **2001**, *12*, 275–295.
[8] D. W. Connell and D. W. Hawker, Use of polynomial expressions to describe the bioconcentration of hydrophobic chemicals in fish, *Ecotox*. *Environ*. *Safety* **1988**, *16*, 242–257.
[9] R. Brüggemann, E. Halfon, and C. Bücherl, Theoretical base of the program "Hasse", *GSF–Bericht 20/95*, *Neuherberg*, **1995**; The software may be obtained by contacting Dr. R. Brüggemann, Institute of Freshwater Ecology and Inland Fisheries, Berlin.
[10] B. A. Davey and H. A. Priestley, *Introduction to lattices and order*. Cambridge University Press, Cambridge,

1990.
[11] L. Carlsen, P. B. Sørensen, and M. Thomsen, Partial order ranking based QSAR's: Estimation of solubilities and octanol–water partitioning, *Chemosphere* **2001**, *43*, 295–302.
[12] R. Brüggemann, S. Pudenz, L. Carlsen, P. B. Sørensen, M. Thomsenand, R. K. Mishra, The use of Hasse diagrams as a potential approach for inverse QSAR, *SAR QSAR Environ*. *Res*. **2001**, *11*, 473–487.
[13] L. Carlsen and J. D. Walker, QSARs for Prioritizing PBT Substances to Promote Pollution Prevention, *QSAR Comb*. *Sci*. **2003**, *22*, 49–57.
[14] L. Carlsen, P. B. Sørensen, and D. B. Lerche, A decision support tool to prioritize chemical substances. Partial order ranking using QSAR generated descriptors in: P. B. Sørensen R. Brüggemann, D. B. Lerche, K. Voigt, G. Welzl, U. Simon, M. Abs, M. Erfmann, L. Carlsen, S. Gyldenkærne, M. Thomsen, P. Fauser, B. B. Mogensen, S. Pudenz and B. Kronvang, 'Order theory in environmental sciences. Integrative approach', *NERI Technical Report No*. *479*, **2003**.
[15] R. Brüggemann, D. Lerche, P. B.Sørensen, and L. Carlsen, Estimation of average ranks by a local partial order model, *J*. *Chem*. *Inf*. *Comput*. *Sci*. **2004**, *44*, 618–625.
[16] Database on pesticide and the environment, http://www.fadinap.org/pesticide/.
[17] Chemical Weapons – threat, effects and protection, *Swedish Defense Research Agency*, *Briefing Book Number 2*, **2002**
[18] D. Lerche, R. Brüggemann, P. Sørensen, L. Carlsen, and O. J. Nielsen, A comparison of partial order technique with three methods of multi–criteria analysis for ranking of chemical substances, *J*. *Chem*. *Inf*. *Comput*. *Sci*. **2002**, *42*, 1086–1098.
[19] L. Carlsen, Partial Order Ranking of Organophosphates with Special Emphasis on Nerve Agents. *Communications in Mathematical and in Computer Chemistry*, *MATCH*, accepted for publication
[20] J. D. Walker and L. Carlsen, QSARs for identifying and prioritising substances with persistence and bioconcentration potential, *SAR QSAR Environ*. *Res*. **2002**, *13*, 713–726.

Biographies

Lars Carlsen is owner and Director of Awareness Center that offers services within the environmental field as on management support (www.AwarenessCenter.dk). He has more than 25 years of experience in R&D and training activities within the area of environmental chemistry. Active in the development and application of model tools, such as structure–activity relationships and ranking techniques as decision support tools for the evaluation and prioritization of chemicals to be carried out by industry and regulatory bodies and how these tools can be applied to prioritize the possible action by industry and regulatory bodies in the environmental area, including studies on how these tools can be integrated in environmental management. Holds a position as assigned professor in environmental chemistry at Roskilde University and lectures in environmental chemistry at the Universities in Roskilde and Copenhagen.