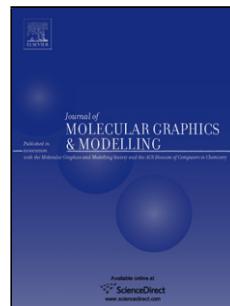


Accepted Manuscript

Title: Ligand-based 3D QSAR analysis of reactivation potency of mono- and bis-pyridinium aldoximes toward VX-inhibited rat acetylcholinesterase

Author: Rafael Dolezal Jan Korabecny David Malinak Jan Honegr Kamil Musilek Kamil Kuca



PII: S1093-3263(14)00198-3

DOI: <http://dx.doi.org/doi:10.1016/j.jmgm.2014.11.010>

Reference: JMG 6489

To appear in: *Journal of Molecular Graphics and Modelling*

Received date: 23-4-2014

Revised date: 8-11-2014

Accepted date: 21-11-2014

Please cite this article as: R. Dolezal, J. Korabecny, D. Malinak, J. Honegr, K. Musilek, K. Kuca, Ligand-based 3D QSAR analysis of reactivation potency of mono- and bis-pyridinium aldoximes toward VX-inhibited rat acetylcholinesterase, *Journal of Molecular Graphics and Modelling* (2014), <http://dx.doi.org/10.1016/j.jmgm.2014.11.010>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Ligand-based 3D QSAR analysis of reactivation potency of mono- and bis-pyridinium aldoximes toward VX-inhibited rat acetylcholinesterase

Rafael Dolezal^{a,e}, Jan Korabecny^{a,b,e}, David Malinak^{a,d}, Jan Honegr^{a,d}, Kamil Musilek^c, Kamil Kuca^{*a}.

^a Biomedical Research Center, University Hospital Hradec Kralove, Sokolska 581, 500 05 Hradec Kralove, Czech Republic

^b Department of Toxicology, Department of Public Health, Centre for Advanced Studies, Faculty of Military Health Sciences, University of Defence, Trebesska 1575, 500 01 Hradec Kralove, Czech Republic

^c Department of Chemistry, Faculty of Science, University of Hradec Kralove, Rokitanskeho 62, 500 03 Hradec Kralove, Czech Republic

^d Department of Cybernetics and Biomedical Engineering, Faculty of Electrical Engineering and Computer Science, VSB-Technical University of Ostrava, 17. listopadu 15, 708 33 Ostrava-Poruba, Czech Republic

^e Center for Basic and Applied Research, Faculty of Informatics and Management, University of Hradec Kralove, Rokitanskeho 62, 500 03 Hradec Kralove, Czech Republic

*Corresponding author: Fax: +420-495-832-100; Tel: +420-495-832-923; E-mail: kamil.kuca@fnhk.cz (K. Kuca)

Keywords: 3D QSAR, acetylcholinesterase, molecular docking, computational chemistry, VX, reactivators.

Highlights

1. 3D QSAR analysis of mono- and bis-pyridinium aldoximes was performed
2. Molecular mechanical and quantum chemical interaction fields were generated
3. Significance of 3D QSAR models was improved through data noise reduction methods
4. 12 reactivation potencies for VX-inhibited acetylcholinesterase were predicted
5. SARs were proposed through analyzing 3D contour maps of β -PLS x SD coefficients

Abstract

To predict unknown reactivation potencies of 12 mono- and bis-pyridinium aldoximes for VX-inhibited rat acetylcholinesterase (*rAChE*), three-dimensional quantitative structure-activity relationship (3D QSAR) analysis has been carried out. Utilizing molecular interaction fields (MIFs) calculated by molecular mechanical (MMFF94) and quantum chemical (B3LYP/6-31G*) methods, two satisfactory ligand-based CoMFA models have been developed: 1. $R^2 = 0.9989$, $Q^2_{LOO} = 0.9090$, $Q^2_{LTO} = 0.8921$, $Q^2_{LMO(20\%)} = 0.8853$, $R^2_{ext} = 0.9259$, $SDEP_{ext} = 6.8938$; 2. $R^2 = 0.9962$, $Q^2_{LOO} = 0.9368$, $Q^2_{LTO} = 0.9298$, $Q^2_{LMO(20\%)} = 0.9248$, $R^2_{ext} = 0.8905$, $SDEP_{ext} = 6.6756$. High statistical significance of the 3D QSAR models has been achieved through the application of several data noise reduction techniques (i.e. smart region definition SRD, fractional factor design FFD, uninformative/iterative variable elimination UVE/IVE) on the original MIFs. Besides the ligand-based CoMFA models, an alignment molecular set constructed by flexible molecular docking has been also studied. The contour maps as well as the predicted reactivation potencies resulting from 3D QSAR analyses help better understand which structural features are associated with increased reactivation potency of studied compounds.

Abbreviations

ACh – acetylcholine; AChE – acetylcholinesterase; AM1 – Austin Model 1; BBB – blood-brain barrier; CADD – computer assisted drug design; CAS – catalytic anionic site; CoMFA – comparative molecular field analysis; DFT – density functional theory; ESP – electrostatic potential; FFD – fractional factor design; *hAChE* – human acetylcholinesterase; HOMO – highest occupied molecular orbital; IVE – iterative variable elimination; LV – latent variable; LMO – leave many out; LOO – leave one out; LTO – leave two out; LUMO – lowest unoccupied molecular orbital; *mAChE* – mouse acetylcholinesterase; MIF – molecular interaction field; MLR – multiple linear regression; OP – organophosphates; PAS – peripheral anionic site; PLS – partial least square; Q^2 – cross-validated coefficient of determination; QMD – quenched molecular dynamics; QSAR – quantitative structure-activity relationship; *rAChE* – rat acetylcholinesterase; R – Pearson's pairwise correlation coefficient; R^2 – coefficient of determination; RMSD – root mean square deviation; RP – reactivation potency; SAR – structure-activity relationship; SD – standard deviation; SDEC – standard error of calculation; SDEP – standard error of prediction; SRD - smart region definition; SSE – error sum of squares; SSR – regression sum of squares; SST – total sum of squares; TCL – tool command language; UVE – uninformative variable elimination; VDW – van der Waals energy; VP – Voroni polyhedra.

1. Introduction

The function of the human nervous system is based on a complex interaction of nerves through neurotransmitters by which the activity of multiple organs, glands, or other neurons is kept in a dynamic balance [1]. Amongst these highly important small molecules we should lay emphasis particularly on acetylcholine (ACh) which is involved not only in the autonomic and but also in the peripheral nervous systems as a major skeletal muscle activator. Through three main subdivisions of the autonomic nervous system, it can act on several receptor types in the body eliciting excitatory as well as inhibitory physiological effects. In the central nervous system, ACh significantly participates in a variety of essential processes such as sensory perception, remembering, sustaining attention, sleeping, arousal and the rewarding system. Thus, the neural signal pathways may be severely compromised when ACh receptors are being blocked or the hydrolytic activity of acetylcholinesterase (AChE, EC 3.1.1.7) is inhibited. The vital role of AChE in the nervous system is to terminate the transmission of nerve signals by ACh degradation. Should the hydrolytic activity of AChE be

irreversibly disabled, a serious pathophysiological state associated with prolonged depolarizing neuromuscular blockade, severe convulsions, bronchial collapse, diaphragm paralysis and suffocation develops due to the excess of ACh [2].

The most notoriously known irreversible inhibitors of AChE are the organophosphate (OP) nerve agents such as VX, RVX, sarin, cyclosarin, soman or tabun (Fig. 1) [3]. Their extremely toxic effect derives from rapid phosphorylation (or phosphorylation) of the Ser203 hydroxyl group (this relates to human AChE - *h*AChE) at the bottom of the active site with consequent inhibition of the hydrolytic activity of AChE (Fig. 2). As a result, the accumulated ACh saturates both nicotinic and muscarinic receptors causing a critical cholinergic condition throughout the entire body. Therefore, organophosphorus substances pose a threat as poisons exploitable especially for military and terrorist purposes [4]. The median human lethal dose of VX (*O*-ethyl *S*-[2-(diisopropylamino)ethyl] methylphosphonothioate), being the most toxic nerve agent ever synthesized, is approximately 0.14mg/kg absorbed through the skin [5]. The largest terrorist attack with sarin gas was committed in the Tokyo subway on the 20th March 1995 [6]. Nevertheless, substances that inhibit AChE are utilized beneficially as pesticides (e.g. paraoxon, parathion, malathion; Fig. 1) [7] or symptomatic drugs in Alzheimer's disease (e.g. tacrine, donepezil) [8, 9, 10], Lewy body dementia (e.g. rivastigmine) [11] and Myasthenia gravis (e.g. pyridostigmine bromides) [12].

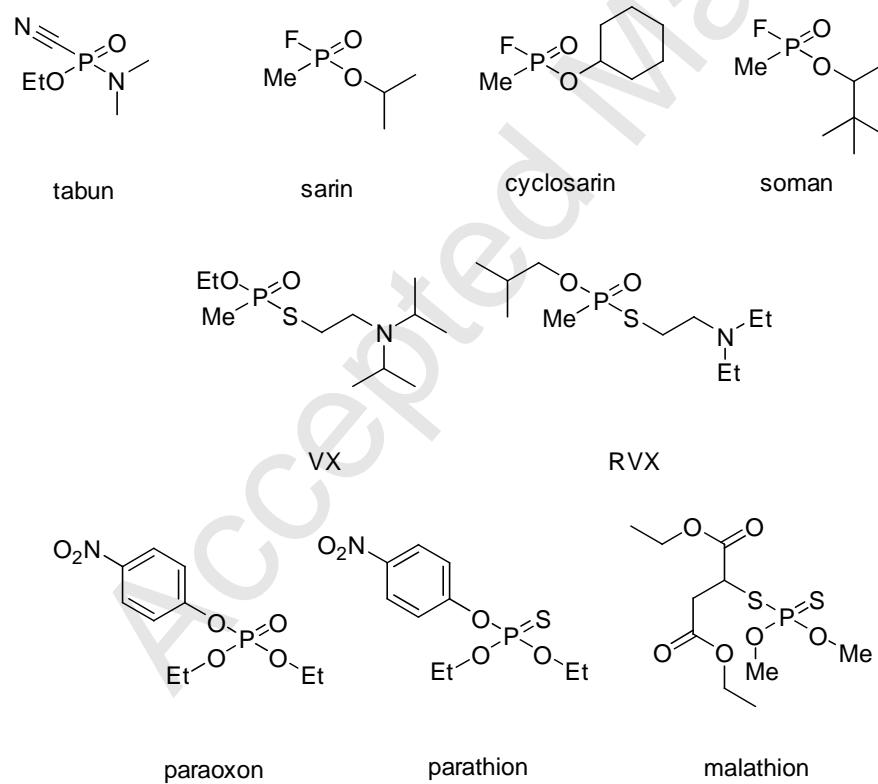


Figure 1. Structures of organophosphorus nerve agents.

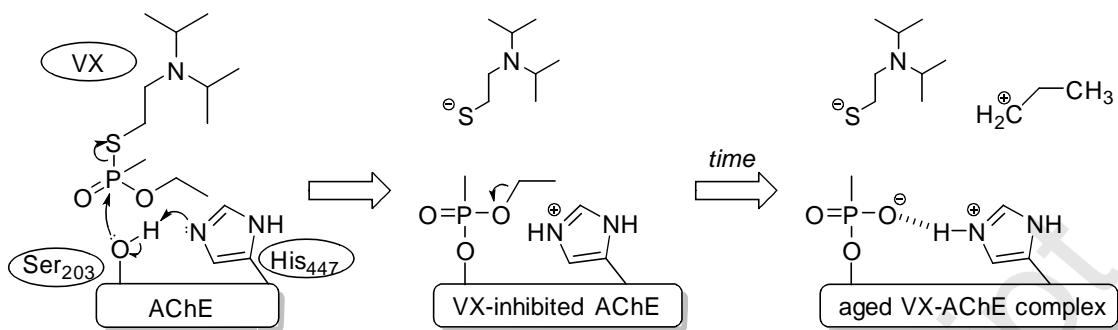


Figure 2. Mechanism of *h*AChE inhibition by VX followed by the aging process. Arrows denote electron shifts.

The current treatment of acute organophosphate poisoning is accomplished *de facto* by reactivation of inhibited AChE using various oximes (e.g. pralidoxime, trimedoxime, Hlö-7, obidoxime, HI-6, Fig. 3). Besides oximes, parasympatholytics (e.g. atropine) and anticonvulsants (e.g. diazepam) are administered in a combination therapy. Parasympatholytic agents serve to competitively inhibit the muscarinic receptors from ACh stimulation, and thus to compensate for the negative effect of an excess of ACh. An oxime, specifically its oximate anion, acts as a nucleophile that attacks the electron-deficient phosphorus atom attached to the Ser203 residue in the OP-AChE complex of *h*AChE (Fig. 4) [13].

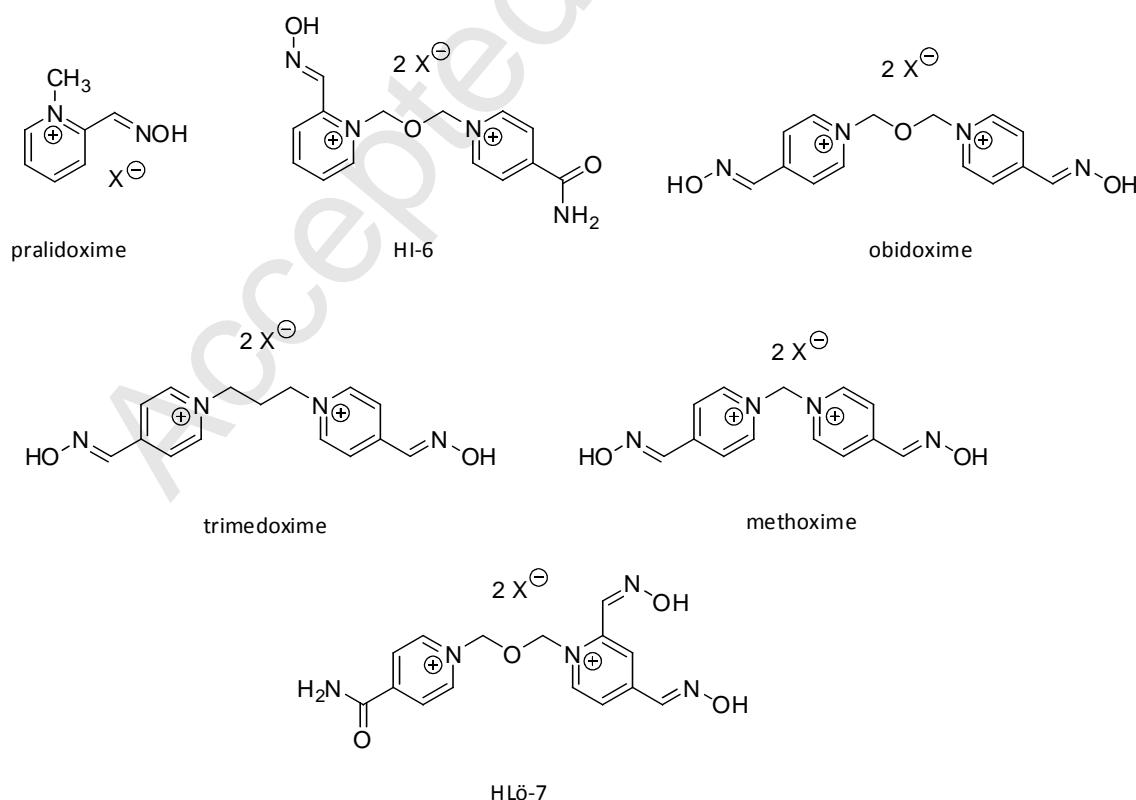


Figure 3. Oxime reactivators used as antidotes for OP intoxication.

Depending on the chemical structure, OPs bind to AChE with different affinity, inducing numerous changes in the conformation of residues at the active site. Furthermore, each OP-AChE complex exhibits a different half-life of aging (e.g. 2-4 min soman, 5-12h sarin, 46h tabun, 48h VX) [14] which goes hand-in-hand with significant decline in willingness to reactivate [15]. The essence of this chemical process is irreversible dealkylation of alkoxy group bonded to the phosphorus atom and subsequent intramolecular stabilization of the OP-AChE complex by a donor-acceptor bond with the proximal histidine residue (i.e. His447 in *h*AChE) (Fig. 2) [16]. The faster the aging of the OP-AChE complex evolves, the worse is the prospect for reactivation as time proceeds. The consequence is that no approved reactivator of inhibited AChE has such a high broad-spectrum effectiveness to be used as a universal antidote for all OP poisoning [17]. Therefore, it is highly desirable to search for better reactivators with increased reactivation potency toward OP inhibited AChE. Current scientific attention in this area is especially paid to the development of new antidotes with: 1) a better ability to cross the blood-brain barrier (BBB), 2) a higher effectiveness toward different OP-AChE complexes, 3) a capability to restore “aged” OP-AChE complexes[18, 19] . Nonetheless, these requirements for improved OP-AChE reactivators need to be solved rationally through understanding the relationship between the structure, its physicochemical properties and the resulting biological activity. Moreover, it is also essential to reflect in the development such reactivator properties as are known to correlate with their toxicity and negative side effects [20, 21].

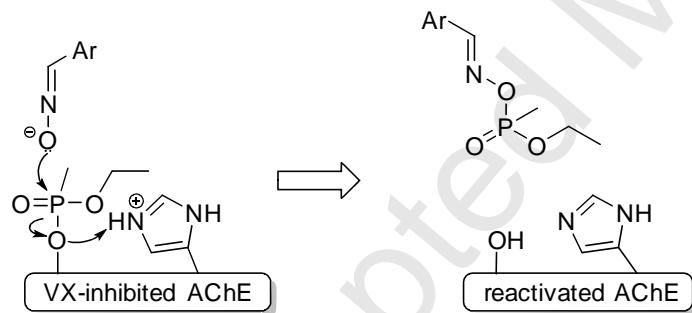


Figure 4. Reactivation of VX-inhibited AChE. Arrows denote electron shifts.

A significant part of these demands for improved reactivators of inhibited AChE may be approached through computational chemistry. Undesirably low lipophilicity (i.e. $\log P$) of reactivators can be revealed by a simple QSPR (quantitative structure properties relationship) filter, whilst optimal pharmacodynamic properties are to a certain extent deducible from molecular docking [22]. So far, quite a few SAR&QSAR (structure activity relationship; quantitative SAR) studies have been published in this area [23,24,25,26,27,28,29,30]. For example, computer analysis of the crystallographic structure of AChE from *Torpedo californica* (*Tc*AChE) using programs for molecular docking led to discovery of two major binding sites for quaternary ligands. The main binding site containing Trp84 and Phe330 was localized close to the catalytic triad (i.e. catalytic anionic site – CAS); the other binding site appeared at the entrance to the active-site gorge of *Tc*AChE (i.e. peripheral anionic site – PAS). The PAS involved aromatic residues such as Tyr70, Tyr121 and Trp279 [31]. Based on these *in silico* studies, novel dual binding-site reactivators of isofluorophate- and echothiophate-inhibited *h*AChE have been prepared, some of which reached in experimental measurements up to 100 times

higher reactivation potency than pralidoxime [32]. The optimal length of the linker for dual binding site reactigators proposed by Pang et al. was seven methylene groups.

Musilek et al. found out *via* molecular docking that for an effective reactivation of sarin-inhibited *mAChE* (from *Mus musculus*) by mono-aldoxime bis-quaternary reactigators π - π and cation- π interactions were crucial between the aromatic cycles of the reactigator and three aromatic residues in CAS (Tyr337, Phe338, Tyr341) and four aromatic residues in PAS (Tyr72, Tyr124, Trp286, Phe297) [22]. For the formation of hydrogen bonds with reactigators a close proximity of Glu285 and Ser298 residues was considered important. The increased potency of one of the studied compounds containing a biphenyl skeleton, was ascribed by the authors to other π - π interactions at the active site of *mAChE*. Although working only with qualitative SARs, the authors assumed the reactivation was facilitated by increased binding affinity of bis-quaternary aldoximes toward *mAChE*.

Maxwell et al. tried to elucidate the relationship between the reactivation potency of different aldoximes and the steric properties of the OPs. They concluded that substituents bound to OPs which had a larger molecular volume (according to the Connolly computational method [33]) than that of isopropyl ($MV = 63.3 \text{ \AA}^3$) significantly hindered access of mono-pyridinium aldoximes to the active site and prevented them from non-bonding interactions with the key residues in CAS [34]. However, in the case of a bis-quaternary reactigator possessing sufficiently high affinity for PAS, the negative consequences of steric hindrance in CAS can be to some extent compensated. Analogous results have been recently published by the same authors demonstrating that the presence of a substituent larger than dimethylamine in the OP molecule (i.e. sarin, cyclosarin, tabun and their derivatives) elicits resistance to oxime reactivation of OP-AChE complexes [35]. In the latter work, they also showed through a multiple linear regression (MLR) based QSAR model that the steric properties of the OP agents made a greater impact on the reactivation of inhibited AChE than the electronic properties of substituents in the OPs or the pK_o for ionization of the aldoxime group in the reactigators. The MLR QSAR model reported was statistically significant (internal Pearson's $R = 0.85$, $n = 75$, 5 molecular descriptors involved), but no external validation was performed.

Bhattacharjee et al. elaborated a SAR study evaluating various quantum-chemical properties (e.g. maps of electrostatic potential, energies and delocalization of frontier molecular orbitals, geometrical properties, etc.) of 11 bis-pyridinium aldoxime reactigators of tabun-inhibited AChE [30]. Since no simple relationship was apparent between the binding affinities of the reactigators for tabun-inhibited AChE and the quantum-chemical molecular descriptors, the authors only hypothesized the binding affinity be supported by lower nucleophilic power of the aldoxime oxygen and by specific location of highest occupied and lowest unoccupied molecular orbitals (HOMO, LUMO). Besides the analysis of quantum-chemical properties, the authors developed a quantitative pharmacophore model which was able to predict the binding affinities in the training set with Pearson's $R = 0.93$. However, in external prediction of binding affinities of four novel bis-pyridinium aldoxime reactigators the model, from the quantitative point of view, failed (external Pearson's $R = 0.05$).

Recently, a quite extensive QSAR study on aldoxime reactigators of various OP-inactivated AChE complexes has been published by Esposito et al. [36]. Accidentally, the authors processed data which had been determined and published before within our workgroup [37]. Esposito et al. developed several QSAR models articulating the reactivation potency of various pyridinium aldoximes for

cyclosarin-, sarin-, tabun- and VX-inactivated AChE. The study was based on building MLR QSAR models from a 209-member molecular descriptor pool (7 AM1 quantum-chemistry descriptors, 207 2D descriptors and 76 VolSurf descriptors) and Tanimoto similarity indexes. Prior to the calculation of molecular descriptors, the aldoximes were flexibly superimposed on a template constructed from four X-ray AChE-ligand structures available from the www.pdb.org databank under the following PDB codes: Hlö-7 from 2GYV, ortho-7 from 2GYW, K027 from 2WHR and obidoxime from 2JEZ. The resulting MLR QSAR models for four types of OP-inhibited AChEs contained two to five molecular descriptors and exhibited acceptably high levels of internal statistical significance criteria ($R^2 = 0.77 - 0.99$, $Q^2_{LOO} = 0.68 - 0.98$). External validation of the QSAR models, calculation of variance inflation factors of involved descriptors, or determining the domain of applicability of the models were not performed. In simplified terms, the found QSAR models suggest that reactivation potency for sarin- and tabun-inactivated AChE is generally strengthened by increasing the polar-positive surface area of aldoxime reactivators. Additionally, reducing the number of oxygen atoms in the reactuator molecule improves reactivation of cyclosarin-inactivated AChE, while for high reactivation potency toward tabun-inhibited AChE a distance of 5 - 8 Å between the most hydrophobic surface locations in the reactuator should be preserved. Sarin-inhibited AChE is reactivated more easily if the reactuator is flexible enough and is able to adopt a conformation similar to the co-crystallized structures used as the template in the molecular alignment. Regarding the reactivation of VX-inhibited AChE, the authors have concluded that an efficient reactuator should contain two positively charged pyridine rings and highly hydrophilic substituents.

A major goal of our work is firstly to carry out ligand-based 3D QSAR comparative molecular field analysis (CoMFA) of 17 mono/bis-quaternary pyridinium aldoxime reactivators of VX-inhibited AChE from rat (*r*AChE), and to decide which data processing method implemented in Open3DQSAR program [38] provides the most optimal 3D QSAR model. In terms of the strategy for QSAR analyses focused on OP-AChE reactivation, our approach tries to connect advanced molecular property description with thorough statistical validation to build a 3D QSAR model reliable for external prediction.

To this end, we utilized an open-source package developed by Paolo Tosco which includes a set of command-line based applications aimed at most common computational methods for chemometric analysis. It can perform, for instance, unsupervised molecular alignment, quenched molecular dynamics, calculation of steric potential, electron density, electrostatic potential fields, principal component analysis, partial least square regression, cross-validation, variable selection, data scrambling procedures, etc. High computational performance of the applications is reached through employing parallelized algorithms (for further information see: <http://open3dqsar.sourceforge.net/>; <http://open3dalig.sourceforge.net/>).

In the present study, particular attention is given especially to data noise reduction techniques implemented in Open3DQSAR program. These tools might be very beneficial to 3D QSAR analyses if used prior to developing a model by regression analysis. Since one of the most desirable traits of a QSAR model is its predictive power, the second aim of our study is to predict externally by the best obtained 3D QSAR model the reactivation potencies of 12 mono/bis-pyridinium aldoxime reactivators for VX-inhibited *r*AChE. To develop the 3D QSAR model, the chemical structures and reactivation potencies of 17 mono/bis-quaternary pyridinium aldoxime reactivators reported in the literature were taken as training and evaluating data [39]. Although 3D QSAR analysis of such a

limited data set may be risky if not properly revised from a statistical point of view, this study fills a gap in the actual research of reactivators of OP-inhibited AChE. If 3D QSAR methodology is considered with reasonable skepticism and the limits imposed by the used computational methods are respected, the obtained 3D QSAR models clarify in quantitative terms which molecular features significantly influence the reactivation process. As a quantitative insight into the relationship between the structure and reactivation potency for VX-inhibited *r*AChE is still very limited, the present work may shed new light on rational design of improved reactivators.

2. Materials and methods

2.1. Chemical structures and reactivation potencies for 3D QSAR study

In the paper, two sets of mono/bis-pyridinium aldoxime reactivators were used. The first data set contained 17 structures of VX-inhibited *r*AChE reactivators along with their reactivation potencies (RP), both having been taken from the literature [39]. As a source of AChE, rat (*Rattus norvegicus*, Wistar) brain homogenate was used in that study. The reactivation potencies were determined under the same laboratory protocol and were expressed as an average relative restoration of the intact *r*AChE hydrolytic activity caused by 10⁻³ M reactuator concentration (Eq. 1). The activity of *r*AChE was measured by potentiometric acid-base titration. The majority of the compounds reached reactivity potency higher than 5-10% and may be, according to previous toxicological assessments [40], sufficient in OP poisoning treatment. All reactivators involved in the present paper are believed to act at the same binding site of *r*AChE and by the same mechanism.

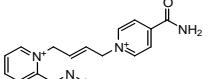
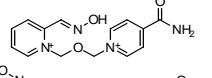
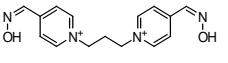
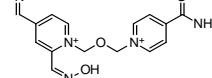
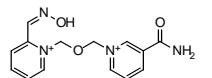
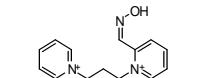
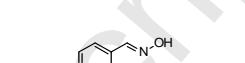
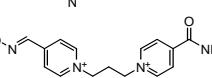
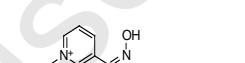
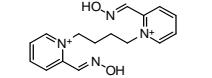
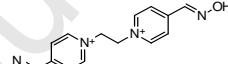
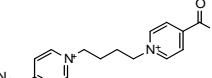
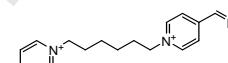
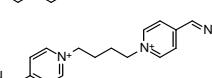
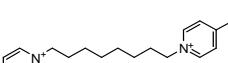
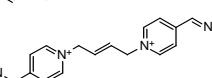
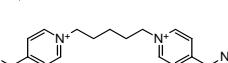
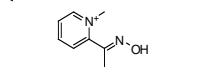
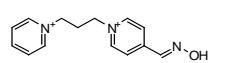
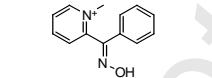
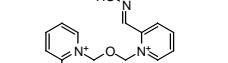
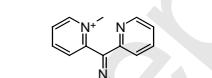
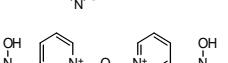
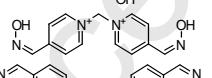
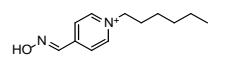
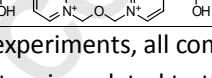
The second data set included 12 other mono/bis-pyridinium aldoxime reactivators whose reactivation potencies for VX-inhibited *r*AChE have to our knowledge not been determined. However, all the substances have already been published in the literature focused on different AChE reactivation studies. These 12 additional structures were chosen for the present work as a supplementary external set for prediction of VX-inhibited *r*AChE reactivation potency. All 29 structures and the known reactivation potencies used in the following 3D QSAR analyses are shown in Table 1. To simplify the text of the study, all compounds are generally called aldoximes, though 11, 12 and 13 are ketoximes. When applicable, aldoximes and ketoximes were differentiated.

$$RP[\%] = \left(1 - \frac{a_r - a_i}{a_o - a_i} \right) * 100 \quad (\text{Eq. 1})$$

RP – reactivation potency; *a_r* – activity of reactivated *r*AChE; *a_i* – activity of VX-inhibited *r*AChE; *a_o* – activity of intact *r*AChE

Table 1 Structures and reactivation potencies (RP) of VX- inhibited *r*AChE reactivators.

Code	Structure ^a	RP [%] ^b	Code	Structure	RP [%]
------	------------------------	------------------------	------	-----------	-----------

1		22	16		34
2		28	17		85
3		32	18		? ^c
4		39	19		?
5		4	20		?
6		70	21		?
7		18	22		?
8		71	23		?
9		60	24		?
10		87	25		?
11		9	26		?
12		9	27		?
13		3	28		?
14		59	29		?
15		79			

^aIn biochemical experiments, all compounds used in halide salt form.

^bReactivation potencies related to the effect of 10⁻³M oxime concentration.

^cReactivation potency unknown (compound used for external prediction).

2.2. 3D QSAR analysis

2.2.1. Preparation of starting molecular models

Three dimensional models of all 29 reactivators in their ionized form (e.g. charge +1 or +2, and multiplicity 1 as appropriate) were drawn in HyperChem 7.52 [41]. All oxime groups were modeled as fully protonated. To find minimum energy conformations, the structures were subjected to conformational analysis during which all flexible dihedral angles were randomly set and then

geometrically optimized using the MM+ force field (Polak-Ribiere conjugate gradient, RMS gradient threshold: 0.03 kcal/mol.Å). A self-developed TCL script for HyperChem was used for this purpose. After 1000 iterations of the random search, the conformers with the lowest potential energy (i.e. global minimum) were again geometrically optimized on DFT [42] B3LYP/6-31G* level of theory in Gaussian 09 [43] and then submitted to the next step of 3D QSAR analysis. Graphical representations of all molecular models in the paper were rendered in PyMOL [44].

2.2.2. Quenched molecular dynamics

Before 3D QSAR analysis is started, it is necessary to bear in mind that molecular models under study are to represent the available conformer population at usual laboratory temperature (i.e. 300K), and simultaneously should be well superimposable to a template structure while retaining important physicochemical properties. In ligand-based 3D QSAR studies where the conformation eliciting the biological response is unknown, one should at least generate a conformational library of all studied structures and then examine which conformers provide the best QSAR model. For this purpose, a quenched molecular dynamics procedure (QMD) was carried out in Open3DAign program [45] which used Tinker [46] as the molecular mechanics engine. The QMD protocol applying the MMFF94 force field was set to: 1) keep the initial structure obtained by DFT B3LYP/6-31G* optimization in the QMD database; 2) minimize the starting structure; 3) run a 10ps molecular dynamics period at 1000K in vacuum (step size: 0.1fs); 4) minimize the geometry until the RMS gradient reaches 0.001 kcal/mol.Å or the number of minimization iterations exceeds 1000; 5) store the conformer if its RMS distance deviation with respect to already stored conformers is greater than 0.2 Å; 6) go to point 3 and repeat this procedure 200 times; 7) discard all conformers (except the initial which corresponded to DFT B3LYP/6-31G* optimized structures) with potential energy 3.0 kcal/mol above the found minimum.

2.2.3. Molecular alignment

Being of the utmost importance in classical 3D QSAR, alignment can be divided into two computational tasks: 1) feature matching and 2) conformational search [47]. Feature matching, which can be accomplished according to several subtly different scoring schemes, serves only to superimpose two molecular models (one considered as a template and the other as a candidate) in the best way. However, due to the flexibility of molecules many conformer pairs need to be taken into consideration before satisfactory alignment is gained. Thus, for the alignment a mixed algorithm in Open3DAign program was finally utilized, combining the Pharao pharmacophore-based approach [48] with an atom-based LAMBDA-like algorithm [49]. The alignment algorithm consists of finding such geometrical orientation of template-candidate pairs where O3A score is maximal. A simplified formula of O3A score derived from the source code of Open3DAign program [50] is given in Eq. 2. In comparison to RMSD (Eq. 2), O3A score depends not only on Euclidian distance of the matching atom pairs in the candidate and template molecules but also on their electric charges. For a more precise description of O3A score the reader is referred to the literature [45].

$$O3A = \sum_{i=1}^n \left(\frac{(1.0 + \alpha_{O3A} + q_{O3A} * Q_i^{sum})}{(1.0 + Q_i^{diff})} * e^{(-\beta_{O3A} * dist_i)} \right)$$

(Eq. 2.)

$$RMSD = \sqrt{\frac{\sum_{i=1}^n [(x_{ai} - x_{bi})^2 + (y_{ai} - y_{bi})^2 + (z_{ai} - z_{bi})^2]}{n}}$$

α_{O3A} , q_{O3A} , β_{O3A} – constants; Q_i^{sum} – sum of MMFF94 charges of corresponding atoms in the template and in the aligned candidate molecule; Q_i^{diff} - absolute value of the difference between the MMFF94 charges of corresponding atoms in the template and in the aligned candidate molecule; $dist_i$ – distance between corresponding atoms in the template and in the aligned candidate molecule; x_{ai} , y_{ai} , z_{ai} – coordinates of i-th atom in the template molecule; x_{bi} , y_{bi} , z_{bi} – coordinates of i-th matching atom in the candidate molecule; n – number of found matching atom pairs.

In order to account for molecular flexibility, the QMD conformational database (see above) was processed by a multiple template/candidate batch alignment method. The three most active reactivators (10, 15, 17) were chosen as templates while all remaining candidates were aligned to each template conformer, and then picking the best fitting candidate conformers. Geometrical minimization within the aligning was done by the Jonker-Volgenant algorithm [51]. As a result, 149 overlaid molecular sets were produced. To the next computations only the alignment sets with the highest cumulative O3A score were passed.

2.2.4. Molecular interaction field calculations and PLS regression

The three best-scored alignment molecular sets superimposed on conformers of compounds 10, 15 and 17 were subsequently analyzed in Open3DQSAR program [38] using classical Coulombic and van der Waals energy molecular interaction fields (MIFs) computed by molecular mechanics method MMFF94 [52]. Having finished preliminary PLS analyses, the alignment set related to 17 was, on the basis of best achieved statistical characteristics (i.e. coefficient of determination R^2 , leave-one/two/many-out cross-validated coefficients of determination Q^2 , standard error of calculation $SDEC$ and standard error of prediction $SDEP$), selected for an extra calculation of DFT B3LYP/6-31G* electrostatic potential (ESP) MIFs in Gaussian 09. The final 3D QSAR models were developed on molecular mechanics based van der Waals (VDW) energy and DFT based electrostatic potential MIFs. As the dependent variable, untreated percentage reactivation potencies (RPs) were used.

In a more detailed view, the alignment molecular ensemble superimposed on a conformer of 17 was first placed into a 2 Å step size grid box leaving a 5 Å gap around the largest molecule. The grid box points were spaced with an alkyl carbon probe atom to calculate VDW MIFs of all 29 reactivators. Similarly, Coulomb electrostatic MIFs using the MMFF94 force field were generated, but it was only a prerequisite to test the quality of the alignment sets that were last substituted by more advanced ESP MIFs created by Gaussian 09 cubegen utility from single point B3LYP/6-31G* energy calculation checkpoint files. The checkpoint files necessary for ESP calculations were generated by g09 program, which is the main application of Gaussian 09. The rationale behind this was to introduce higher precision into the 3D QSAR analysis. Following this intention, a finer grid box with 1.0 Å spacing (29.0 x 20.0 x 21.0 Å, 13860 points) was designed for both MIFs calculations, which is especially advisable to achieve sufficiently high resolution of quantum-chemistry based ESP. The obtained ESP MIFs were multiplied by the factor 23.06035 (to convert eV into kcal/mol) and all grid points where the VDW MIFs exceeded 50 kcal/mol were zeroed in the ESP MIFs. This procedure was performed to avoid the influence of large values of ESP closest to nuclei. Accordingly, VDW MIFs were transformed so that all energies of absolute value greater than a cut-off of 30 kcal/mol had this cut-off value. Other

transformations applied to both VDW and ESP MIFs data included: 1) zeroing all values below +/- 0.05 kcal/mol threshold; 2) removing all X variables with standard deviation lower than 0.1; 3) removing all X variables with a skewed distribution assuming only 1 – 4 different levels, as recommended by Baroni et al. [53]; 4) block unscaled weighting (BUW) [54] of all remaining X variables.

On such pretreated data three different procedures were applied in order to find the most robust 3D QSAR model:

- 1) PLS regression employing 1 – 5 latent variables without any other data modifications;
- 2) smart region definition (SRD) and fractional factor design (FFD) to extract favorable variables out of the original data set; PLS regression employing 1 – 5 latent variables;
- 3) hybrid uninformative/iterative variable elimination (UVE/IVE) to rule out unimportant variables in the original data set; PLS regression employing 1 – 5 latent variables.

In all QSAR methodologies statistical analysis to discriminate chance correlations and reveal significant dependencies between the structure description data and the biological activity in mathematical terms, plays a crucial role. In the case of 3D QSAR where several thousand interrelated variables need to be processed, only multivariate regression techniques operating on orthogonally transformed data can be used. The method of choice in 3D QSAR has become partial least square (PLS) regression [55], despite the fact that a few cognate approaches may be also applicable [56,57]. A main advantage of PLS consists in allowing the pseudo β -PLS coefficients to be mapped in the molecular 3D space so as to disclose those regions most influential to the biological activity [58]. Nevertheless, PLS is well known to suffer from noise and redundancy in the input data, and therefore it is highly recommended to eliminate beforehand such variables that do not contribute to improving the stability and predictive power of the model. From several known variable pruning methods three were used in the present study: SRD, FFD and UVE/IVE.

Since the calculation of MIFs leads to a huge number of variables which are indeed arranged but not according to the spatial molecular structure, SRD was used to extract from the MIFs data matrix those descriptors related to the same spatial region and to reorganize them into groups bearing nearly the same chemical information. These variable groups reflect the spatial continuity constraints in order to make 3D QSAR analysis less prone to chance correlations and easier to interpret [59]. The SRD algorithm consists of three substantial steps: 1) selecting a set of as independent as possible X variables (seeds) bearing important information with respect to activity, 2) building Voronoi polyhedra (VP) of X variables surrounding seeds in a defined Euclidian distance, 3) collapsing of highly correlated VP into a single region. For a deeper insight into SRD the reader is kindly referred to Ref. [59]. Within the SRD procedure 3000 seeds picked by the D-optimal selection algorithm [53] (accomplished in the space of PLS weights using 5 latent variables), a critical distance of 1.0 Å for VP building and a critical collapsing distance of 2.0 Å were used.

In the next step, the dimensionality of the SRD grouped data set was reduced by FFD variable selection procedure and the resulting data set was regressed by PLS using 1 - 5 latent variables on the RPs. Basically, the FFD is aimed at selecting such SRD grouped variables as improve the model robustness (i.e. Q^2_{LMO}) in comparison to dummy variables introduced in the FFD design matrix [60].

As a criterion for accepting a variable group, the average leave-20%-out cross-validated coefficient of determination Q^2_{LMO} within 20 repetitions was used. The number of dummies was set to form 20% of the design matrix (e.g. 1005 groups of real variables, 252 dummies). To select a subset of variables with favorable impact on Q^2_{LMO} 2048 random combinations of X variable groups and dummies were tested (by Yates algorithm) [61]. After the FFD variable selection 1773 single X variables were left.

Another instrument to select significant X variables is a hybrid UVE/IVE algorithm, which combines approaches developed by Centner [62] and Gieleciak [63]. The principle behind the UVE method is to exclude those variables whose c_j ratio of the β -PLS coefficient median to the interquartile range of Q^2_{LMO} (resulting from 20 runs of leave-20%-out cross-validation) is lower compared to a user-defined percentile of the sorted C_d value vector corresponding to dummy variables obtained from the PLS analysis (Eq. 3) [64].

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}^{PLS} + \mathbf{E} \quad c_j = \left| \frac{\hat{\beta}_j^{PLS}}{Q_3^{PLS} - Q_1^{PLS}} \right| \quad (\text{Eq. 3})$$

\mathbf{Y} – matrix of biological activities; \mathbf{X} – matrix of descriptors; $\boldsymbol{\beta}^{PLS}$ – matrix of pseudo β -PLS coefficients; \mathbf{E} – matrix of residuals; c_j – relative magnitude of j-th $\hat{\beta}_j^{PLS}$; $\hat{\beta}_j^{PLS}$ – median of j-th pseudo β -PLS coefficients obtained by leave-20%-out cross-validation; Q_3^{PLS} – Q_1^{PLS} – interquartile range of pseudo β -PLS coefficients.

In this study however, an iterative variable elimination (IVE) algorithm with a threshold of 50% maximally removed variables was employed. Contrary to the UVE algorithm, in a UVE/IVE hybrid algorithm the X variables are not rejected according to c_j values calculated in one pass but an X variable having the lowest c_j in an iteration is sequentially excluded from the X matrix until a user-defined percentage (i.e. 50% in this study) of excluded variables is reached. After the iterative procedure was completed, the model yielding highest Q^2_{LMO} was chosen and the pre-selected X variables were definitely removed from the dataset. Within the UVE/IVE procedure, we employed PLS with 1 - 5 latent variables, leave-20%-out cross-validation in 20 repetitions, UVE β -PLS medians and interquartile ranges for c_j computation and 90-percentile of the dummy sorted C_d vector. The UVE/IVE selection extracted 8170 important variables out of the X matrix. As only a concise description of used algorithms can be given here, for detailed information the reader is referred to the cited literature.

2.2.5. Validation of 3D QSAR models

Without a proper statistical validation no model can be reliably used for biological activity interpretation and prediction and thus considered true. Recently, from several validation techniques recommended for QSAR analyses external test set prediction has become generally the most accepted [65]. Therefore, the initial structure set in the study containing 17 reactivators was extended by 12 previously synthesized and published reactivators whose RPs for VX-inhibited rAChE

was unknown. These additional structures were chosen to belong to the same chemical space as the training structures. Their RPs for VX-inhibited rAChE will be determined later and the *in vitro* results confronted with the *in silico* values predicted within the present study. However in order to evaluate the stability and predictivity of the 3D QSAR models in terms of external coefficient of determination R^2_{ext} , 4 (~25%) randomly chosen reactivators with known RP (5, 8, 9, 16) had to be transferred to the external set. Besides the external coefficient of determination R^2_{ext} , the 3D QSAR models under investigation were evaluated also by internal coefficient of determination R^2 , cross-validated coefficient of determination Q^2 computed with leave-one/two/many-out approaches, standard error of calculation $SDEC$, and standard error of prediction $SDEP$ (Eq. 4). As for the leave-many-out cross-validation, the training set was randomly split into five groups and the RPs in each groups were predicted by the model resulting from PLS regression on the 4 groups remaining (corresponding to 20%-leave out cross-validation). After 50 repetitions the average Q^2_{LMO} was calculated. The 3D QSAR models with highest Q^2 and R^2_{ext} were subjected to a progressive Y-scrambling test to challenge their robustness [66]. In the scrambling tests, 300 shuffled Y vectors, starting with distribution of their elements into 4 bins (quantiles), were tested by leave-20%-out cross-validation (within 50 runs).

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_i^n (y_i^{real} - y_i^{calc})^2}{\sum_i^n (y_i^{real} - \bar{y}^{real})^2}; SDEC = \sqrt{\frac{\sum_i^n (y_i^{real} - y_i^{calc})^2}{n - LV - 1}}$$

(Eq. 4)

$$Q^2 = 1 - \frac{\sum_i^n (y_i^{real} - y_{i(-1)}^{pred})^2}{\sum_i^n (y_i^{real} - \bar{y}^{real})^2}; SDEP = \sqrt{\frac{\sum_i^n (y_i^{real} - y_{i(-1)}^{pred})^2}{n - LV - 1}}$$

Here, y_i^{real} denotes the observed RP for i-th sample, y_i^{calc} denotes the calculated RP for i-th sample, $y_{i(-1)}^{pred}$ denotes the RP for i-th sample calculated by 3D QSAR model built without the i-th sample (or k samples in the case of leave-two- or many-out methods) in the training process, and LV denotes the number of latent variables included in the model. SSE and SST stand for the error sum of squares and the total sum of squares, respectively. If $SST = SSE + SSR$ (SSR means regression sum of squares $\sum_i^n (y_i^{calc} - \bar{y}^{real})^2$), then $R^2 = SSR/SST$. The meaning of the other statistical terms is as usual.

A necessary condition of a valid regression model is concerned with the coefficient of determination R^2 . In an optimal model R^2 is close to 1.0, while in models to be rejected R^2 approaches 0.0 or may be even negative in some cases [67, 68]. Coefficient of determination is a measure of the quality of fit between real and calculated values. It estimates the portion of y^{real} variation explained by the model, but it does not describe the mathematical ability of the model to predict the end point variable for data outside the training set. Internal statistical criteria similar to R^2 are therefore insufficient to express the validity of the model. Principally, the only way to evaluate the predictivity of a QSAR model is by external validation [69, 70]. However, the cost of such rigorous validation is a loss of valuable information from the training set.

A flow chart outlining the major steps in building the models is given in Fig. 5.

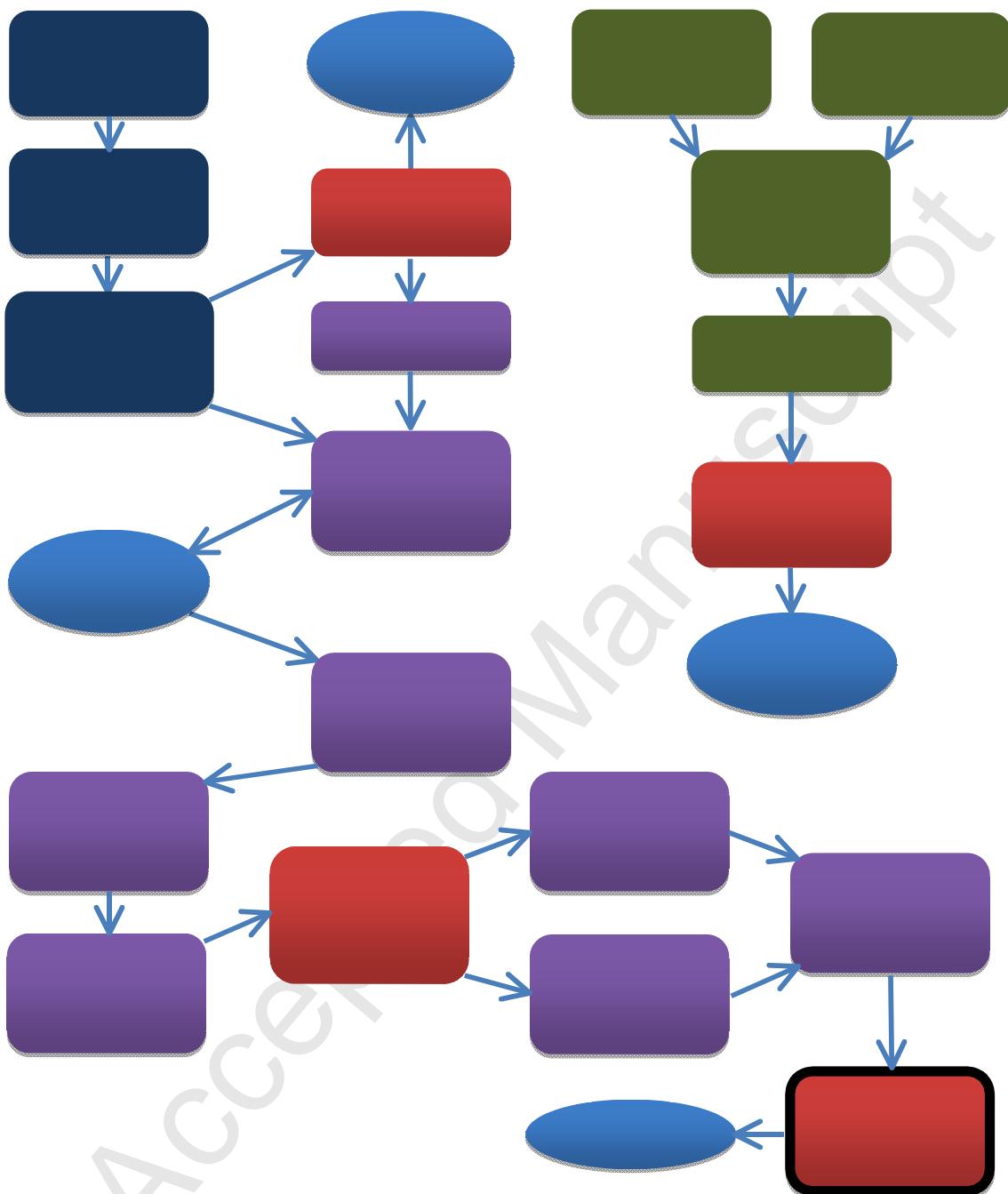


Figure 5. Major steps in building the 3D QSAR models.

3. Results and discussion

3.1. Molecular interaction field calculations

29 molecular models of mono-/bis-pyridinium aldoxime reactivators were subjected to QMD to obtain conformational libraries (29×200 models). From all possible combinations, an alignment set superimposed on a conformer of compound 17 was selected and used for MIF calculation (Fig. 6). According to our previous 3D QSAR analysis (not published), a combined “pharmacophore-atom” alignment method was chosen as generally the most convenient in the sense of statistical significance of the resulting 3D QSAR models. This alignment set, referenced to the second most active reactivator in the study, also gained the highest O3A score ($O3A(17) = 3213.95$) in Open3DAlign program. Cumulative RMSD of this alignment set referenced to 17 conformer template was 9.60 \AA . Individual contributions of all candidate-template pairs to the cumulative $O3A(17)$ score and RMSD(17) are given in Table S1 (Supporting Information).

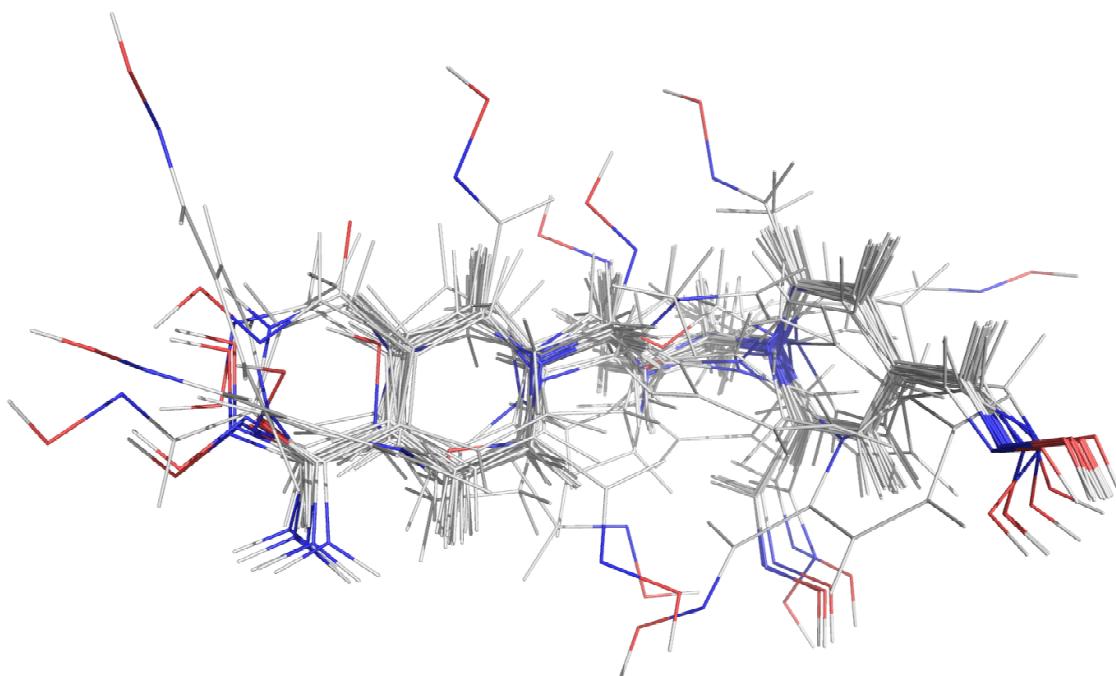


Figure 6. Alignment set of 29 reactivator models superimposed on a 17 conformer template. A combined Pharao-pharmacophore and atom based alignment algorithm implemented in Open3DAlign program was employed. Cumulative $O3A(17)$ score = 3213.95, cumulative $\text{RMSD}(17)$ = 9.60 \AA . Carbon – white, nitrogen – blue, oxygen – red, hydrogen – white.

The alignment molecular set was placed into a 1\AA -spaced gridbox and VDW MIFs were generated in Open3DQSAR program. In a similar manner, ESP MIFs of all 29 compounds were extracted from cube files obtained by single point B3LYP/6-31G* computations. For illustration, interpolated VDW MIF and ESP MIF of the 17 conformer template are depicted in Fig. 7.

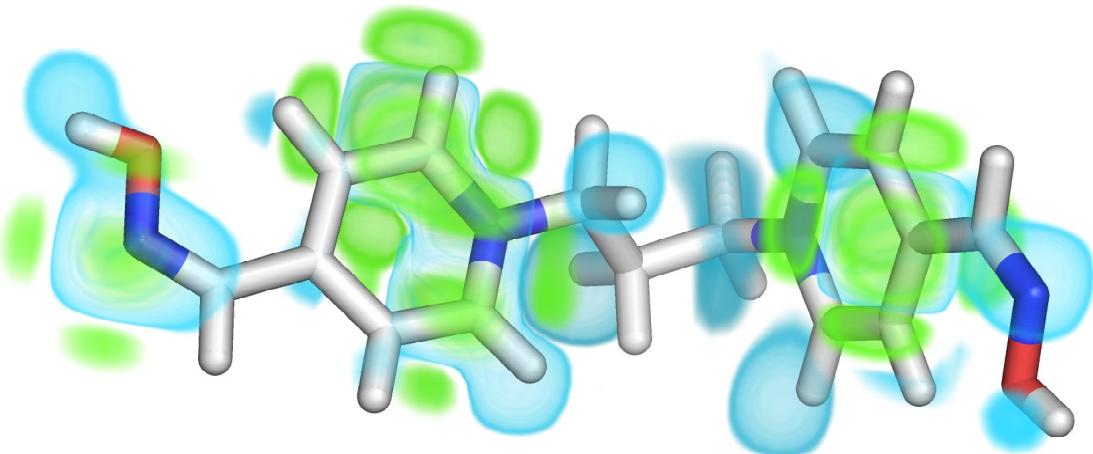
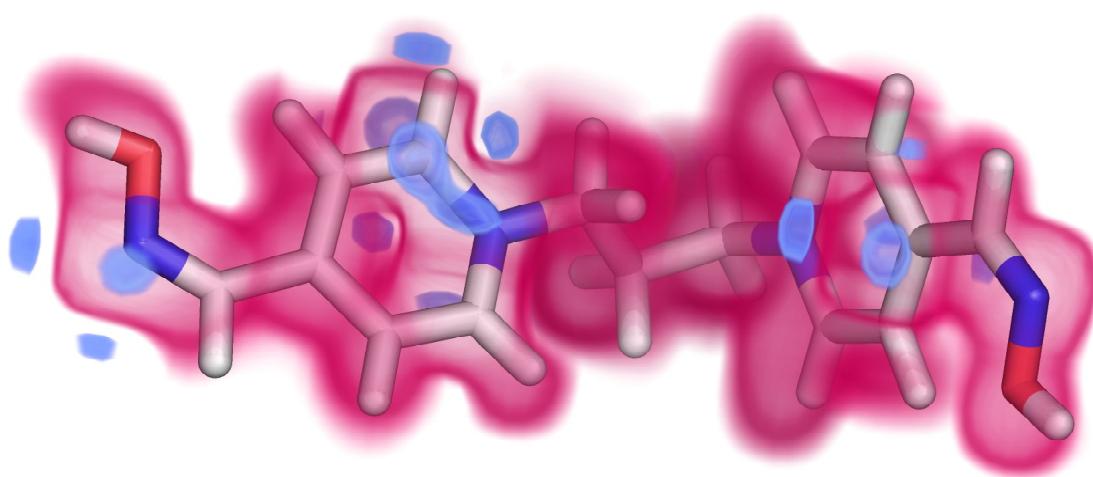
A**B**

Figure 7. Colored volume representations of VDW MIF (A; higher energy – teal, lower energy – green) and ESP MIF (B; higher potential – magenta, lower potential – blue) for the template conformer of compound 17.

It is noteworthy that both VDW and ESP MIFs reflecting the electro-topological properties of the studied molecules have to be more or less correlated as they are derived from the same structures. On the other hand, a high degree of similarity between the fields introduces redundant information into the data increasing the risk of biased statistical evaluation and interpretation. A common

measure which quantifies the correlation between the MIFs is the Tanimoto similarity index T (Eq. 5). Here, X_i and Y_i stand for the MIFs values at the i-th grid point.

$$T(x, y) = \frac{\sum_i^n X_i Y_i}{\sum_i^n X_i^2 + \sum_i^n Y_i^2 - \sum_i^n X_i Y_i} \quad (\text{Eq. 5})$$

The value of T for similar fields is close to 1.0, whereas for completely different fields it is close to 0.0. For 29 pairs of original VDW and ESP MIFs for the studied compounds these T are representative: $T(\max) = 0.0253$, $T(\min) = 0.0115$ and $T(\text{average}) = 0.0182$. Accordingly, it is obvious that the original MIFs with 13860 components contain rather different information.

However, the original MIFs were adjusted by several techniques to remove information noise. An overview of MIFs and performed modifications is given in Table 2.

Table 2 Overview of MIFs.

Adjustment	VDW MIFs			ESP MIFs		
	No. of X var.	Min	Max	No. of X var.	Min	Max
Original data set ^a	13860	-1.04	7.63×10^{10}	13860	0.51	468.90
Pretreated ^b	1796	-1.04	30.00	13500	0.00	5.78
SRD/FFD ^c	1107	-1.04	30.00	666	0.00	5.78
UVE/IVE ^c	427	-1.04	30.00	7743	0.00	5.78

^aIn the original data set, ESP MIFs were multiplied by factor 23.06035; ^bdata pretreatment included: exclusion of high energy gridpoints closest to nuclei, setting cutoffs, zeroing, exclusion of X variables with low variance and assuming only several levels (see Material and methods); ^cSRD/FFD and UVE/IVE were done in parallel on the pretreated data.

3.2. Alignment set based on flexible molecular docking

Since the construction of the molecular alignment ensemble is a crucial point in 3D QSAR analysis, ligand conformers resulting from flexible molecular docking were also explored to this end. All 29 reactivators were docked in the active site of VX-phosphonylated *m*AChE (from mouse) model (PDB: 2U2Y, resolution: 2.60 Å) by AutoDock Vina 1.1.2 [71]. A model of AChE enzyme isolated from rat was not available in the online PDB database (www.pdb.org). 40 amino acid residues within a sphere of $R = 10$ Å around the phosphorus atom in the VX-*m*AChE complex were set as a flexible enzyme moiety (grid box size: 20 × 20 × 20 Å; center of grid box: 28.21, 15.53, 12.40; exhaustiveness = 9). Each ligand was docked five times in the VX-*m*AChE active site (always using default random seed in the computations) and the lowest energy binding modes were stored in an sdf file. However, this molecular docking based alignment set (Fig. 8) exhibited relatively low O3A(17) score = 459.01 and too large RMSD(17) = 145.76 Å (i.e. taking compound 17 as a reference template). Even the initial molecular set provided better values of these criteria (O3A(17) score = 712.11, RMSD(17) = 138.36 Å, see Table S1 in Supporting Information). High cumulative O3A score and RMSD of the docked molecular set is likely a consequence of searching the global energy minimum on a complex hypersurface of potential energy related to flexible ligands in the *m*AChE binding site delineated with flexible residues. Due to manifold interactions in the *m*AChE binding site a ligand having aromatic cycles may be docked toward more than one aromatic acid residue and accordingly take up different positions. Thanks to different surroundings and forces within flexible docking it is not surprising that

the alignment of QMD conformers, ruled only by a specific scoring function (i.e. O3A score), resulted in a different geometrical arrangement. Since high disorder of the alignment set considerably disturbs the 3D QSAR calculation, the alignment set produced by Open3DAlign program was preferred, despite the fact that docked ligand structures may contain very useful information concerning the binding modes in the receptor.

Undoubtedly, structure based QSAR analysis of OP-inhibited AChE is worthy of detailed research. So far, our preliminary analyses of the top-scoring binding modes in VX-mAChE have revealed that the aldoxime functional group is directed toward VX-Ser203 in 13 reactivators (3, 6, 8, 9, 10, 15, 17, 22, 23, 24, 25, 26, 29). In the remaining 16 ligands the aldoxime group has been found in approximately the opposite direction, pointing out of the mAChE gorge. The estimates of binding energy fluctuated from -9.5 kcal/mol to -5.0 kcal/mol. It is interesting that the lowest binding energies were calculated for compounds 10 (-9.5 kcal/mol), 23 (-9.5 kcal/mol), 24 (-9.4 kcal/mol) and 25 (-9.4 kcal/mol) all of which have very high reactivation potency. These results seem to be in good agreement with Pang's hypothesis on dual-binding reactivators (see Chapter 3.3.). However, further studies have to be carried out in order to elucidate the relationship between *in silico* binding energies of mono-/bis-pyridinium aldoxime reactivators and their reactivation potency.

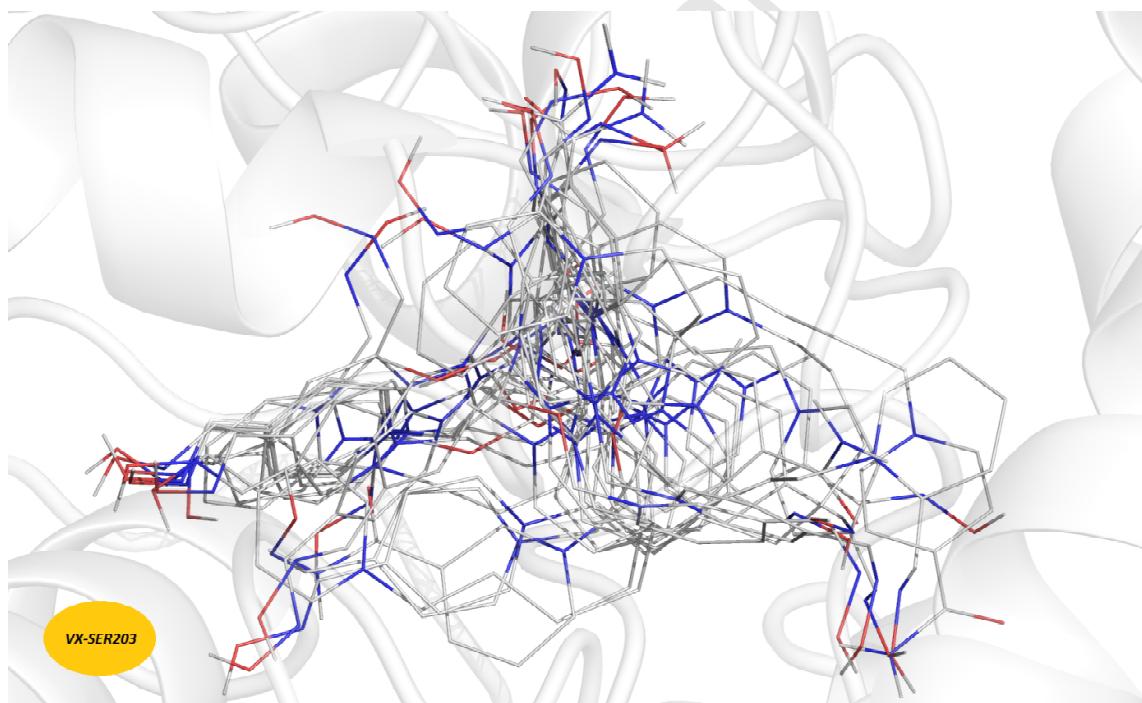


Figure 8. Alignment set of 29 reactivator models resulting from flexible molecular docking in the VX-mAChE active site (PDB: 2U2Y). Cumulative O3A(17) score = 459.01, cumulative RMSD(17) = 145.76 Å. Carbon – white, nitrogen – blue, oxygen – red, polar hydrogen – white.

3.3. PLS analyses and predictions

On the variously adjusted input data set regular PLS analyses using 1 – 5 latent variables (*LV*) and LOO/LTO/LMO cross-validations were performed in Open3DQSAR program. The predictivity of 3D

QSAR models were estimated by coefficient of determination R^2_{ext} and $SDEP_{ext}$ on an external test set of four randomly chosen compounds. As only 17 reactivators with known RPs were available in the study, the 3D QSAR models were developed on 13 samples. Another 12 reactivators chosen to belong to the same chemical structure class served for supplementary prediction purposes. The distribution of the compounds over chemical space projected by PCA of the pretreated data set into two dimensions is displayed in Fig. 9. A 9-membered group of points with negative values of PC1 (on the left of Fig. 9) involves substances containing only a single pyridinium aromatic ring with one aldoxime group (18, 21, 19, 20, 11, 29, 16), or an analogous ketoxime substituted by another aromatic ring on the imino carbon (12, 13). The remaining points in the positive sector of the PC1 axis represent 20 substituted bis-pyridinium aldoximes. The more or less outlying points 1 and 5 probably express a unique combination of molecular features (e.g. double bond in the linking chain between pyridinium rings and/or the aldoxime function in the *ortho* position). However, the precise interpretation of these outlying points would require more investigation than could be given in the present study. The most important conclusions from this PCA analysis are that the three compound sets, especially the supplementary one, occupy very similar chemical space, and that the quantitative predictions deduced from it will not encompass all extrapolations. Additionally, PCA analysis confirmed that information included in the training and test sets (except for compounds 5 and 1) was not significantly different, and the 3D QSAR models built on 13 samples should be representative for all studied reactivators. The results of PLS and cross-validations of the original, pretreated, SRD/FFD and UVE/IVE adjusted data sets are summarized in Tables S2 – S9 (Supporting Information).

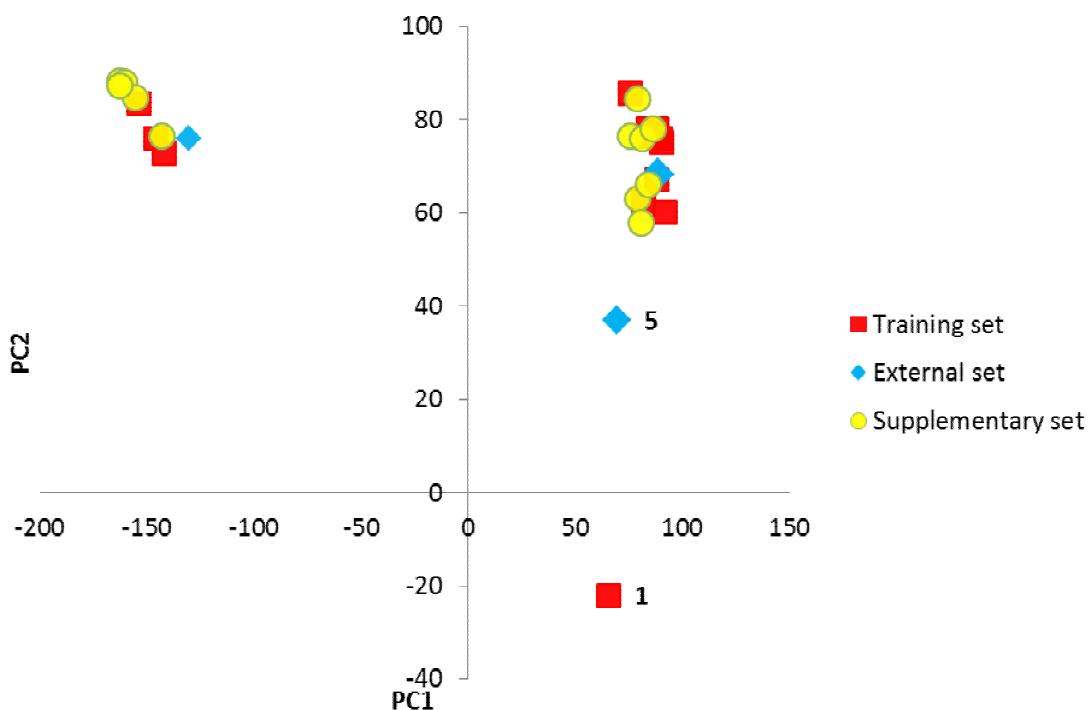


Figure 9. Chemical space represented by PCA of the pretreated data set. The PC1 and PC2 accounted for 38.6% and 17.6% of the variance in the data, respectively.

According to statistical criteria given in Tables S2 – S9 (Supporting Information), the SRD/FFD adjusted data set showed the best external predictivity ($R^2_{ext} = 0.7606 - 0.9331$), whereas UVE/IVE corrected input showed excellent stability in cross-validation ($Q^2_{LMO} = 0.4870 - 0.9248$). However, their stability in cross-validation is understandable since both SRD/FFD and UVE/IVE data noise reduction methods are controlled in selecting significant X variables by the increment of Q^2_{LMO} . Although the 3D QSAR model built on the original data set exhibited quite good external predictivity ($R^2_{ext} = 0.4766 - 0.8217$), in cross-validation tests it completely failed ($Q^2_{LMO} = -0.4258 - 0.2364$). Analyses of the pretreated data set likewise led to very satisfying 3D QSAR models, though they did not surpass the models obtained *via* SRD/FFD or UVE/IVE. Except for the 3D QSAR model based on the original data, introducing more LV generally increases the stability and predictivity of the models. However, it still remains questionable how many LV should then be kept in the model. A compromise balancing the internal stability and external predictivity seems to be the best approach to the optimal number of LV. In the above cases, 3D QSAR models with 4 LV may be considered as such an optimum middle way exhibiting the lowest fitting error. Summarily, it is obvious that most statistically significant results were provided by SRD/FFD and UVE/IVE procedures for noise reduction in the input data. As both results are well comparable, it is not possible to prefer one to another. The RPs of all 29 reactivators calculated by 3D QSAR analyses of SRD/FFD and UVE/IVE adjusted data sets using 1 – 5 LV are given in Table 3.

Table 3 Predicted RPs by SRD/FFD and UVE/IVE 3D QSAR models.

Set	Comp.	Real RP [%]	Calculated RP [%]									
			LV – SRD/FFD					LV – UVE/IVE				
			1	2	3	4	5	1	2	3	4	5
Training	1	22	32.7	19.8	21.1	23.5	21.6	45.3	22.3	19.2	22.9	21.9
	2	28	38.6	32.5	27.4	28.6	28.9	46.5	34.5	30.5	28.6	29.3
	3	32	52.9	40.2	34.6	32.1	31.9	52.1	36.8	36.0	34.8	33.0
	4	39	44.7	34.3	37.7	37.1	38.4	49.8	31.0	38.0	37.9	37.9
	6	70	65.2	67.6	65.9	68.0	69.1	57.3	66.4	64.4	64.9	67.3
	7	18	27.1	13.9	19.4	18.3	18.6	45.0	16.4	17.8	16.7	17.5
	10	87	76.3	87.8	90.0	87.6	87.0	63.5	83.1	87.6	88.9	89.5
	11	9	-0.8	9.7	9.2	9.0	9.1	-0.1	9.4	9.9	9.8	9.5
	12	9	1.6	10.1	9.2	8.9	9.0	-0.1	7.9	8.7	8.6	9.2
	13	3	-4.0	2.4	2.8	2.9	2.7	-2.4	3.2	2.6	2.8	2.3
	14	59	57.9	60.8	59.5	58.9	58.9	56.3	62.3	59.4	59.4	60.7
External	15	79	73.9	79.9	80.2	79.9	78.8	62.9	81.5	80.5	79.7	77.5
	17	85	74.1	81.2	83.1	85.4	85.8	63.9	85.3	85.3	85.1	84.4
	5	4	25.4	14.3	17.6	18.7	19.0	43.5	13.0	13.8	15.0	16.0
	8	71	59.3	59.1	57.3	58.8	59.3	55.7	60.4	57.5	57.4	59.0
Supplement	9	60	62.3	64.1	63.0	64.0	63.9	58.8	68.3	65.7	65.2	65.5
	16	34	23.4	32.2	33.9	34.6	34.8	7.1	23.0	28.2	29.2	28.3
	18	- ^a	34.2	43.4	40.4	39.2	39.1	10.8	32.5	30.3	30.2	28.5
	19	-	35.5	45.2	42.9	42.0	42.0	11.1	33.0	31.0	31.0	29.5
	20	-	34.5	43.9	42.0	41.6	41.7	11.2	33.4	31.5	31.4	30.1
	21	-	33.5	41.8	38.3	37.2	36.8	10.7	30.9	27.6	27.0	24.5
	22	-	38.9	42.8	44.3	45.6	45.5	48.3	48.6	49.3	49.6	52.4
	23	-	67.1	73.2	73.8	74.3	74.2	59.7	72.8	72.4	73.3	73.6

24	-	70.8	76.9	78.5	78.2	78.0	63.2	80.2	83.1	84.0	83.2
25	-	71.5	78.3	79.1	79.3	79.4	63.9	80.7	82.7	83.0	81.8
26	-	58.2	61.0	60.7	62.2	63.0	57.9	66.2	64.0	64.4	66.6
27	-	29.6	18.1	18.0	15.9	16.4	46.0	19.6	22.1	21.2	21.3
28	-	69.7	77.0	78.4	79.1	79.2	62.4	81.0	82.7	83.1	82.7
29	-	35.2	43.7	42.2	42.3	42.6	11.0	33.3	31.4	31.3	29.9

^aReactivation potencies in the supplementary set are unknown.

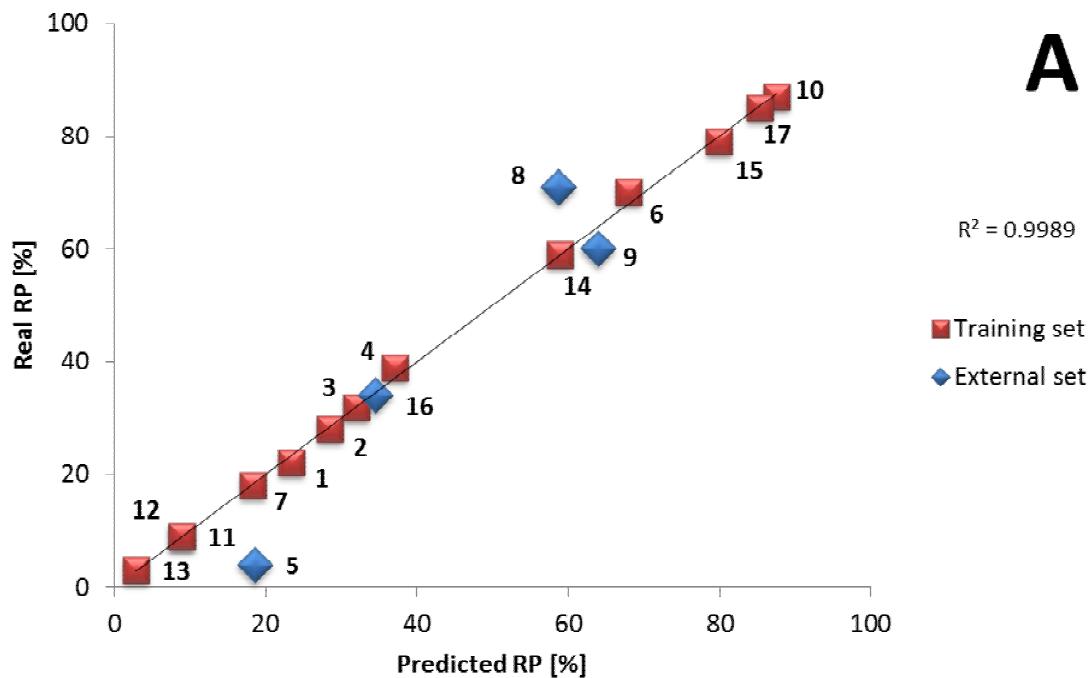
From Table 3 it is clear that within the supplementary set the highest reactivation potencies were predicted for 24 (SRD/FFD – 4LV model: 79.3%) and 25 (UVE/IVE – LV4 model: 84.0%). High reactivation potencies were also calculated for 23 (74.3%; 73.3%), 26 (62.2%; 64.4%) and 28 (79.1%; 83.1%). Although none of these newly predicted RPs exceeded the experimental RP of 17 (85%) or 10 (87%) present in the training set, it is necessary to bear in mind that both experimental and calculated values involve some degree of uncertainty. While the precision and accuracy of experimental RPs used in the present study is unknown, $SDEP_{ext}$ of the 3D QSAR models with 4 LV is approximately 7% of RP. Thus, the highest RPs predicted in the supplementary set might be considered *bona fide* similar to those in the training set.

The structures of 24 and 25 both contain two *p*-aldoxime pyridinium moieties connected by an oligomethylene chain with 8 and 5 carbons, respectively. Thus, the reason for the high RPs predicted undoubtedly resides in the structural similarity of 24 and 25 to 17 which is also built up of two *p*-aldoxime pyridinium moieties but connected by a trimethylene linker. On the other hand, the lowest RP was predicted for 27 (15.9%; 21.2%) which has two aldoxime functions bound in the *ortho* position on the bis-pyridinium scaffold. Consistent with this, compound 5 containing two aldoxime groups in *ortho* positions on a similar scaffold reactivated only 4% of the original rAChE activity *in vitro*. Within the set of 29 VX-rAChE reactivators studied, the *p*-aldoxime substituted bis-pyridinium scaffold seems to be a more potent pharmacophore than *ortho* isomers. Since the *para* position of the aldoxime group in the bis-pyridinium structure is sterically less hindered, the formation of the transition state within the reactivation process may proceed more easily, increasing the RP. Furthermore, from comparison of the predicted RPs for 22, 23, 24 and 25 it is obvious that extending the length of the linker between the two *p*-aldoxime pyridinium moieties generally boosts the activity.

Since within the framework of ligand-based 3D QSAR methodology we cannot infer how the length of linker influences the ligand settling in the AChE binding sites, the strengthening of RPs in these reactivators has to be judged as a result of alignment of the candidates to the template so that their calculated MIFs may be mathematically transformed by the 3D QSAR model to relatively high RP values. The fact that bis-pyridinium aldoximes with a sufficiently long linker may interact simultaneously in CAS and PAS of AChE is an important clue in the rational design of novel dual binding reactivators of OP-inhibited AChE [72], but deducing this conclusion is beyond the methodology of ligand-based 3D QSAR. The rationale for design of dual-binding reactivators of OP-inhibited AChE was inferred from a structure-based model and expressed by Pang et al.: *reactivators with increased affinity for the active site of the phosphorylated AChE accelerate the rate of enzyme dephosphorylation* [32]. The statement was later adopted by Hammond et al. who further generalized that increasing the linker length of bis-pyridinium aldoximes led to strengthened affinity for inhibited AChE, but decreased the rate constant of reactivation [73]. Unfortunately, the

relationship between the binding affinity, length of the linker and reactivation potency of bis-pyridinium aldoximes toward inhibited AChE has not been proved by any statistically significant QSAR model. From 3D QSAR models found in this paper we might confirm the favorable influence of the linker length on reactivation, though the principle of it is not based on the nature of intermolecular interactions in the binding site of AChE.

A comparison between the prediction by SRD/FFD and UVE/IVE based 3D QSAR models using 4 LV is shown in Fig. 10.



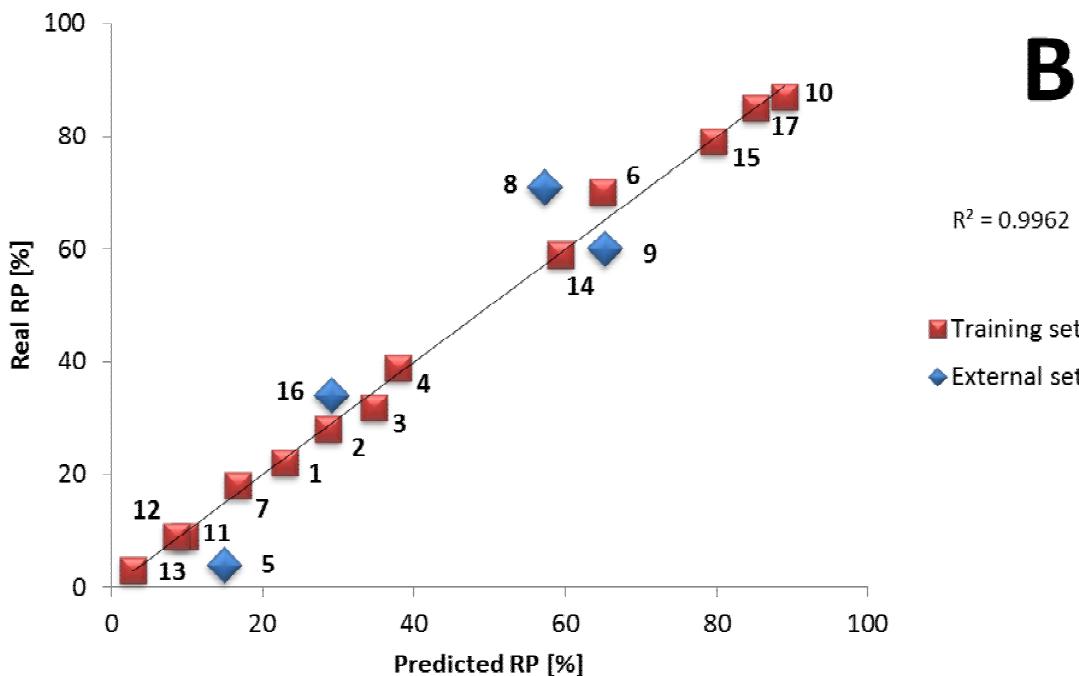


Figure 10. Prediction by 3D QSAR models resulting from PLS ($LV = 4$) of SRD/FFD (A) and UVE/IVE (B) adjusted data sets. R^2 stands for the coefficient of determination of the real and predicted RPs for the compounds within the training set.

As can be seen above, both SRD/FFD and UVE/IVE 3D QSAR models demonstrated statistically significant internal as well as external predictions. Relying on these criteria, the chosen 3D QSAR methodology employing SRD/FFD or UVE/IVE may be accepted as a suitably sophisticated procedure for removing abundant information from the input data matrix.

Stability of the 3D QSAR models was finally analyzed by progressive Y-scrambling. The results proving sufficient stability/sensitivity of the models can be found in Supporting Information.

3.4. Maps of significant regions in MIFs

Another attractive feature of 3D QSAR models, besides the possibility to predict unseen RPs, is the visualization of the results as spatial contour maps [74]. The contour plots indicate which regions in MIFs of the molecules included in the training set contribute significantly to increasing or decreasing the RP. These maps are calculated as scalar products of pseudo β -PLS coefficients and standard deviations (SD), associated with each vector in MIFs. Fig. 11 and 12 show stereo color maps of pseudo β -PLS x SD for ESP and VDW MIFs, respectively, which were obtained through PLS analysis of the SRD/FFD adjusted data set. The pseudo β -PLS x SD coefficients with positive sign are depicted in red, whilst negative values are colored in teal (for ESP MIFs, Fig. 11) or yellow (for VDW MIFs, Fig.

12). As the vector elements in the MIFs and pseudo β -PLS x SD coefficients generally have a sign, a region in the molecule increases the RP only if the product $x_i^{MIF}\beta_i$ is positive and preferably of a high value. In the pseudo β -PLS x SD coefficients contour map associated with ESP MIFs (Fig. 11), the majority of the red area is localized along the rim of the molecular alignment set, suggesting that molecules aligning to this border by higher substituent ESP are able to elicit a stronger RP. The size requirement is fulfilled mainly by the *p*-substituted bis-pyridinium reactivators which are larger than mono-pyridinium derivatives. As regards the molecular features which diminish the RP, the distribution of teal-colored regions in the pseudo β -PLS x SD coefficients contour map associated with ESP MIF cannot be interpreted easily. However, the low value of the pseudo β -PLS x SD coefficients isosurface (-0.001) suggests that the teal-colored regions in the ESP MIFs do not influence the RP dramatically.

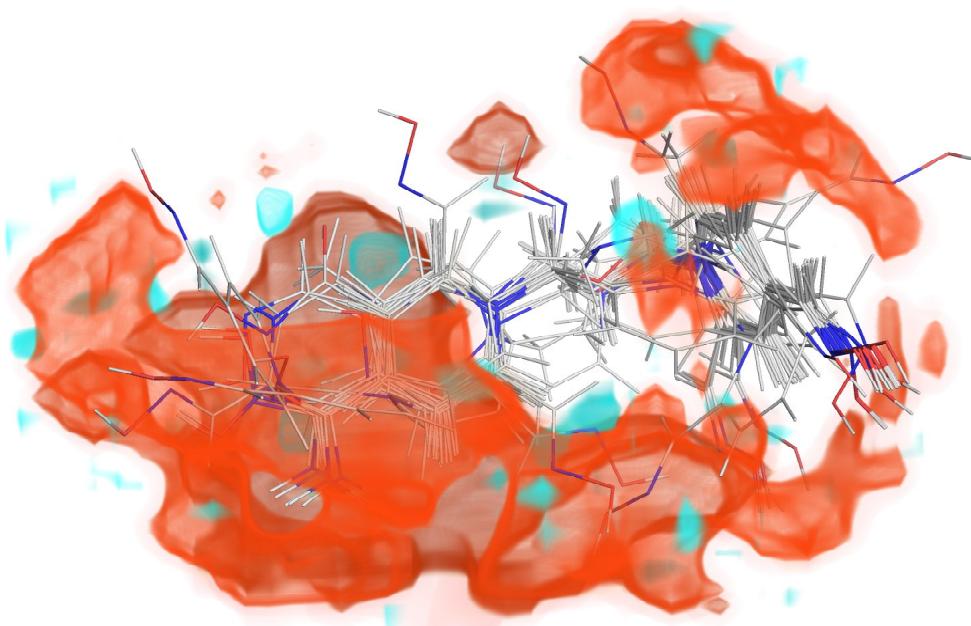


Figure 11. A color contour map of pseudo β -PLS x SD coefficients within ESP MIFs resulting from 3D QSAR analysis of SRD/FFD adjusted data set. Red volume represents positive pseudo β -PLS x SD coefficients (isosurface level: 0.008), while teal symbolizes areas where these values are negative (isosurface level: -0.001).

A better resolution of pseudo β -PLS x SD coefficients was obtained for VDW MIFs (Fig. 12) where the red regions improving the RP contour cluster around the substituents in the *para* position in the bis-pyridinium scaffold, whilst the yellow colored regions detrimental to RP are situated closer to the *p*-amido amino group and to substituents in the middle part of the alignment set. Therefore, according to the pseudo β -PLS x SD coefficients contour map related to ESP and VDW MIFs the RP is enhanced in compounds which are of a larger size (e.g. in bis-pyridinium derivates) having preferably an aldoxime function in the *para* position in the aromatic rings. Substitutions in the linker connecting the pyridinium rings as well as the presence of substituents in *ortho* and *meta* positions probably decrease RP. Visual analysis of individual VDW MIFs adjusted by the SRD/FFD noise reduction

method revealed that relatively high positive VDW energy regions were preserved close to functional groups. Since the majority of the compounds studied contain *p*-aldoxime groups it might generalized that high positive VDW energy, which is associated with high repulsive force, remaining around *p*-aldoxime groups is beneficial to RP.

Noteworthy, in the majority of RP predictions calculated by SRD/FFD 3D QSAR model the relative contributions of VDW MIFs are greater than those of ESP MIFs (see Table S10, Supporting Information). Higher contributions of ESP MIFs (48.4 – 79.1%) to RP predictions were observed for compounds 16, 18, 19, 20, 21, 22 and 29. These compounds are mostly of a low, but not the lowest, or medium RP and contain predominantly a mono-pyridinium aldoxime scaffold.

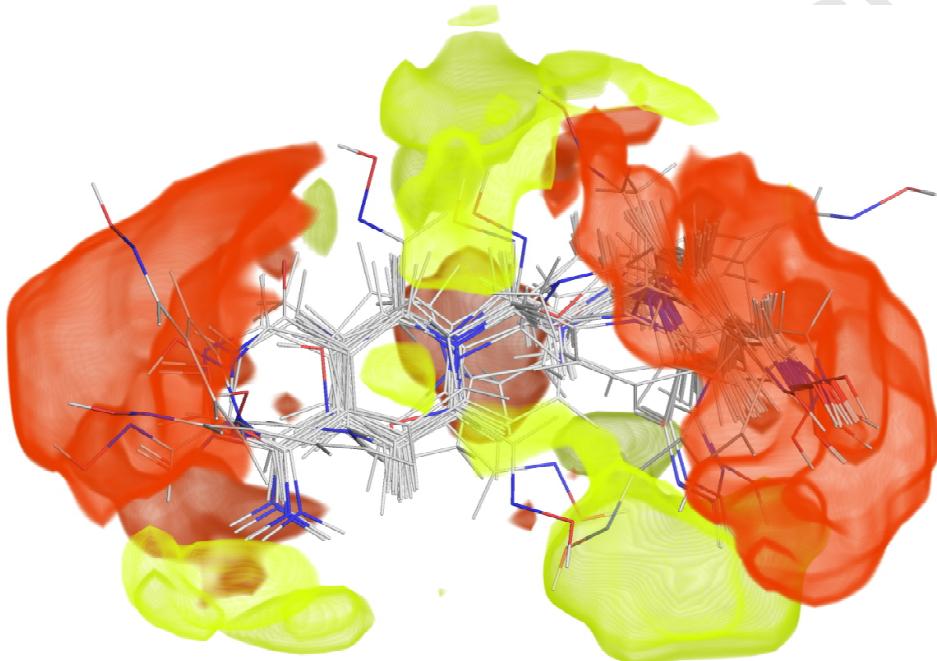


Figure 12. A color contour map of pseudo β -PLS x SD coefficients within VDW MIFs resulting from 3D QSAR analysis of SRD/FFD adjusted data set. Red volume represents positive pseudo β -PLS x SD coefficients (isosurface level: 0.02), while yellow symbolizes areas where these values are negative (isosurface level: -0.022).

Similar features can be observed in the contour plots describing the results of PLS analysis of ESP (Fig. 13) and VDW MIFs (Fig. 14) adjusted by the UVE/IVE method. The meaning of red, teal and yellow colored regions is the same as in the previous contour maps. In Fig. 13, the red contours around *p*-aldoxime functions on the left and right side of the bis-pyridinium scaffold suggest a high significance of these substituents for strengthening the RP. In contrast to Fig. 12, a beneficial influence on the RP is ascribed also to a higher ESP of substituents (i.e. aldoxime groups) occupying the *ortho* position in the bis-pyridinium skeleton. A small red region in the proximity of the NH_2 group indicates that increased ESP in this place is favorable for the RP. A lower level of ESP is preferred only in the right bottom part of the alignment set. However, this teal area with isosurface value -0.002 covering the

red region close to the aldoxime groups in the *ortho* position does not contribute significantly to the resulting RP.

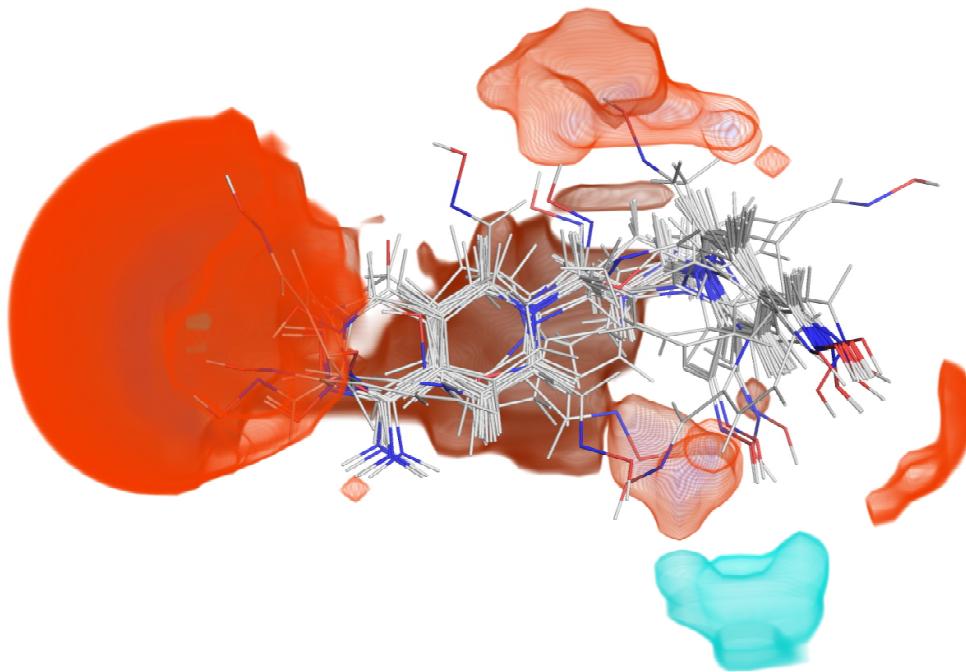


Figure 13. A color contour map of pseudo β -PLS x SD coefficients within ESP MIFs resulting from 3D QSAR analysis of UVE/IVE adjusted data set. Red volume represents positive pseudo β -PLS x SD coefficients (isosurface level: 0.004), while teal symbolizes areas where these values are negative (isosurface level: -0.002).

The contour plot of pseudo β -PLS x SD coefficients in Fig. 14 confirms that a higher VDW energy in the red regions around the *p*-aldoxime function enhances the RP. The regions where decreased VDW energy is favorable for the RP are indicated in yellow. Two yellow regions occur around functions bound in *ortho* or *meta* positions. Another small yellow region marks the VDW MIF close to the NH_2 group. Thus, if the VDW energy is decreased in the yellow-colored areas, the RP may be improved. Nearly the same conclusions may be deduced from the contour map of pseudo β -PLS x SD coefficients associated with VDW MIFs (Fig. 12) provided by PLS analysis of SRD/FFD adjusted data set. However, it is still necessary to keep in mind that VDW and ESP MIFs in both SRD/FFD and UVE/IVE 3D QSAR models were subjected to noise reduction which removed many MIF vectors from the final PLS analysis. A comparison of VDW and ESP MIF contributions to the RPs predicted by the UVE/IVE 3D QSAR model is given in Table S11 (Supporting Information).

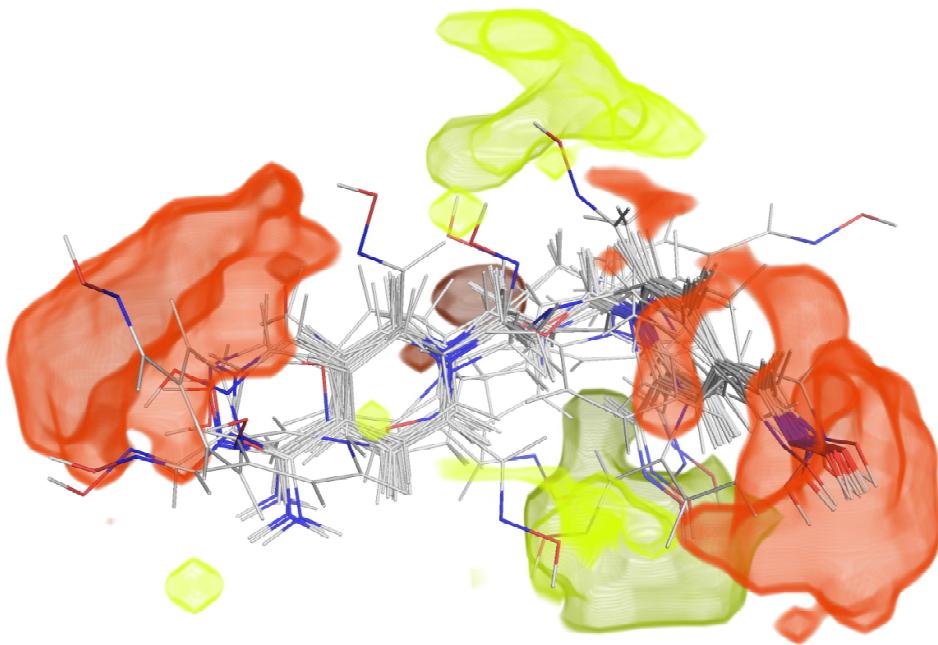


Figure 14. A color contour map of pseudo β -PLS x SD coefficients within VDW MIFs resulting from 3D QSAR analysis of UVE/IVE adjusted data set. Red volume represents positive pseudo β -PLS x SD coefficients (isosurface level: 0.02), while yellow symbolizes areas where these values are negative (isosurface level: -0.023).

The results suggested by the contour maps of pseudo β -PLS x SD coefficients may be summarized by several points: 1) to enhance the RP, increased ESP and VDW MIF levels are required especially around the two *p*- and *p'*-aldoxime functions in the bis-pyridinium scaffold; 2) larger molecules (i.e. bis-pyridinium aldoximes with the heteroaromatic rings separated by a linker) able to reach the red colored areas can induce improved RP; 3) substitutions in the linker between the pyridinium rings or those in *ortho* and *meta* positions are not favorable for increasing the RP; 4) according to the population of pseudo β -PLS x SD coefficients (Table 4) and individual MIF contributions (Table S10 – S11, Supporting Information), VDW MIFs influence more significantly the predicted RP than ESP MIFs; 5) as for the mechanistic interpretation of the 3D QSAR models, PLS analysis of SRD/FFD and UVE/IVE adjusted data sets leads to very similar results.

Table 4 Statistical characteristics of pseudo β -PLS x SD coefficients.

Characteristics	β -PLS x SD – SRD/FFD		β -PLS x SD – UVE/IVE	
	ESP MIF	VDW MIF	ESP MIF	VDW MIF
Maximum	2.6545E-02	1.0785E+00	2.0833E-02	1.3818E+00
Minimum	-5.1200E-03	-6.8686E-01	-4.5805E-03	-1.2195E+00
Mean	3.6139E-04	3.1723E-03	5.3282E-04	1.5053E-03
Total	37.5912	329.9733	55.4220	156.5720
Mean (+) ^a	4.1672E-03	9.5284E-02	1.0240E-03	1.2703E-01
Mean (-)	-3.6162E-04	-6.2710E-02	-1.8658E-04	-9.8212E-02

Total (+)	39.2297	994.5786	56.6948	775.5017
Total (-)	-1.6385	-664.6053	-1.2728	-618.9297

^a"+" or "-" symbolizes that the calculation was performed only for positive or negative values of pseudo β -PLS x SD coefficients.

4. Conclusion

The aim of this study was: 1) to develop a 3D QSAR model for reactivation of VX-inhibited *r*AChE based on a 17-member set of mono-/bis-pyridinium aldoxime reactivators and 2) to predict reactivation potencies of 12 other reactivators belonging to the same chemical space. To obtain the alignment molecular set, a QMD ligand conformer database was generated and screened to find the top scored superimposition molecular ensemble. As templates, the three most active reactivators (10, 15, 17) were examined. Eventually, the optimum alignment set superimposed on a conformer of compound 17 was used to calculate van der Waals (molecular mechanics level) and electrostatic potential (B3LYP/6-31G* quantum mechanics level) MIFs. Advanced quantum chemical calculations were performed in order to introduce higher accuracy into the 3D QSAR study. The alignment molecular set based on flexible molecular docking was not preferred due to high RMSD (145.76 Å) and low O3A score (459.01) though the promising results of molecular docking itself encourage further structure-based research. Both calculated MIFs were pre-processed by three data noise reduction methods and analyzed by regular PLS. The most statistically significant 3D QSAR models were achieved by SRD/FFD ($R^2 = 0.9989$, $Q^2_{LOO} = 0.9090$, $Q^2_{LTO} = 0.8921$, $Q^2_{LMO(20\%)} = 0.8853$, $R^2_{ext} = 0.9259$, $SDEP_{ext} = 6.8938$) and UVE/IVE ($R^2 = 0.9962$, $Q^2_{LOO} = 0.9368$, $Q^2_{LTO} = 0.9298$, $Q^2_{LMO(20\%)} = 0.9248$, $R^2_{ext} = 0.8905$, $SDEP_{ext} = 6.6756$) techniques utilizing only 4 LV. In Y-scrambling, the SRD/FFD as well as UVE/IVE-based 3D QSAR model showed good stability - sensitivity ratio after introducing perturbation in the RPs. The construction of 3D QSAR models was described in detail to ensure reproducibility of the results. In Supplementary material the alignment molecular set, contour maps of pseudo β -PLS x SD coefficients, input command file for Open3DQSAR program and the data file containing the MIFs can be found. Among the 12 reactivators with unknown RP, these two compounds achieved the highest predicted activities: 24 (SRD/FFD – 4LV model: 79.3%) and 25 (UVE/IVE – LV4 model: 84.0%). Though neither of these predicted RPs has overwhelmed the nominal experimental activity of compounds 17 (85%) or 10 (87%) present in the training set, due to uncertainty in calculated as well as observed values compounds 24 and 25 may still be considered potentially very efficient. Analyses of pseudo β -PLS x SD coefficients suggest that the RPs are significantly influenced by increased ESP and VDW MIFs around the *p*-aldoxime functional groups in the bis-pyridinium scaffold. Being in good agreement with qualitative ideas on the reactivation process, mechanistic interpretation of the obtained 3D contour maps enabled a deeper understanding of the structure-activity relationship of mono-/bis-pyridinium aldoximes. Comparing the present study to the results provided by Esposito et al. [36], we did not recognize the presence of quaternary nitrogen as a significant molecular feature. Thus, we propose a follow-up *in silico* analysis of non-quaternary aldoximes for further improvement of the RP of OP-AChE reactivators. The 3D QSAR models developed in the study showed excellent performance and may be easily utilized in ligand-based virtual screening focused on a broader range of compounds.

5. Acknowledgement

The work was supported by Ministry of Health project (NT14270), Post-doctoral project (No. CZ.1.07/2.3.00/30.0044), Ministry of Defence of the Czech Republic (Long Term Development plan – 1011), Ministry of Education, Youth and Sports (No. LH13009) and by MH CZ - DRO (University Hospital Hradec Kralove, No. 00179906). This paper has been elaborated in the framework of the project Opportunity for young researchers, reg. no. CZ.1.07/2.3.00/30.0016, supported by Operational Programme Education for Competitiveness and co-financed by the European Social Fund and the state budget of the Czech Republic. The access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the program "Projects of Large Infrastructure for Research, Development, and Innovations" (LM2010005) is highly appreciated. The authors are very grateful to Paolo Tosco, PhD. (Università degli Studi di Torino, Italy) for his assistance in 3D QSAR calculations. Many thanks are also due to Ian McColl, PhD. who improved the English of the article and clarified its text.

6. References

-
- [1] K. Starke, Presynaptic receptors, *Annu. Rev. Pharmacol. Toxicol.* 21 (1981) 7-30. DOI: <http://dx.doi.org/10.1146/annurev.pa.21.040181.000255>.
 - [2] M. Moshiri, E. Darchini-Maragheh, M. Balali-Mood, Advances in toxicology and medical treatment of chemical warfare nerve agents, *Daru* 20 (2012) 81. DOI: <http://dx.doi.org/10.1186/2008-2231-20-81>.
 - [3] A. Watson, D. Opresko, R. Young, V. Hauschild, J. King, K. Bakshi, Organophosphate Nerve Agents. In *Handbook of Toxicology of Chemical Warfare Agents*; Gupta, R. C., Ed.; Academic Press Inc: London, 2009, Chapter 6.
 - [4] J. Bajgar, J. Fusek, K. Kuca, L. Bartosova, D. Jun, Treatment of organophosphate intoxication using cholinesterase reactivators: facts and fiction, *Mini Rev. Med. Chem.* 7 (2007) 461-466. DOI: <http://dx.doi.org/10.2174/138955707780619581>.
 - [5] N. Munro, Toxicity of the organophosphate chemical warfare agents GA, GB, and VX: implications for public protection, *Environ. Health. Perspec.* 102 (1994) 18-38.
 - [6] M. Nagao, T. Takatori, Y. Matsuda, M. Nakajima, H. Iwase, K. Iwadate, Definitive evidence for the acute sarin poisoning diagnosis in the Tokyo subway, *Toxicol. Appl. Pharmacol.* 144 (1997) 198-203. DOI: <http://dx.doi.org/10.1006/taap.1997.8110>.
 - [7] G.A. Meneely, C.R. Wyttenbach, Effects of the organophosphate insecticides diazinon and parathion on bobwhite quail embryos: skeletal defects and acetylcholinesterase activity, *J. Exp. Zool.* 252 (1989) 60-70. DOI: <http://dx.doi.org/10.1002/jez.1402520109>.
 - [8] A. Nordberg, A.L. Svensson, Cholinesterase inhibitors in the treatment of Alzheimer's disease, *Drug Saf.* 19 (1998) 465-480. DOI: <http://dx.doi.org/10.2165/000002018-199819060-00004>.
 - [9] J. Korabecny, R. Dolezal, P. Cabelova, A. Horova, E. Hruba, J. Ricny, L. Sedlacek, E. Nepovimova, K. Spilovska, M. Andrs, K. Musilek, V. Opletalova, V. Sepsova, D. Ripova, K. Kuca, 7-MEOTA-donepezil like compounds as cholinesterase inhibitors: Synthesis, pharmacological evaluation, molecular modeling and QSAR studies, *Eur. J. Med. Chem.* 82 (2014) 426-438. DOI: <http://dx.doi.org/10.1016/j.ejmech.2014.05.066>.
 - [10] P. Ambure, K. Roy, Advances in quantitative structure-relationship models of anti-Alzheimer's agents, *Expert Opin. Drug. Discov.* 9 (2014) 697-723. DOI: <http://dx.doi.org/10.1517/17460441.2014.909404>.
 - [11] I.G. McKeith, D.W. Dickson, J. Lowe, M. Emre, J.T. O'Brien, H. Feldman, J. Cummings, J.E. Duda, C. Lippa, E.K. Perry, D. Aarsland, H. Arai, C.G. Ballard, B. Boeve, D.J. Burn, D. Costa, T. Del Ser, B. Dubois, D. Galasko, S. Gauthier, C.G. Goetz, E. Gomez-Tortosa, G. Halliday, L.A. Hansen, J. Hardy, T. Iwatsubo, R.N. Kalaria, D. Kaufer, R.A. Kenny, A. Korczyn, K. Kosaka, V.M. Lee, A. Lees, I. Litvan, E. Londos, O.L. Lopez, S. Minoshima, Y. Mizuno, J.A. Molina, E.B. Mukaetova-Ladinska, F. Pasquier, R.H. Perry, J.B. Schulz, J.Q. Trojanowski, M. Yamada, D.L.B. Consortium on, Diagnosis and management of dementia with Lewy

- bodies: third report of the DLB Consortium, *Neurology* 65 (2005) 1863-1872. DOI: <http://dx.doi.org/10.1212/01.wnl.0000187889.17253.b1>.
- [12] L. Maggi, R. Mantegazza, Treatment of myasthenia gravis: focus on pyridostigmine, *Clin Drug Investig* 31 (2011) 691-701. DOI: <http://dx.doi.org/10.2165/11593300-00000000-0000>.
- [13] M.C. de Koning, M. van Grol, D. Noort, Peripheral site ligand conjugation to a non-quaternary oxime enhances reactivation of nerve agent-inhibited human acetylcholinesterase, *Toxicol. Lett.* 206 (2011) 54-59. DOI: <http://dx.doi.org/10.1016/j.toxlet.2011.04.004>.
- [14] F. Worek, H. Thiermann, L. Szinicz, P. Eyer, Kinetic analysis of interactions between human acetylcholinesterase, structurally different organophosphorus compounds and oximes, *Biochem. Pharmacol.* 68 (2004) 2237-2248. DOI: <http://dx.doi.org/10.1016/j.bcp.2004.07.038>.
- [15] M. Jokanovic, M. Prostran, Pyridinium oximes as cholinesterase reactivators. Structure-activity relationship and efficacy in the treatment of poisoning with organophosphorus compounds, *Curr. Med. Chem.* 16 (2009) 2177-2188. DOI: <http://dx.doi.org/10.2174/092986709788612729>.
- [16] E. Carletti, H. Li, B. Li, F. Ekstrom, Y. Nicolet, M. Loiodice, E. Gillon, M.T. Froment, O. Lockridge, L.M. Schopfer, P. Masson, F. Nachon, Aging of cholinesterases phosphorylated by tabun proceeds through O-dealkylation, *J. Am. Chem. Soc.* 130 (2008) 16011-16020. DOI: <http://dx.doi.org/10.1021/ja804941z>.
- [17] R.M. Dawson, Review of oximes available for treatment of nerve agent poisoning, *J. Appl. Toxicol.* 14 (1994) 317-331. DOI: <http://dx.doi.org/10.1002/jat.2550140502>.
- [18] G. Mercey, T. Verdelet, J. Renou, M. Kliachyna, R. Baati, F. Nachon, L. Jean, P.Y. Renard, Reactivators of acetylcholinesterase inhibited by organophosphorus nerve agents, *Acc. Chem. Res.* 45 (2012) 756-766. DOI: <http://dx.doi.org/10.1021/ar2002864>.
- [19] J. Korabecny, O. Soukup, R. Dolezal, K. Spilovska, E. Nepovimova, M. Andrs, T. Duong Nguyen, D. Jun, K. Musilek, M. Kucerova-Chlupacova, From Pyridinium-based to Centrally Active Acetylcholinesterase Reactivators, *Mini Rev. Med. Chem.* 14 (2014) 215-221.
- [20] D.E. Lorke, G.A. Petroianu, Minireview: does in-vitro testing of oximes help predict their in-vivo action after paraoxon exposure?, *J. Appl. Toxicol.* 29 (2009) 459-469. DOI: <http://dx.doi.org/10.1002/jat.1457>.
- [21] G.F. Makhaeva, E.V. Radchenko, I.I. Baskin, V.A. Palyulin, R.J. Richardsonc, N.S. Zefirov, Combined QSAR studies of inhibitor properties of O-phosphorylated oximes toward serine esterases involved in neurotoxicity, drug metabolism and Alzheimer's disease, *SAR QSAR Environ. Res.* 23 (2012) 627-647. DOI: <http://dx.doi.org/10.1080/1062936X.2012.679690>.
- [22] K. Musilek, M. Komloova, O. Holas, A. Horova, M. Pohanka, F. Gunn-Moore, V. Dohnal, M. Dolezal, K. Kuca, Mono-oxime bisquaternary acetylcholinesterase reactivators with prop-1,3-diyl linkage-Preparation, in vitro screening and molecular docking, *Bioorg. Med. Chem.* 19 (2011) 754-762. DOI: <http://dx.doi.org/10.1016/j.bmc.2010.12.021>.
- [23] O. Holas, K. Musilek, A. Horova, V. Opletalova, K. Kuca, SAR Study on Reactivators of Ethyl-Paraoxon Inhibited Acetylcholinesterase, *Lett. Drug Des. Discov.* 9 (2012) 587-594. DOI: <http://dx.doi.org/10.2174/157018012800673083>.
- [24] F. Worek, T. Wille, M. Koller, H. Thiermann, Reactivation kinetics of a series of related bispyridinium oximes with organophosphate-inhibited human acetylcholinesterase-Structure-activity relationships, *Biochem. Pharmacol.* 83 (2012) 1700-1706. DOI: <http://dx.doi.org/10.1016/j.bcp.2012.03.002>.
- [25] K. Musilek, M. Dolezal, F. Gunn-Moore, K. Kuca, Design, evaluation and structure-activity relationship studies of the AChE reactivators against organophosphorus pesticides, *Med. Res. Rev.* 31 (2011) 548-575. DOI: <http://dx.doi.org/10.1002/med.20192>.
- [26] N. Aurbek, N.M. Herkert, M. Koller, H. Thiermann, F. Worek, Kinetic analysis of interactions of different sarin and tabun analogues with human acetylcholinesterase and oximes: is there a structure-activity relationship?, *Chem., Biol., Interact.* 187 (2010) 215-219. DOI: <http://dx.doi.org/10.1016/j.cbi.2010.01.035>.
- [27] K. Kuca, V. Racakova, D. Jun, J. Bajgar, Structure-activity relationship of acetylcholinesterase reactivators-antidotes against nerve agents, *Lett. Org. Chem.* 4 (2007) 212-217. DOI: <http://dx.doi.org/10.2174/157017807780737282>.
- [28] A.P. Guimaraes, T.C.C. Franca, T.C. Ramalho, M.N. Renno, E.F.F. da Cunha, K.S. Matos, D.T. Mancini, K. Kuca, Docking studies and effects of syn-anti isomery of oximes derived from pyridine imidazol bicycled systems as potential human acetylcholinesterase reactivators, *J. Appl. Biomed.* 9 (2011) 163-171. DOI: <http://dx.doi.org/10.2478/v10136-009-0037-1>.

- [29] T.C. Ramalho, C.C. Tanos, M.N. Renno, A.P. Guimaraes, E.F. da Cunha, K. Kuca, Development of new acetylcholinesterase reactivators: molecular modeling versus *in vitro* data, *Chem. Biol. Interact.* 187 (2010) 436-440. DOI: <http://dx.doi.org/10.1016/j.cbi.2010.05.016>.
- [30] A.K. Bhattacharjee, K. Kuca, K. Musilek, R.K. Gordon, *In silico* pharmacophore model for tabun-inhibited acetylcholinesterase reactivators: a study of their stereoelectronic properties, *Chem. Res. Toxicol.* 23 (2010) 26-36. DOI: <http://dx.doi.org/10.1021/tx900192u>.
- [31] H.M. Greenblatt, H. Dvir, I. Silman, J.L. Sussman, Acetylcholinesterase: a multifaceted target for structure-based drug design of anticholinesterase agents for the treatment of Alzheimer's disease, *J. Mol. Neurosci.* 20 (2003) 369-383. DOI: <http://dx.doi.org/10.1385/JMN:20:3:369>.
- [32] Y.P. Pang, T.M. Kollmeyer, F. Hong, J.C. Lee, P.I. Hammond, S.P. Haugabouk, S. Brimijoin, Rational design of alkylene-linked bis-pyridiniumaldoximes as improved acetylcholinesterase reactivators, *Chem. Biol.* 10 (2003) 491-502. DOI: [http://dx.doi.org/10.1016/S1074-5521\(03\)00126-1](http://dx.doi.org/10.1016/S1074-5521(03)00126-1).
- [33] M.L. Connolly, Computation of molecular volume, *J. Am. Chem. Soc.*, 107 (1985) 1118-1124. DOI: <http://dx.doi.org/10.1021/ja00291a006>.
- [34] D.M. Maxwell, I. Koplovitz, F. Worek, R.E. Sweeney, A structure-activity analysis of the variation in oxime efficacy against nerve agents, *Toxicol. Appl. Pharm.* 231(2008) 157-164. DOI: <http://dx.doi.org/10.1016/j.taap.2008.04.007>.
- [35] D.M. Maxwell, K.M. Brecht, R.E. Sweeney, A common mechanism for resistance to oxime reactivation of acetylcholinesterase inhibited by organophosphorus compounds, *Chem.-biol. Interact.* 203 (2013) 72-76. DOI: <http://dx.doi.org/10.1016/j.cbi.2012.08.024>.
- [36] E.X. Esposito, T.R. Stouch, T.R. Stouch, T. Wymore, J.D. Madura, Exploring the physicochemical properties of oxime-reactivation therapeutics for cyclosarin, sarin, tabun, and VX inactivated acetylcholinesterase, *Chem. Res. Toxicol.* 27 (2014) 99-110. DOI: <http://dx.doi.org/10.1021/tx400350b>.
- [37] V. Dohnal, K. Kuca, D. Jun, Prediction of a new broad-spectrum reactivator capable of reactivating acetylcholinesterase inhibited by nerve agents, *J. Appl. Biomed.*, 3 (2005) 139-145.
- [38] P. Tosco, T. Balle, Open3DQSAR: a new open-source software aimed at high-throughput chemometric analysis of molecular interaction fields, *J. Mol. Model.* 17 (2011) 201-208. DOI: <http://dx.doi.org/10.1007/s00894-010-0684-x>.
- [39] K. Kuca, K. Musilek, D. Jun, J. Karasova, O. Soukup, J. Pejchal, M. Hrabinova, Structure-Activity Relationship for the Reactivators of Acetylcholinesterase Inhibited by Nerve Agent VX, *Med. Chem.* 9 (2013) 689-693.
- [40] J. Bajgar, J. Fusek, J. Kassa, D. Jun, K. Kuca, P. Hajek, An attempt to assess functionally minimal acetylcholinesterase activity necessary for survival of rats intoxicated with nerve agents, *Chem. Biol. Interact.* 175 (2008) 281-285. DOI: <http://dx.doi.org/10.1016/j.cbi.2008.05.015>.
- [41] HyperChem(TM) Professional 7.52, Hypercube, Inc., 1115 NW 4th Street, Gainesville, Florida 32601, USA.
- [42] R.G. Parr, Density functional theory, *Annu. Rev. Phys. Chem.* 34 (1983) 631-656. DOI: <http://dx.doi.org/10.1146/annurev.pc.34.100183.003215>.
- [43] Gaussian 09, Revision C.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2010.
- [44] The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.
- [45] P. Tosco, T. Balle, F. Shiri, Open3DALIGN: an open-source software aimed at unsupervised ligand alignment, *J. Comput. Aided. Mol. Des.* 25 (2011) 777-783. DOI: <http://dx.doi.org/10.1007/s10822-011-9462-9>.
- [46] TINKER—Software tools for molecular design, version 5.1; <http://dasher.wustl.edu/tinker/>. Accessed 13th of June 2013.
- [47] A.R. Leach, V.J. Gillet, R.A. Lewis, R. Taylor, Three-dimensional pharmacophore methods in drug discovery, *J. Med. Chem.* 53 (2010) 539-558. DOI: <http://dx.doi.org/10.1021/jm900817u>.

- [48] J. Taminau, G. Thijs, H. De Winter, Pharao: pharmacophore alignment and optimization, *J. Mol. Graph. Model.* 27 (2008) 161-169. DOI: <http://dx.doi.org/10.1016/j.jmgm.2008.04.003>.
- [49] N.J. Richmond, P. Willett, R.D. Clark, Alignment of three-dimensional molecules using an image recognition algorithm, *J. Mol. Graph. Model.* 23 (2004) 199-209. DOI: <http://dx.doi.org/10.1016/j.jmgm.2004.04.004>.
- [50] <http://sourceforge.net/projects/open3dalignt/files/source/>
- [51] R. Jonker, A. Volgenant, A shortest augmenting path algorithm for dense and sparse linear assignment problems, *Computing* 38 (1987) 325-340. DOI: <http://dx.doi.org/10.1007/BF02278710>.
- [52] T.A. Halgren, Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions, *J. Comput. Chem.* 17 (1996) 520-552. DOI: [http://dx.doi.org/10.1002/\(SICI\)1096-987X\(199604\)17:5/6<520::AID-JCC2>3.0.CO;2-W](http://dx.doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<520::AID-JCC2>3.0.CO;2-W).
- [53] M. Baroni, G. Costantino, G. Cruciani, D. Riganelli, R. Valigi, S. Clementi, Generating Optimal Linear PLS Estimations (GOLPE): An Advanced Chemometric Tool for Handling 3D-QSAR Problems, *Quant. Struct.-Act. Relat.* 12 (1993) 9-20. DOI: <http://dx.doi.org/10.1002/qsar.19930120103>.
- [54] M.A. Kastenholz, M. Pastor, G. Cruciani, E.E. Haaksma, T. Fox, GRID/CPCA: a new computational tool to design selective ligands, *J. Med. Chem.* 43 (2000) 3033-3044. DOI: <http://dx.doi.org/10.1021/jm000934y>.
- [55] S. Wold, M. Sjostrom, L. Eriksson, PLS-regression: a basic tool of chemometrics, *Chemometr. Intell. Lab.* 58 (2001) 109-130. DOI: [http://dx.doi.org/10.1016/S0169-7439\(01\)00155-1](http://dx.doi.org/10.1016/S0169-7439(01)00155-1).
- [56] J. Verma, V.M. Khedkar, E.C. Coutinho, 3D-QSAR in drug design--a review, *Curr. Top. Med. Chem.* 10 (2010) 95-115. DOI: <http://dx.doi.org/10.2174/156802610790232260>.
- [57] R.D. Cramer III, Partial least squares (PLS): its strengths and limitations, *Perspectives in Drug Discovery and Design* 1 (1993) 269-278. DOI: <http://dx.doi.org/10.1007/BF02174528>.
- [58] P. Tosco, T. Balle, A 3D-QSAR-driven approach to binding mode and affinity prediction, *J. Chem. Inf. Model.* 52 (2012) 302-307. DOI: <http://dx.doi.org/10.1021/ci200411s>.
- [59] M. Pastor, G. Cruciani, S. Clementi, Smart region definition: a new way to improve the predictive ability and interpretability of three-dimensional quantitative structure-activity relationships, *J. Med. Chem.* 40 (1997) 1455-1464. DOI: <http://dx.doi.org/10.1021/jm9608016>.
- [60] M. Baroni, S. Clementi, G. Cruciani, G. Costantino, D. Riganelli, E. Oberrauch, Predictive ability of regression models. Part II: Selection of the best predictive PLS model, *J. Chemometr.* 6 (1992) 347-356. DOI: <http://dx.doi.org/10.1002/cem.1180060605>.
- [61] G.E.P. Box, J.S. Hunter, The 2 k-p fractional factorial designs, *Technometrics* 3 (1961) 311-351.
- [62] V. Centner, D.L. Massart, O.E. de Noord, S. de Jong, B.M. Vandeginste, C. Sterna, Elimination of uninformative variables for multivariate calibration, *Anal. Chem.* 68 (1996) 3851-3858. DOI: <http://dx.doi.org/10.1021/ac960321m>.
- [63] R. Gieleciak, J. Polanski, Modeling robust QSAR. 2. iterative variable elimination schemes for CoMSA: application for modeling benzoic acid pKa values, *J. Chem. Inf. Model.* 47 (2007) 547-556. DOI: <http://dx.doi.org/10.1021/ci600295z>.
- [64] R. Grohmann, T. Schindler, Toward robust QSPR models: Synergistic utilization of robust regression and variable elimination, *J. Comput. Chem.* 29 (2008) 847-860. DOI: <http://dx.doi.org/10.1002/jcc.20831>.
- [65] V. Consonni, D. Ballabio, R. Todeschini, Evaluation of model predictive ability by external validation techniques, *J. Chemometr.* 24 (2010) 194-201. DOI: <http://dx.doi.org/10.1002/cem.1290>.
- [66] R.D. Clark, P.C. Fox, Statistical variation in progressive scrambling, *J. Comput. Aided Mol. Des.* 18 (2004) 563-576. DOI: <http://dx.doi.org/10.1007/s10822-004-4077-z>.
- [67] A.C. Cameron, F.A.G. Windmeijer, An R-squared measure of goodness of fit for some common nonlinear regression models, *J. Econometrics* 77 (1997) 329-342. DOI: [http://dx.doi.org/10.1016/S0304-4076\(96\)01818-0](http://dx.doi.org/10.1016/S0304-4076(96)01818-0).
- [68] G. Tóth, Z. Bodai, K. Héberger, Estimation of influential points in any dat set from coefficient of determination and its leave-one-out cross-validated counterpart, *J. Comput. Aid. Mol. Des.* 27 (2013) 837-844. DOI: <http://dx.doi.org/10.1007/s10822-013-9680-4>.
- [69] A. Golbraikh, A. Tropsha, Beware of $q^2!$, *J. Mol. Graph. Model.* 20 (2002) 269-276. DOI: [http://dx.doi.org/10.1016/S1093-3263\(01\)00123-1](http://dx.doi.org/10.1016/S1093-3263(01)00123-1).
- [70] A. Cherkasov, E.N. Muratov, D. Fourches, A. Varnek, I.I. Baskin, M. Cronin, J. Dearden, P. Gramatica, Y.C. Martin, R. Todeschini, V. Consonni, V.E. Kuz'min, R. Cramer, R. Benigni, C. Yang, J. Rathman, L. Terfloth, J. Gasteiger, A. Richard, A. Tropsha, QSAR modeling: where have you been? Where are you going to?, *J. Med. Chem.* 57 (2014) 4977-5010. DOI: <http://dx.doi.org/10.1021/jm4004285>.

-
- [71] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.* 31 (2010) 455-461. DOI: <http://dx.doi.org/10.1002/jcc.21334>.
 - [72] M. Kliachyna, G. Santoni, V. Nussbaum, J. Renou, B. Sanson, J.-P. Colletier, M. Arboléas, M. Loiodice, M. Weik, L. Jean, P.-Y. Renard, F. Nachon, R. Baati: Design, synthesis and biological evaluation of novel tetrahydroacridine pyridine- aldoxime and -amidoxime hybrids as efficient uncharged reactivators of nerve agent-inhibited human acetylcholinesterase, *Eur. J. Med. Chem.* 78 (2014) 455-467. DOI: <http://dx.doi.org/10.1016/j.ejmech.2014.03.044>.
 - [73] P.I. Hammond, C. Kern, F. Hong, T.M. Kollmeyer, Y.P. Pang, S. Brimijoin, Cholinesterase Reactivation in Vivo with a Novel Bis-Oxime Optimized by Computer-Aided Design, *J. Pharmacol. Exp. Ther.* 307 (2003) 190-196. DOI: <http://dx.doi.org/10.1124/jpet.103.053405>.
 - [74] H. Zeng, H. Zhang, Combined 3D-QSAR modeling and molecular docking study on 1,4-dihydroindeno[1,2-c]pyrazoles as VEGFR-2 kinase inhibitors, *J. Mol. Graph. Model.* 29 (2010) 54-71. DOI: <http://dx.doi.org/10.1016/j.jmgm.2010.04.004>.

